

Cristian Ciurea
Paul Pocatilu
Florin Gheorghe Filip *Editors*



Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)

Recent Results in Education, Research,
and Applications



Springer

Smart Innovation, Systems and Technologies

Volume 426

Series Editors

Robert J. Howlett, KES International, Shoreham-by-Sea, UK

Lakhmi C. Jain, KES International, Shoreham-by-Sea, UK

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

Cristian Ciurea · Paul Pocatilu ·
Florin Gheorghe Filip
Editors

Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)

Recent Results in Education, Research,
and Applications

 Springer

Editors

Cristian Ciurea
Faculty of Economic Cybernetics, Statistics
and Informatics
Bucharest University of Economic Studies
Bucharest, Romania

Paul Pocatilu
Faculty of Economic Cybernetics, Statistics
and Informatics
Bucharest University of Economic Studies
Bucharest, Romania

Florin Gheorghe Filip
Romanian Academy
Bucharest, Romania

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-96-0160-8 ISBN 978-981-96-0161-5 (eBook)
<https://doi.org/10.1007/978-981-96-0161-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Organization

The IE 2024 conference was organized and hosted at the West University of Timișoara, Romania, on 23rd–24th May, 2024.

Conference Chair

Acad. Florin-Gheorghe Filip, Romanian Academy, Romania

Honorary Chair

Prof. Lakhmi C. Jain, KES International, UK

Keynote Speakers

Prof. Jan W. Owsinski, System Research Institute, Polish Academy of Sciences, Poland

Prof. Ulrich Norbistrath, University of Tartu, Estonia

Organizing Committee

Prof. Cristian Ciurea, Bucharest University of Economic Studies, Romania

Prof. Claudiu Brândaș, West University of Timișoara, Romania

Prof. Paul Pocatilu, Bucharest University of Economic Studies, Romania

Assoc. Prof. Mihai Doinea, Bucharest University of Economic Studies, Romania

Assoc. Prof. Alin Zamfiroiu, Bucharest University of Economic Studies, Romania
Prof. Cătălin Boja, Bucharest University of Economic Studies, Romania
Prof. Viorel Negrut, West University of Timișoara, Romania
Assoc. Prof. Otniel Didraga, West University of Timișoara, Romania
Assist. Prof. Ionuț Nica, Bucharest University of Economic Studies, Romania
Assist. Prof. Tiberiu Georgescu, Bucharest University of Economic Studies, Romania

Program Committee

Prof. Anda Belciu, Bucharest University of Economic Studies, Romania
Prof. Constanta-Nicoleta Bodea, Bucharest University of Economic Studies, Romania
Prof. Ana Ramona Bologa, Bucharest University of Economic Studies, Romania
Asst. Antonio Clim, Bucharest University of Economic Studies, Romania
Prof. Catalina Lucia Cocianu, Bucharest University of Economic Studies, Romania
Assoc. Prof. Alexandra Corbea, Bucharest University of Economic Studies, Romania
Prof. Liviu-Adrian Cotfas, Bucharest University of Economic Studies, Romania
Prof. Adriana Dârdală, Bucharest University of Economic Studies, Romania
Prof. Camelia Delcea, Bucharest University of Economic Studies, Romania
Prof. Vlad Diaconita, Bucharest University of Economic Studies, Romania
Assoc. Prof. Irina Georgescu, Bucharest University of Economic Studies, Romania
Assoc. Prof. Sorina Gramatovici, Bucharest University of Economic Studies, Romania
Assoc. Prof. Monica Maer-Matei, Bucharest University of Economic Studies, Romania
Prof. Ioana Manafi, Bucharest University of Economic Studies, Romania
Prof. Marinela Mircea, Bucharest University of Economic Studies, Romania
Assoc. Prof. Radu Ioan Mogos, Bucharest University of Economic Studies, Romania
Prof. Nora Chiriță, Bucharest University of Economic Studies, Romania
Assoc. Prof. Simona Oprea, Bucharest University of Economic Studies, Romania
Assoc. Prof. Lorena Bătăgan, Bucharest University of Economic Studies, Romania
Prof. Mădălina Ecaterina Popescu, Bucharest University of Economic Studies, Romania
Prof. Mihai Daniel Roman, Bucharest University of Economic Studies, Romania
Assoc. Prof. Iuliana Simonca, Bucharest University of Economic Studies, Romania
Prof. Marian Stoica, Bucharest University of Economic Studies, Romania
Prof. Răzvan Zota, Bucharest University of Economic Studies, Romania
Asst. Prof. Mădălina Zurini, Bucharest University of Economic Studies, Romania

Preface

The volume contains a selection of papers presented at the 23rd *International Conference on Informatics in Economy* (IE 2024) entitled “Recent Results in Education, Research, and Applications,” which was held in Timișoara, Romania, in a hybrid format, on 23–24 May 2024.

The IE conference has a quite long tradition. Its first edition was organized in 1993 by the *Bucharest University of Economic Studies* (ASE) in collaboration with academics of the *Institut National des Sciences Appliquées* (INSA) de Lyon, France, as an event meant to stimulate the exchange of ideas and scientific results in the context of new evolutions in Romania. While, in the period 1993–2011, the conference has been initially organized in odd years since 2012, IE has become an annual event.

The International Conference on Informatics in Economy was, in the early 90s, one of the first scientific events in this thematic domain ever held in Romania and, over years, has gained an international scientific recognition. At the national level, it remains one of the most important scientific events that gather the entire Romanian Economic Informatics community. Many of the papers presented at the conference have been quality works and, consequently, the volumes of the 11th–16th editions were indexed by ISI Thomson Reuters as ISI Proceedings.

In 2021, 2022, and 2023, Springer accepted to publish the proceedings of the IE 2021, IE 2022, and IE 2023 conferences under the title *Education, Research and Business Technologies* in the book series, named *Smart Innovation, Systems and Technologies*, led by professors Robert J. Howlett and Lakhmi C. Jain.

The 23rd edition of the conference promoted research results in Business Informatics and related Computer Science topics such as:

- Machine Learning Theory and Applications
- Big Data Management, Processing and Analytics
- Smart Cities and Sustainable Communities
- IoT, Mobile and Multimedia Solutions
- Cloud, Distributed and Parallel Computing
- Cybersecurity and Critical Infrastructures

- Quantitative Economics
- Digital Business and e-Transformation
- IT Deployment in Cultural Institutions

The conference program included two valuable keynote speeches, and 60 regular papers grouped into 5 parts. Many papers reflected the international cooperation research works and were authored by researchers from 12 countries (Bulgaria, Croatia, Estonia, Germany, India, Israel, Moldova, Poland, Romania, Slovenia, Switzerland, and Turkey).

From the total number of regular manuscripts submitted for evaluation, a number of 39 papers were accepted by the *International Programme Committee*.

A number of 39 final versions of the accepted papers are included in the 5 parts of the current volume:

1. Big Data Management, Processing and Analytics, Smart Cities and Sustainable Communities (8 papers)
2. Digital Business and e-Transformation, Cloud, Distributed and Parallel Computing, IT Deployment in Cultural Institutions (6 papers)
3. IoT, Mobile and Multimedia Solutions, Cybersecurity and Critical Infrastructures (7 papers)
4. Quantitative Economics (10 papers)
5. Machine Learning Theory and Applications (8 papers)

The study of the papers contained in the volume can be useful for researchers, consultants, and postgraduate students in *Computer science and design, Business informatics, Economics, and Management*.

The editors are grateful to Springer for accepting to publish the proceedings of IE conference in the *Smart Innovation, Systems and Technologies* series, and ensuring a high-quality technical version of the volume. Special thanks are due to *Prof. Lakhmi C. Jain* for permanent guidance and the technical assistance of colleagues from Springer in the book production is also acknowledged.

Bucharest, Romania
June 2024

Cristian Ciurea
Paul Pocatilu
Florin Gheorghe Filip

Contents

Part I Big Data Management; Processing and Analytics; Smart Cities and Sustainable Communities		
1	Data Integration Techniques for Building a Comprehensive Air Quality Dataset	3
	Andreea-Mihaela Niculae and Adela Bâra	
2	The Perceived Economic Impact of Data Science Solutions on Business Performance	15
	Marian Pompiliu Cristescu, Dumitru Alexandru Mara, Lia Cornelia Culda, and Raluca Andreea Nerișanu	
3	Performance Evaluation of Data Vault and Dimensional Modeling: Insights from TPC-DS Dataset Analysis	27
	Andreea Vines	
4	Revolutionizing Healthcare: Harnessing Natural Language Processing and Big Data for Predictive Disease Diagnosis	39
	Diana Andreea Căuniac and Andreea Alexandra Cîrnaru	
5	Enhancing the Involvement Level of Volunteer Students in the Academic Community: Efficient Usage of Digital Resources	49
	Laurentiu-Gabrie Dinca and Vlad Diaconita	
6	Intelligent Models in Power Delivery Management for Economic Damage Computation	59
	Mariana Dumitrescu	
7	Smart City Parking Applications in Romania: An Analysis and Conclusions	73
	Răzvan Daniel Zota and Mihai Adrian Lungu	

8	Measurement and Verification Solution for Energy Performance Contracting in the Building Sector	83
	George Suci, Lucian Necula, Cosmina Stalidi, and Mădălin Silion	
Part II Digital Business and e-Transformation; Cloud, Distributed and Parallel Computing; IT Deployment in Cultural Institutions		
9	Analyzing the Antecedents of Deceived Buying and Disinformation Risk	97
	Luigia-Gabriela Sterie and Daniel Mican	
10	Driving Factors of Social Commerce Intention: The Role of Social Commerce Constructs, Social Influence, and Trust	107
	Diana-Ștefania Ghenie, Virgil-Vasile Șoncutean, and Dan-Andrei Sitar-Tăut	
11	Strategic Collaboration in International University Networks—Case Study	119
	Margarita Bogdanova, Evelina Parashkevova-Velikova, and Yuriy Kuznetsov	
12	Comparative Analysis of Natural Language Query Responses on BPMN Model Serializations: RDF Graphs Versus BPMN XML	129
	Damaris Naomi Dolha and Robert Andrei Buchmann	
13	A Hybrid Retrieval-Augmented Generation Approach for Heterogeneous Knowledge Bases	141
	Andrei Chiș and Ana-Maria Ghiran	
14	Digitalization in Higher Education—Designing and Implementing an Application to Facilitate the University Final Thesis Collaboration Process	153
	Cristina Iancu, Simona Vasilica Oprea, and Cristian-Eugen Ciurea	
Part III IoT, Mobile and Multimedia Solutions; Cybersecurity and Critical Infrastructures		
15	An Overlook of IoT Finance and Pre-existing Security Challenges in Financial Transactions via IoT Devices on Blockchain Networks	167
	Alisa Harkai	
16	Models for Network Traffic Behavioral Analysis in IoT Systems ...	179
	Alin Zamfiroiu, Iustin Floroiu, Daniel Savu, Lorena Bătăgan, and Anupriya Sharma Ghai	

17 Augmented Reality in Education—Prototype for the Undergraduate System in Romania 191
 Andreea-Cristina Stroe

18 Empowering Sustainability: Upcycling Smartphones as the Future of IoT and Edge Computing in Emerging Economies 203
 Ulrich Norbistrath, Perseverance Munga Ngoy, Renato Perotto Machado, Huber Flores, and Mohan Liyanage

19 Quantum Computing and Cybersecurity: Threat or Opportunity? 215
 Alexandru Butoi

20 New Real-Time Encryption Mechanism for Financial Transactions 225
 Ionuț-Eduard Matei, Costin-Radu Boldea, and Bogdan-Boldea

21 Enhanced Blockchain-Based e-Voting System Using Zero-Knowledge Proofs 237
 Marcela Lupu and Iulian Aciobăniței

Part IV Quantitative Economics

22 The Environmental Kuznets Curve for Deforestation in Romania: An ARDL-Based Evaluation 249
 Irina Georgescu, Ionuț Nica, and Nora Chiriță

23 Modeling Financial Contagion: Exploring Interconnectedness and Systemic Risk Through Agent-Based Simulation 263
 Ștefan Ionescu, George Tătaru, Gabriel Dumitrecu, Nicole Bianca Stanică, and Liviu-Adrian Cofas

24 Modeling Socio-economic Phenomena: Some Critical Thoughts and a Positive Example 277
 Jan W. Owsiański, Jarosław Stańczak, Przemysław Śleszyński, and Rafał Wiśniewski

25 Enhancing Social Acceptance in Microgrid Implementation: A Share-of-Choice Optimization Approach 289
 Emmanuel Fragniere, Francesco Moresino, Sarah Sandoz, and Nabil Abdennadher

26 Greening the Workplace for Happiness: Exploring CSR’s and Sustainable Practice’s Effects on Employees Well-being 301
 Alexandra-Nicoleta Ciucu Durnoi, Adina-Iuliana Jigani, Bianca Cibu, Andra Sandu, and Camelia Delcea

27	Digital Skills and Labor Market Prospects: A Conjoint Experiment on Students' Preferences	311
	Maer Matei Monica Mihaela, Mocanu Cristina, Zamfir Ana Maria, and Năstasă Anamaria	
28	Estimating Probability of Default. A New Qualitative Perspective	321
	Florin George Dragu	
29	Brand Associations Bibliometric Analysis	333
	Alexandra Raluca Jelea, Adriana Manolică, Teodora Roman, and Patricea Elena Bertea	
30	Community Detection on an European Airports Network	343
	Mioara Băncescu and Adina Gabriela Iliuță	
31	The Impact of Minimum Wage on Employment. The Case of Romania	353
	Georgiana Stănciulescu and Mădălina Ecaterina Popescu	
 Part V Machine Learning Theory and Applications		
32	The Analysis of the Real Estate Market Using the k-Nearest Neighbors Algorithm	365
	Ion-Florin Răducu, Stelian Stancu, Andreea Pernici, and Denisa Bălă	
33	Viability of Open Source Assistants for Adaptive Learning Systems	375
	Bobocea Andrei and Corina Marina Mirea	
34	Different Entropy Measurements in Machine Learning	385
	Dror Ben-Ami	
35	Algorithmic Trading Bots Using Artificial Intelligence	397
	Florentin Șerban and Bogdan-Petru Vrînceanu	
36	Multinomial Naïve Bayes Classifier for Romanian News in the Context of a Television Station	407
	Adrian Vintila and Constanta-Nicoleta Bodea	
37	Text Summarization and Sentiment Analysis Pipelines Using Large Language Models for Financial News	421
	Alin-Gabriel Văduva and Anca-Ioana Andreescu	
38	Develop a Sentiment Analysis for Teaching Performance Using a Machine Learning Pipeline	431
	Dimitrie-Daniel Plăcintă	

39	Are LLMs Hallucinating When Answering Specific Questions? Experiments with a RAG Pipeline	441
	Isabela Iacob and Gheorghe Cosmin Silaghi	

About the Editors

Cristian Ciurea is a Professor at the Department of Economic Informatics and Cybernetics at Bucharest University of Economic Studies. He is also the Head of department. He has graduated from the Faculty of Economic Cybernetics, Statistics and Informatics at the Bucharest University of Economic Studies in 2007. He has a master's in Informatics Project Management (2010) and a Ph.D. in Economic Informatics (2011) from the Bucharest University of Economic Studies. He has a solid background in computer science and is interested in collaborative systems-related issues. Other fields of interest include intelligent systems, software metrics, data structures, object-oriented programming, windows applications programming, mobile devices programming, and testing process automation for software quality assurance.

Paul Pocatilu graduated from the Faculty of Cybernetics, Statistics and Economic Informatics in 1998. He achieved the Ph.D. in Economics in 2003 with a thesis on Software Testing Cost Assessment Models. He has published as the author and co-author over 45 articles in journals and over 40 articles on national and international conferences. He is the author and co-author of 10 books, (Mobile Devices Programming and Software Testing Costs are 2 of them). He is a Professor at the Department of Economic Informatics and Cybernetics in the Bucharest University of Economic Studies, Bucharest. He teaches courses, seminars, and laboratories on Mobile Devices Programming, Economic Informatics, Computer Programming, and Project Quality Management to graduate and postgraduate students. His current research areas are software testing, software quality, project management, and mobile application development.

Florin Gheorghe Filip graduated in Automation at Politehnica University of Bucharest in 1970 and received his Ph.D. degree from the same university in 1982. He was elected as a Corresponding Member of the Romanian Academy in 1991 and became a Full Member of the Academy in 1999. During 2000–2010, he was the Vice-President of the Romanian Academy (elected in 2000, re-elected in 2004 and 2006). In 2010, he was the Elected President of the Information Science and

Technology section of the Academy (re-elected in 2015 and 2019). He was the Managing Director of the National Institute for R&D in Informatics-ICI, Bucharest (1991–1997). He is an Honorary Member of the Romanian Academy of Technical Sciences, and Academy of Sciences of the Republic of Moldova.

Part I
Big Data Management; Processing
and Analytics; Smart Cities
and Sustainable Communities

Chapter 1

Data Integration Techniques for Building a Comprehensive Air Quality Dataset



Andreea-Mihaela Niculae  and Adela Bâra 

Abstract In today's data-rich world, merging diverse datasets is essential for thorough analysis. This article delves into integrating air quality data, focusing on the challenges and methods involved. It presents a framework covering preprocessing steps like data cleaning, reduction, transformation, and relevance determination, leading to the fusion of varied datasets into a unified format. Using a case study of integrating static datasets from the United States, this paper offers practical insights into the process. Key considerations include data organization, consistency, spatial–temporal analysis, and metadata access. Challenges like format variations and data discrepancies are addressed systematically. The case study underscores understanding metadata, spotting redundant information, and making informed decisions for data consistency and spatial analysis. Highlighting the importance of integrated air quality datasets in research and decision-making, the article provides a comprehensive guide for researchers and practitioners involved in air quality data integration and analysis.

1.1 Introduction

We live in a data-intensive world, with numerous continuous and heterogeneous data sources. Amid this abundance of data, one topic stands out: the development of multiple IoT sensor-based approaches for collecting and monitoring different themes that yield both advantages and disadvantages [1]. With advantages such as an unceasing increase in Big Data sets and the growth of availability of their sources, knowing how to utilize them at their full potential is an important step in Big Data analysis. However, the disadvantages should not be overlooked either. The main issues include discovering which data one should use for research, as not all data

A.-M. Niculae (✉) · A. Bâra
Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania
e-mail: andreea.niculae@csie.ase.ro

is of quality, relevant, or even extensive, raising the need to properly unite different sources.

In practice, several businesses have already developed solutions that utilize the concept of integration, whether to emphasize the use of multiple services into a single end solution or to combine multiple sources into a single final dataset [2]. In this paper, however, the focus is on the latter. Using multiple sources comes with numerous benefits for analyzing the data, as we will see in this paper.

With so many available datasets at every step, the ability to select, obtain, and integrate pertinent information becomes indispensable. Indeed, data integration serves as the final missing puzzle piece required to fully harness the riches of available data [3], by combining multiple sources into a comprehensive dataset.

One such domain where comprehensive datasets are indispensable for analysts is air pollution. Air quality is continuously monitored using remote sensors, which are made available by researchers, scientists, or governments. As of January 2024 [4], the World Health Organization (WHO) records and offers data for more than 7000 settlements across 120 states, making it a valuable source of information. Integrating data from these sensors serves as a crucial starting point for numerous research articles [1, 5–8]. Some articles focus on integrating air pollution sensors with external data, such as satellites [1], mobile sensors [5], weather [7], and hazardous particles [8]. The common thread across these articles is the recognition of the importance of using integrated air quality data, as relying solely on independent measurements is insufficient [9].

The ability to utilize an integrated dataset offers numerous benefits. Focusing on the medical domain, for example, the availability of an extensive dataset about air quality enables a wide range of studies on exposure to different pollutants. An increasing number of studies [10] focus on the dangers of air pollution on overall health, its effects on the lungs, or the connection between pollution and cardiovascular diseases. This emphasis on the health effects of air pollution served as a starting point in the decision to prioritize air quality-related data integration in this paper. Our aim is to contribute to and strengthen future analyses related to pollution.

This paper aims to address an essential step in data preprocessing: the integration of data for air quality research, with a focus on integrating information from various sensors. These sensors provide data on key topics, including the Air Quality Index (AQI), pollutants (particulates or gases), weather conditions, and hazardous materials. Given the diverse nature of the collected data, it is crucial to integrate them in a manner that facilitates and enhances high-level analysis of air quality.

This paper highlights the importance of data integration in the air quality domain, focusing on the differences between integrating static and dynamic data, with emphasis on static data. The main objective is to underscore the necessity of integrating diverse data sources to enable comprehensive and high-level analysis of air quality.

The paper is structured as follows: first, the framework for integrating air quality data is presented, along with detailed steps and explanations. Next, a case study on integrating US static datasets is performed, combining multiple sets into a comprehensive dataset, ending with a map to illustrate the utility of the combined data.

Finally, a conclusion is drawn related to reiterate the importance of integrating air quality data.

The novelty of this paper lies in its approach to data integration in the air quality domain, the datasets used to demonstrate the implementation of integration, and the overall example it provides.

1.2 Integration Framework

Data integration is one of the methods contained in the data preprocessing step. Preprocessing [11] the data is frequently considered a mandatory step in various types of analyses, and its adoption has notably increased in domains such as Big Data, Machine Learning, and Data Mining. In addition to integrating heterogeneous data to obtain a single dataset for easy manipulation and use, this step encompasses several other important methods, including data cleaning, relevance determination, data reduction, and data transformation [12].

- *Data cleaning* enhances the quality of the data by employing techniques for imputing missing values, detecting outliers, and removing duplicates. Missing data cannot be directly used in machine learning models, as it may lead to biased results. Similarly, outliers can negatively influence the outcomes of models by skewing the analysis.
- *Data reduction* is used to avoid the existence of potential noise in the data, by using only a part of the original dataset. This can involve reducing the number of rows, columns, or both. Sample reduction techniques focus on selecting a representative subset of rows, while column reduction techniques involve combining attributes or retaining only relevant attributes.
- *Data transformation* means converting the original data into a format that is suitable for use in analytical models. This process may include normalization, standardization, scaling, and transforming variables from non-numerical to numerical types or from continuous to discrete forms.
- *Data Relevance Determination* identifies and selects the most relevant features or variables for the chosen analysis. It filters irrelevant or redundant information and focuses only on the most contributing features.
- *Data Integration* is the process of combining data from multiple heterogeneous sources to create a single, unified, and centralized dataset. However, integrating numerous sources can be challenging [13], particularly due to variations in data formats, measurements, files, and types. There are two types of integrations, for static and for dynamic data. Integrating static data focuses on identifying common information to map all the data together. In contrast, dynamic data integration involves the use of wrappers to translate different types of inputs into a centralized dataset, being able to adapt to changes as they occur.

While all methods included in the preprocessing step are equally important, data integration poses a greater challenge compared to the others. This is particularly clear

in the air quality domain, where diverse sensors are deployed to monitor various aspects. Figure 1.1 illustrates the four primary areas of interest in air pollution monitoring: major pollutants (particulate matter and harmful gases), weather conditions, dangerous components, and Air Quality Index (AQI, derived from specific pollutants). The AQI itself relies on data integration to identify the most harmful air pollutants and compute the index's value.

Dynamic air quality data is typically organized into the four areas of interest depicted in Fig. 1.1, with one unstructured dataset dedicated to each area. In contrast, static air quality data is usually organized into separate files, each focusing on individual components. Figure 1.2 depicts a comparison between challenges identified in integrating static and integrating dynamic datasets in the air quality domain.

Integrating dynamic air pollution data possesses two types of challenges: combining multiple sources to better monitor each pollutant, and combining multiple pollutants, weather conditions, and harmful components to improve air quality analysis. Both issues are based on understanding how data is structured and recorded for each data source (sensor).

- To obtain accurate measurements of a specific pollutant, it is necessary to integrate data from multiple sensors located in close proximity. This requires establishing

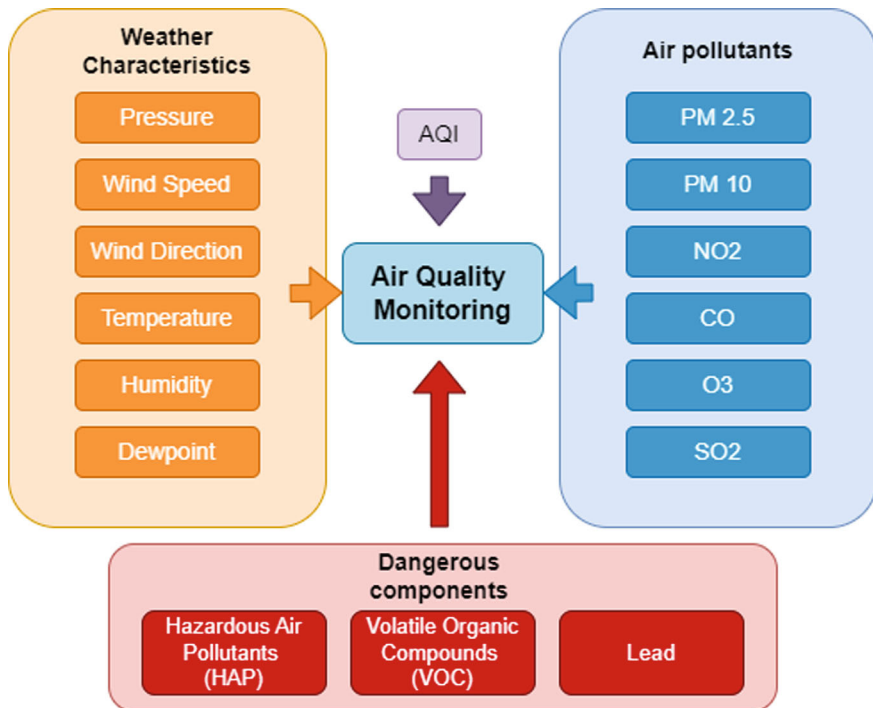


Fig. 1.1 Air quality dataset's integration for comprehensive monitoring and analysis

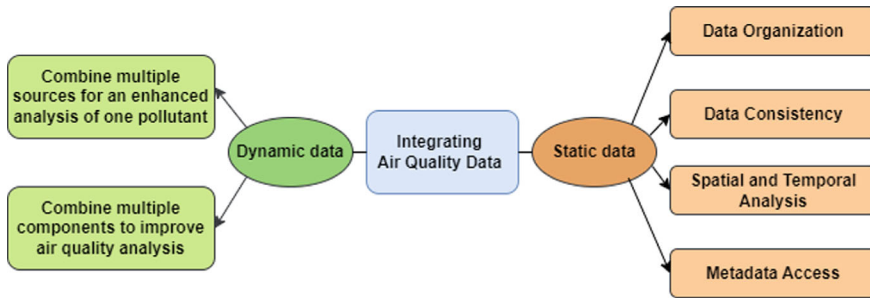


Fig. 1.2 Comparison in challenges identified when integrating static and dynamic air quality data

the spatial scope of the additional measurements needed, whether around individual sensors, within the city, county, or even at the state/country level, in order to increase accuracy by computing the mean of all measurements. Additionally, it is crucial to consider the frequency at which each sensor provides data, whether daily, every 8 h, hourly, or even every other minute. Data integration must take into account how each sensor emits and stores its data. A key consideration is whether the sensor data use consistent keywords or if a dictionary for similar keywords is required.

- Combining data from multiple components presents similar challenges regarding harmonizing keywords from sensor data. It is essential to consider the frequency of data, as some sensors may be more active (e.g., pollutant sensors recording data more frequently) than others (e.g., rain sensors transmitting data only during precipitation events). This type of integration often requires additional preprocessing methods, as a more detailed analysis is necessary to determine how each dataset should be introduced and integrated. Furthermore, multiple types of sensors are typically required, as some sensors focus solely on weather components while others specialize in monitoring specific pollutants—resulting in the need to unite disparate datasets with varying information. Managing these diverse datasets can pose significant challenges in the integration process.

Integrating *static air pollution data* presents several distinct challenges, including data organization, consistency, spatial and temporal analysis, and access to metadata. Successfully integrating static data requires a thorough understanding of how the data is stored and how one intends to integrate it.

- Data organization is the first step which encourages studying the source of the data, including its file format, database schema, or data model. Different methods are employed for each type of source. Dictionaries of keywords might be needed. Understanding these aspects is essential for ensuring proper alignment and integration.
- Data consistency is crucial for air quality analysis, particularly because it encompasses diverse components measured at potentially different times, frequencies, standards, time zones, and, importantly, units of measurement.

- Spatial analysis involves choosing the integration level: latitude-longitude, city, county, or country, to perform necessary calculations. Temporal analysis determines the relevant time period: hourly, daily, or otherwise.
- Metadata access is required when dealing with static datasets, as it enables the understanding of the characteristics and limitations of the sources. It provides insights into data sources, collection methods, and quality assurance procedures.

Table 1.1 presents an algorithm for integrating static air quality data, assuming that the initial step of analyzing data format and ensuring compatibility has already been undertaken.

Table 1.1 Algorithm for integrating static air quality data

	Integrating air quality data from several sources	Observations
1	<i>obtain a list of all documents</i>	
2	<i>determine the list of common columns to merge</i>	(geographic level: latitude-longitude, city, county, country)
3	for each document in the list:	
4	read file	
5	perform exploratory data analysis (EDA)	(optional step to visualize components structure, distribution, outliers, and others)
6	preprocess document:	[14]
7	select only relevant columns	
8	fill missing values	(usually with 0)
9	expand data for sets with multiple key information	
10	select the desired time zone / time frame	(e.g., WGS84, NAD83)
11	select the desired measuring standard	
12	group information based on the desired level	(only if latitude-longitude/city grouping is not desired)
13	merge with other documents:	(this step does nothing for the first document)
14	<i>determine dataset identifier for suffix</i>	(suffixes for coding merged data)
15	outer join current document with previous results	(add suffix for easy identification)
16	<i>check if datasets were merged correctly</i>	

1.3 Case Study on US Datasets for Air Quality Integration

Utilizing publicly available static datasets of various components used in measuring air quality, an integration was conducted for the United States of America. Having a choice between hourly data and daily data, both from 2023 until 1980, in this paper, daily data only from 2023 was selected as the basis of the integration. The reasoning behind this is that no matter the frequency, the datasets have the same structure, but considerable differences in record numbers, thus using a smaller sample and successfully integrating it will lead to inferences for the bigger sets. Hourly datasets come with the issue of slower processing, as they have millions of records for only 1 year. Furthermore, including data from multiple years would result in a slower process, as it would only increase the number of records while maintaining the same underlying structure.

The following datasets were downloaded as csv files:

- *Gases*: Ozone (coded as 44,201), SO₂ (42,401), CO (42,101), NO₂ (42,602).
- *Particulates*: PM_{2.5} FRM/FEM Mass (Federal Reference Methods and Federal Equivalent Methods for measuring PM_{2.5}—88,101), PM_{2.5} non FRM/FEM Mass (88,502), PM₁₀ Mass (86,101), PMc Mass (86,101), PM_{2.5} Speciation, PM₁₀ Speciation.
- *Meteorological*: Winds (Resultants), Temperature (62,101), Barometric Pressure (64,101), Relative Humidity (RH) and Dewpoint.
- *Toxics, Precursors, and Lead*: Hazardous Air Pollutants (HAPs), Volatile Organic Compounds (VOCs), NONOxNOy, Lead.
- *Air Quality Index (AQI)*.

Information about each component listed above was found in a separate file. The goal was to integrate all these 19 datasets into one comprehensive set. The Data Organization step is yet incomplete, as only the data format was identified. Next, it was crucial to discover similarities between all the columns in the tables to decide how to integrate each component. Fortunately, 18 of them had the same structure. The only one apart was the AQI set.

Before proceeding, it is imperative to understand the metadata, especially regarding the information contained within datasets with similar structures, as well as any contrasts present in the different dataset. Upon inspection, redundant information was identified, such as pairs where only one piece of data is necessary (*State Code—State Name; County Code—County Name; Site Number—Local Site Name—City Name; Parameter Code—Parameter Name; Method Code—Method Name*). *Latitude* and *Longitude* are only required if city-level integration is performed. *Datum* contains two key values, WGS84 (World Geodetic System) and NAD83 (North American Datum): filtering the data based on the desired standard must be performed. *Parameter Name* (or code) was studied in advance: some files contain multiple values in this column and thus should be exploded, since it is a very important column for integration. *Sample Duration* mentions information about how much air (measured in time) is needed to record the mentioned pollutant. *Pollutant Standard* contains multiple

information as well and should be filtered for further analysis: there are different standards for recording pollutants, especially for CO and SO₂. *Units of Measure* helps researchers in understanding how each pollutant is measured, aiding in standardization decisions. *Date Local* is a very important column for integration as well. *Arithmetic Mean* is the mean of all the daily measurements for the specific pollutant, one for each pair of standard and datum. *Date of Last Change* helps in further integrations, to check for consistencies in the data. While additional variables exist, these encompass the most frequently utilized information.

When assessing Data Consistency in this dataset for air quality integration, several decisions needed to be made. These include determining which variables from the pairs should be utilized in integration, selecting the preferred pollutant standard, specifying the desired standard for datum, and considering standardization of pollutants to ensure comparable units of measurement.

Lastly, integrating air quality needs a careful spatial and temporal analysis. While temporal decisions were established during the initial data extraction process, spatial analysis requires determining the desired level of integration, such as station, city, county, or state. Hourly data can be integrated at all levels, except for the AQI dataset, while daily data can be integrated across all levels without limitations. However, integrating daily data at higher levels entails additional computation to derive the mean of combined values from stations within the same city, county, or state. Evidently, by combining values the Big Data dataset will shrink considerably. It is worth noting that joining based on longitude-latitude (or city) results in fewer rows in the integrated dataset compared to county-based joining, likely due to variations in pollutants measured by different stations.

Only in 2023, as of 26 October, the last data update, there is a considerable difference in the number of records in each file.

- Among gases, ozone has the most records (more than 200 thousand rows), while CO has the least recorded days (less than 100 thousand, for two types of pollutant measuring standards).
- For particulates, PM_{2,5} FRM/FEM mass is the dataset with the most records (among all files), with more than 350 thousand rows, while PM_c mass has just 20 thousand records.
- Meteorological datasets have the most records in the temperature dataset (around 150 thousand rows), and the least in Barometric pressure (50 thousand).
- Toxics, precursors, and lead contain information about several dangerous components in the same dataset, thus making it hard to assess which is the most common. Lead dataset has only 4500 rows.

Thus, datasets lacking sufficient information might be better off excluded to prevent the integration of numerous missing values. Additionally, it is worth noting that record counts can be misleading, as they factor in both datum standards and potential variations in the number of monitoring stations for each pollutant, weather parameter, or hazardous component.

Finally, integration of the mentioned dataset was conducted using all available information. Data was filtered for the WGS84 Datum standard. Two datasets required

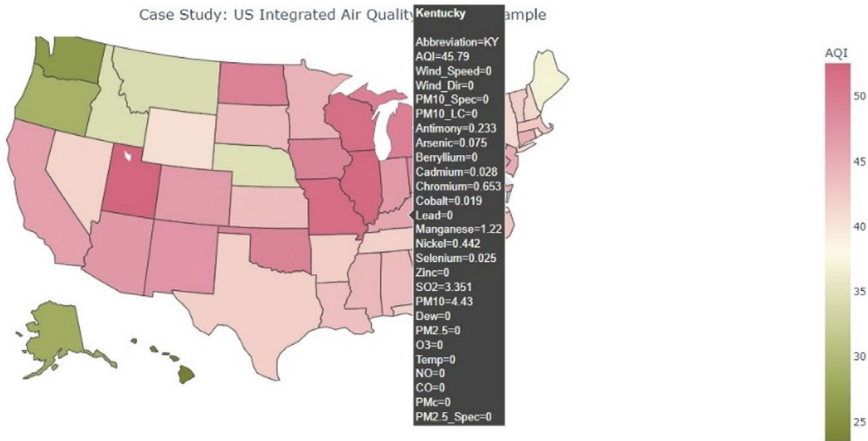


Fig. 1.3 US integrated air quality dataset example—mean values for each state

further filtering based on Pollutant Standard: SO2 for “SO2 3-h 1971 standard” and CO for “CO 8-h 1971 standard”. Following the algorithm presented in Table 1.1, various types of integrations were performed based on levels. For example, after combining all information using outer joins at the county level, the final dataset contained only 25% of counties from all existing. This signals that there are not enough monitoring stations in the US to adequately assess air quality, pollutants, or even weather characteristics. Thus, for a clearer understanding and visualization of the data integration process, the final comprehensive dataset was computed at the state level, ensuring that every desired piece of information was available.

Figure 1.3 illustrates a map generated by integrating data from all 19 CSV files at the state level, incorporating the mentioned filtering criteria and preprocessing methods. The original data was cleaned by removing duplicates and replacing missing values with 0. Data reduction was achieved by selecting only essential columns and aggregating rows to compute mean values for each state. Data transformation and relevance determination were not performed, as no analysis was conducted yet. The primary focus was on data integration. This approach aligns with the framework and objectives outlined in the paper.

1.4 Conclusion

In conclusion, this article has explored the complex process of integrating air quality data, shedding light on the challenges and methodologies involved. By presenting a comprehensive integration framework and a practical case study of integrating US datasets, this paper aims to provide valuable insights for researchers and practitioners in the field. Emphasizing the importance of data preprocessing, including cleaning,

reduction, and transformation, in this paper, the significance of understanding meta-data and ensuring data consistency was highlighted. The integration framework outlined here serves as a systematic approach to address variations in data formats and measurements. As we navigate the complexities of data integration, it becomes evident that integrated air quality datasets play a crucial role in advancing research and decision-making in public health and environmental domains. Moving forward, continued efforts in data integration will contribute to a deeper understanding of air quality dynamics and facilitate informed actions for a healthier environment.

Acknowledgements This paper was co-financed by The Bucharest University of Economic Studies during the PhD program.

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI—UEFISCDI, project number ERANET-ERAMIN-3-ValorWaste-1, within PNCDI IV.

Data Source The data used in the analysis comes from United States Environment Protection Agency (EPA), available at: https://aqs.epa.gov/aqsweb/airdata/download_files.html. It contains pre-generated files ready to download cost-free.

Date of extraction: March 2024.

References

1. Shaddick, G., Thomas, M.L., Green, A., Brauer, M., Donkelaar, A.V., Burnett, R., Chang, H.H., Cohen, A., Dingenen, R.V., Dora, C., Gumy, S., Liu, Y., Martin, R., Waller, L.A., West, J., Zidek, J.V., Pruss, A.: Data integration model for air quality: a hierarchical approach to the global estimation of exposures to ambient air pollution. *Appl. Statist.* **1**(67), 231–253 (2018)
2. Zhou, K., Fu, C., Yang, S.: Big data driven smart energy management: From big data to big insights. *Renew. Sustain. Energy Rev.* **56**, 215–225 (2016)
3. Doan, A., Halevy, A., Ives, Z.: *Principles of Data Integration*, Elsevier Inc. (2012)
4. World Health Organization: WHO Ambient Air quality database. World Health Organization, 31 January 2024. <https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database>
5. Leung, Y., Zhou, Y., Lam, K.-Y., Fung, T., Cheung, K.-Y., Kim, T., Jung, H.: Integration of air pollution data collected by mobile sensors and ground-based stations to derive a spatiotemporal air pollution profile of a city. *Int. J. Geogr. Inf. Sci. Geogr. Inf. Sci.* **33**(11), 2218–2240 (2019)
6. Integration of Data and Predictive Models for the Evaluation of Air Quality and Noise in Urban Environments. *Sensors* **24**(311) (2024)
7. Handhayani, T.: An integrated analysis of air pollution and meteorological conditions in Jakarta. *Sci. Reports* **13**(5798) (2023)
8. Wu, D., Zheng, H., Li, Q., Wang, S., Zhao, B., Jin, L., Lyu, R., Li, S., Liu, Y., Chen, X., Zhang, F., Wu, Q., Liu, T., Jiang, J., Li, L.W.X., Chen, J., Hao, J.: Achieving health-oriented air pollution control requires integrating unequal toxicities of industrial particles. *Nat. Commun.* **14**(6491) (2023)
9. Zusman, E., Unger, C., Borgford-Parnell, N., Mar, K.: One atmosphere: integrating air pollution and climate policy and governance. *Atmosphere* **12**(1570) (2021)
10. Pozzer, A., Anenberg, S.C., Dey, S., Haines, A., Lelieveld, J., Chowdhury, S.: Mortality attributable to ambient air pollution: A review of global estimates. *GeoHealth* (7) (2023)
11. Han, J., Kamber, M., Pei, J.: *Data mining concepts and techniques (Third Edition)*, Elsevier Inc. (2012)

12. Fan, C., Chen, M., Wang, X., Wang, J., Huang, B.: A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Front Energy Res* **9** (2021)
13. Hasselbring, W.: Information System Integration. *Commun. ACM* **43**(6) (2000)
14. Kyriakidis, I., Karatzas, K.D., Papadourakis, G.: Using preprocessing techniques in air quality forecasting with artificial neural networks. *Inform. Technol. Environ. Eng. Environ. Sci. Eng.*, 357–372 (2009)

Chapter 2

The Perceived Economic Impact of Data Science Solutions on Business Performance



Marian Pompiliu Cristescu , Dumitru Alexandru Mara ,
Lia Cornelia Culda , and Raluca Andreea Nerişanu 

Abstract The advent of digital transformation has ushered in an era in which data science solutions are critical to organizational success, particularly in terms of improving economic outcomes for businesses. In a landscape characterized by intense competition and rapid technological advancements, this study seeks to dissect the economic implications of implementing data analytics within corporate structures, with a focus on creating business value and augmenting organizational capabilities. This study, based on a comprehensive review of the literature and empirical data analysis, investigates the role of data analytics in fostering value generation and strengthening companies' analytics capabilities. The methodology includes a systematic literature review to identify the critical resources and factors that contribute to the development of data analytics proficiency within organizations. Furthermore, a survey approach examines the effects of data analytics on company performance and strategic resource allocation. The findings reveal a significant link between strong data analytics capabilities and improved economic performance, emphasizing the importance of tangible, human, and intangible resources in developing these capabilities. Furthermore, the study emphasizes data analytics' transformative potential in operational optimization, strategic decision-making, and market positioning, which will ultimately contribute to long-term economic growth.

This paper aims to provide an in-depth understanding of the economic impact of data science solutions on businesses by combining theoretical insights with empirical evidence, making it a valuable resource for academics, industry practitioners, and decision-makers with an interest in using the power of data analytics for economic advancement.

M. P. Cristescu · D. A. Mara (✉) · L. C. Culda · R. A. Nerişanu
Lucian Blaga University of Sibiu, Calea Dumbrăvii, No. 17, 550324 Sibiu, Romania
e-mail: dumitrualexandru.mara@ulbsibiu.ro

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
C. Ciurea et al. (eds.), *Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)*, Smart Innovation, Systems and Technologies 426,
https://doi.org/10.1007/978-981-96-0161-5_2

2.1 Introduction

As the business world changes rapidly, being able to use data to make smart, strategic decisions have become imperative for an organization's success. This paper addresses the relationship between data analytics and value creation in companies, focusing on the variables that influence this relationship.

2.2 Data Analytics Capability as a Driver for Improving Business Performance

Bharadwaj [1] suggests using the resource-based view of the enterprise to understand how IT affects company performance by linking it to scarce, hard-to-replace resources and skills. Achieving organizational capabilities through resource accumulation and development gives companies a competitive edge. The resource-based view holds that firms can gain a competitive edge or improve performance by sharing resources and strategies [2]. The resource-based view is suitable for this study because it identifies concepts that develop data analysis capacity [3]. State that the resource-based view assumes that capability development can only occur after a resource is acquired, so it depends on what resources a company acquires [4]. Provided the most common conceptualization of resources that can be applied to data analysis capabilities, dividing resources into tangible (financial and physical), human skills (employee knowledge and skills), and intangible (organizational culture and learning).

Gupta and George [5] state that a firm must possess a unique combination of financial, physical, human, and organizational resources and adapt them to market changes to create a capability. Importantly, they identified seven resources needed to build analytics capability [5]. List "data, technology, and other core resources (e.g., time and investment)" as tangible resources, "big data managerial and technical skills" as human resources, and "data-driven culture and organizational learning intensity" as intangible resources. When combined with technical skills to analyze data and management's willingness to use it to make key decisions, data builds analytical capacity. SME managers can compare analysis software solutions using a table with different criteria. Thus, they will score each feature based on its relevance to their business and process [6]. Suggested scoring cloud versus on-premises technology based on user experience, risks, scalability, and data security.

2.3 Areas that Can Be Improved by Developing Data Analysis Capabilities

Data is valuable, but financial reports rarely reflect this. Even the most data-rich organizations treat information assets that are not on the balance sheet as worthless, according to [7]. Without an explicit or defensible requirement to assess data value, organizations have failed to create methodologies that accurately measure information value or monitor value changes over time [8]. Argued that data analytics' combined insights and effective uses demonstrate its value [8]. Stated that "academics have not studied how BDA can lead to competitive advantage and strategic business value." They claimed that strategic business value can be both functional (market share, financial performance) and symbolic (positive brand image and reputation, environmental mitigation) [8]. Identified BDA value targets as improving organizational performance, business processes, product and service innovation, customer experience, and market improvement. The BDA value targets are organization-specific functions and processes in which data analytics can enhance the performance of the business. The company can optimize internal processes and boost business efficiency by improving decision-making. Product and service innovation can give a company an edge, and improving customer experience can boost market share and customer loyalty. BDA helps companies improve performance, innovate, and add value to customers.

Tallon and Kraemer [9] examined how business strategy and IT strategy affect firm strategic alignment and business value. To assess this value, perceptual measures of information systems (IS) were created to capture managers' intuition or understanding of how IT affects different business aspects. Company strategic value disciplines include operational excellence, customer intimacy, and product leadership, according to [9]. IT can affect five processes or functional areas: IT enhances supplier relationships by enabling closer collaboration and trust, leading to cost savings and improved service delivery [10]. Emphasize that strong supplier ties can result in customized solutions and better supply chain performance, benefiting a company financially. Similarly, [11] highlight how effective supplier relationships enhance operational efficiency and reduce costs, ultimately boosting customer satisfaction and revenue. By significantly enhancing production and operations, IT has a positive effect on the budget of an organization. Efficient processes, as noted by [12], result in increased productivity and decreased production times by eliminating superfluous activities and improving workflow. According to [13], "improving production and operations can lead to increased efficiency, which can lead to reduced costs and increased quality of products and services, which can lead to increased revenue and profitability."

Technology can provide valuable insights into customer needs and expectations. This can reveal new opportunities to offer better and more innovative products and services, increasing sales and customer loyalty [9]. IT technology automates sales processes and provides valuable customer information. This can improve customer segmentation and understanding, increasing sales and loyalty [9]. By enhancing

support and service, IT improves customer relations, leading to greater customer loyalty and satisfaction, which has a positive impact on the financial performance of a business [9]. The study conducted by [9] prioritizes these processes according to their significance in generating business value. It offers organizations a structured method for evaluating the influence of each process on their overall operations and financial results.

2.4 Methodology

The objectives of this research are

1. Assessing the importance of resources and organizational capabilities that contribute to the development of data analytics capability for companies in Sibiu in the HoReCa sector.

Hypothesis: Tangible, human and intangible resources are equally important in developing data analytics capability for companies.

2. Analyze the impact of using data analytics on the development of company performance and budget.

Hypothesis: The perception on the use of data analytics is that it has a positive impact on business processes.

The survey method was chosen based on previously used tools identified in the literature review, such as data analytics capability [5, 14] and business value [7, 14]. Stated that a survey-based approach effectively measures a company's big data analytics maturity. They recommend using surveys to examine analytics capabilities and their impact on perceived business value. Similarly, [3] suggested that future research should empirically evaluate this framework through surveys, interviews, observations, focus groups with experts, and industry case studies. They also recommended using both qualitative and quantitative data collection methods, with multiple analysis techniques like structural equation modeling and qualitative comparative analysis.

Survey research, as described by [15], quantitatively describes population trends or opinions through sample studies. Surveys measure variables and explore relationships between them [5, 16]. Used a cross-sectional survey to explore the development of Data Analytic Capabilities (DAC) and its business outcomes, considering experiences over the past 2 years. Our questionnaire, divided into three sections, gathers demographic data, evaluates firms' data analytics capabilities, and assesses the impact on company performance and budget. The second section, validated by [5], includes closed-ended questions on a 7-point Likert scale concerning firm resources, technology, skills, data-driven culture, and organizational learning intensity. The third section measures the impact of data analytics across functional areas, confirmed as an effective method by [6, 17].

This study employs a process-oriented approach to assess business value, using perceptual metrics to approximate the impact of data analytics on firm performance. We grouped survey items into categories to enhance data clarity and interpretation, streamline analysis, and improve the reliability of findings.

The target population for the survey comprises active hotels and restaurants in Sibiu, with an estimated size of 195 firms. We aim for a sample of 130 firms to ensure a 95% confidence level and a 5% margin of error using simple random sampling. This method helps reduce selection bias, ensuring objective and representative results. Questionnaire distribution is executed through various online channels, adhering to ethical and confidentiality standards.

2.5 Results

Despite a low average score for data management practices indicating poor performance, there is notable variation, with some firms performing better than others. Conversely, a higher average in resources allocated to data analysis suggests substantial investment in this area, complemented by good technical and managerial skills and a willingness to learn. SMEs generally view data analytics positively, particularly valuing its impact in marketing and sales, which is crucial for understanding customer behavior and refining marketing strategies. Although data analytics is seen as less beneficial in supplier relationships compared to customer relations, the overall positive attitude towards data suggests readiness for adopting more sophisticated data management techniques in the future (Table 2.1).

In the following, we will present the results following the multiple regression analysis of each independent variable with the dependent ones (Table 2.2).

The regression analysis reveals how various aspects of data analytics capability impact value creation in supplier relationships, with a significant proportion ($R^2 = 0.542$) of the variance explained by the factors considered. Managerial skills emerge as a key component, as indicated by the highest and statistically significant coefficient ($B = 1.232$, $p < 0.000$), suggesting their critical role in leveraging data analysis to enhance supplier relationships. This indicates that without strong managerial skills, even the best resources and technologies may not translate into effective actions or strategic decisions that strengthen these relationships.

Conversely, technical skills show a significant negative impact ($B = -0.438$, $p < 0.002$) on supplier relationship value, suggesting that an overemphasis on technical expertise can undermine the quality of these relationships by potentially neglecting necessary strategic and interpersonal dynamics.

Other factors like data management practices and organizational learning were also tested but showed no significant impact on the value of supplier relationships in this study. This indicates that while these factors are relevant, they did not influence supplier relationships significantly in the current data set (Table 2.3).

According to the model coefficients, “data management practices” ($B = 0.242$, $p < 0.000$) and “resources allocated to data analysis” ($B = 0.198$, $p = 0.003$) have

Table 2.1 Descriptive statistics on independent and dependent variables

Variable	Media	Median	Standard deviation
date_management_practices	2,625,641	2	1,43,718
resources_allocated_to_DA	3,961,538	4	1,135,432
technical_skills	4,556,923	4,4	0,863,734
managerial_abilities	4,570,513	4,333,333	0,747,913
attitude_to_data	4,667,692	4,6	0,598,993
organisational_learning	4,815,385	4,4	0,891,168
value_added_to_relationships_with_suppliers	2,786,538	2,5	1,205,652
value_added_to_production_and_operations	4,175,385	3,8	1,057,132
value_added_to_improved_products_and_services	4,130,769	4	0,926,011
value_added_marketing_and_vanities	4,936,923	4,4	1,028,896
value_added_to_customer_relationships	3,249,231	2,8	1,1789

Table 2.2 Correlation coefficients for the six independent variables chosen in relation to the value added to the relationship with suppliers

Model	Unstandardized coefficients		Standardized coefficients	t	Mr
	B	Std. Error	Beta		
(Constant)	-1,289	,834		-1,546	,125
date_management_practices	,110	,109	,131	1,007	,316
resources_allocated_to_DA	,075	,109	,070	,687	,493
technical_skills	-,438	,138	-,314	-3,162	,002
managerial_abilities	1,232	,172	,764	7,165	,000
attitude_to_data	-,248	,178	-,123	-1,393	,166
organisational_learning	,210	,148	,155	1,419	,159

^aRegression coefficients for the six independent variables chosen in relation to the value added to the relationship with suppliers

a significant and positive impact on value added in production and operations. This suggests that the implementation of sound data management practices and the appropriate allocation of resources to data analysis can amplify the efficiency of production and operations.

“Management skills” is also a significant predictor (B = 0.351, p = 0.001), highlighting the importance of effective management and strategic vision in using data to improve operational performance.

Table 2.3 Correlation coefficients for the six chosen independent variables in relation to value added to production and operations

Model		Unstandardized coefficients		Standardized coefficients	t	Mr
		B	Std. Error	Beta		
1	(Constant)	-,390	,504		-,773	,441
	date_management_practices	,242	,066	,329	3,684	,000
	resources_allocated_to_DA	,198	,066	,213	3,010	,003
	technical_skills	-,098	,084	-,080	-1,167	,245
	managerial_abilities	,351	,104	,248	3,377	,001
	attitude_to_data	,198	,108	,112	1,834	,069
	organisational_learning	,221	,090	,186	2,463	,015

^aRegression coefficients for the six chosen independent variables in relation to value added to production and operations

Although “technical skills” and “data attitude” were included in the model, they did not have a statistically significant impact on value added in production and operations. On the other hand, “organizational learning” had a significant positive effect ($B = 0.221, p = 0.015$), highlighting the idea that adaptability and continuous learning are essential to remain relevant and effective in a changing environment.

This analysis demonstrates that in order to enhance production and operations, businesses should give priority to the adoption of efficient data management practices, allocate sufficient resources to data analytics, build managerial expertise, and promote a culture of continuous organizational learning (Table 2.4).

The results reveal that “data management practices” ($B = 0.242, p < 0.000$) and “resources allocated to data analysis” ($B = 0.198, p = 0.003$) significantly and positively influence the value added in production and operations. This suggests that robust data management practices and appropriate resource allocation for data analysis can significantly enhance operational efficiency.

“Management skills” also emerge as a significant predictor ($B = 0.351, p = 0.001$), underscoring the importance of effective management and strategic vision in leveraging data to improve operational performance.

While “technical skills” and “data attitude” did not show a statistically significant impact, “organizational learning” demonstrated a significant positive effect ($B = 0.221, p = 0.015$). This emphasizes the importance of adaptability and continuous learning in maintaining relevance and effectiveness in a dynamic environment.

Based on this analysis, to boost production and operations, organizations should prioritize implementing efficient data management practices, adequately allocate resources to data analytics, develop managerial expertise, and foster a culture of continuous organizational learning (Table 2.5).

Table 2.4 Correlation coefficients for the six independent variables chosen in relation to the value added to the improvement of products and services

Model		Unstandardized coefficients		Standardized coefficients	t	Mr
		B	Std. Error	Beta		
1	(Constant)	,334	,571		,585	,560
	date_management_practices	,185	,075	,288	2,486	,014
	resources_allocated_to_DA	,245	,075	,300	3,285	,001
	technical_skills	,053	,095	,050	,564	,574
	managerial_abilities	-,026	,118	-,021	-,224	,823
	attitude_to_data	,488	,122	,316	4,000	,000
	organisational_learning	-,013	,102	-,012	-,125	,900

^aRegression coefficients for the six independent variables chosen in relation to the value added to the improvement of products and services

Table 2.5 Correlation coefficients for the six independent variables chosen in relation to value added in marketing and sales

Model		Unstandardized coefficients		Standardized coefficients	t	Mr
		B	Std. Error	Beta		
1	(Constant)	-2,150	,586		-3,666	,000
	date_management_practices	-,157	,077	-,219	-2,050	,042
	resources_allocated_to_DA	,067	,076	,074	,873	,385
	technical_skills	-,209	,097	-,175	-2,145	,034
	managerial_abilities	,397	,121	,289	3,283	,001
	attitude_to_data	,832	,125	,484	6,640	,000
	organisational_learning	,517	,104	,447	4,954	,000

^aRegression coefficients for the six independent variables chosen in relation to value added in marketing and sales

Interpreting the coefficients from the regression analysis, “attitude to data” (B = 0.832, p < 0.000) and “organizational learning” (B = 0.517, p < 0.000) show the largest positive impacts on marketing and sales value. This highlights that a proactive, positive data attitude and the ability to learn and adapt organizationally are crucial for enhancing marketing and sales initiatives.

“Management skills” ($B = 0.397, p = 0.001$) also demonstrate a significant positive effect, emphasizing the importance of effective management in marketing and sales efforts. Conversely, “technical skills” ($B = -0.209, p = 0.034$) and “data management practices” ($B = -0.157, p = 0.042$) present negative coefficients, suggesting they might not directly benefit or could potentially detract from marketing and sales value in the examined model (Table 2.6).

Examining the coefficients individually, we find that “managerial skills” ($B = 0.657, p < 0.000$) and “organizational learning” ($B = 0.524, p < 0.000$) have the strongest positive coefficients. This means that these two aspects are particularly relevant for increasing value in customer relationships. Effective management and an organization’s ability to learn and adapt are key to improving and strengthening customer relationships.

Also, “data management practices” ($B = 0.204, p = 0.014$) have a positive coefficient, indicating that these practices can add value to customer relationships, although to a lesser extent compared to the other two variables.

In contrast, the variables “technical skills” ($B = -0.143, p = 0.170$) and “attitude towards data” ($B = -0.097, p = 0.470$) have no statistically significant impact, although their coefficients are negative. This suggests that, in the context of the model analyzed, these aspects may be less relevant to the value added to customer relationships.

The effective use of data analytics has become a key factor in a company’s performance and success. Analyses have highlighted different aspects of data analytics capability and their impact on various dimensions of company performance.

Table 2.6 Correlation coefficients for the six independent variables chosen in relation to the value added to the customer relationship

Model		Unstandardized coefficients		Standardized coefficients	t	Mr
		B	Std. Error	Beta		
1	(Constant)	-1,794	,625		-2,870	,005
	date_management_practices	,204	,082	,248	2,496	,014
	resources_allocated_to_DA	,022	,082	,021	,264	,792
	technical_skills	-,143	,104	-,105	-1,380	,170
	managerial_abilities	,657	,129	,417	5,096	,000
	attitude_to_data	-,097	,134	-,049	-,725	,470
	organisational_learning	,524	,111	,396	4,719	,000

^aRegression coefficients for the six independent variables chosen in relation to the value added to the customer relationship

To summarize, in all the models analyzed, an organization's ability to learn and adapt was found to be important. This is fundamental for any company that wants to be resilient and continuously adapt to changes in its environment.

Effective and strategic management was also essential in several dimensions, highlighting the importance of sound management in the use of data.

Valuing data and understanding its importance were significant issues, particularly in the context of marketing and sales.

While data management practices and resources allocated to data analysis also had a positive impact, they were not as strong as the issues mentioned above. However, they are essential components to build and sustain data analytics capacity.

Surprisingly, technical skills, while important, did not always have a statistically significant impact in our models. This suggests that, while technology is essential, the success of data analysis in a company depends more on culture, mindset, and the ability to manage and use information than simply possessing technical skills.

Consequently, the budgetary implications of utilizing data analytics to enhance the performance of an organization are significant. In order to optimize their operations, businesses should allocate resources towards fostering a culture of continuous learning within the organization, maintain competent leadership, and advocate for a constructive outlook on data. Understanding the true value of data with these can result in increased expenditures and enhanced performance.

2.6 Conclusion

Data analytics has become an essential part of current corporate operations. In the age of globalization and digitalization, a company's capacity to understand and efficiently use the data it acquires can determine its success or failure. Following careful examination of the findings presented above. Regarding the first research question, which is how data analytics can contribute to value creation for businesses, there is a significant link between data analytics capability and value added in various areas of the business, including manufacturing, product and service improvement, marketing, and customer relations. This suggests that by using data effectively, companies can optimize their processes, identify new market opportunities, improve the quality of products and services, and build stronger customer relationships. All this translates into increased revenues and a stronger market position.

As the business world evolves and becomes increasingly interconnected and technology dependent, the amount of data available grows exponentially. Whether it's data about customers, transactions, user behavior on online platforms, or user feedback, all this data can be used to increase operational efficiency, optimize marketing strategies, improve customer relationships, or identify new business opportunities.

But data collection alone is not enough. It is imperative that this data is interpreted and used appropriately. This is where data analysis capabilities come in. The more a company is able to analyze and understand this data, the more likely it is to create added value for its business.

For example, accurate interpretation of consumer behavior data can help a company identify emerging trends in the market, adapt quickly to changes in consumer preferences, and offer products or services that truly meet their needs and wants. Data analytics can also help identify the effectiveness of different marketing campaigns, allowing resources to be reallocated to the most profitable strategies.

We see from the results that organizational resources and capabilities are vital. Variables such as organizational learning and managerial skills were found to have a significant impact in several of the analyses. This underlines the point that it is not enough for a company to have access to data or to use sophisticated analytical tools. Organizational culture and team skills play an important role in correctly interpreting data and using it to make informed decisions. Ongoing training and investment in employee skills development are therefore essential to maximize the value of data analysis.

To answer the question regarding the impact of using data analytics to develop company performance, data analytics significantly influences performance in key areas, with the potential to improve efficiency and generate growth. For example, in manufacturing, data analytics has been associated with greater efficiency, which can lead to lower costs and higher profit margins. In budget terms, this suggests that investing in data analytics can deliver a significant ROI (return on investment). Companies that understand and act on the insights gained from data analytics can optimize budgets, strategically allocating resources where they generate the greatest value.

In the context of SMEs in Sibiu in the HoReCa sector, our analyses highlighted the vital importance of organizational learning. A learning organization is one that constantly adjusts its approaches and methods based on feedback and information gathered. In this sense, a culture of organizational learning can help companies quickly identify new consumer trends, optimize offerings, and maximize profitability.

Managerial skills also played an important role in the impact of data analysis. Effective management that understands and harnesses the power of data can turn this information into actionable strategies, effectively directing resources and aligning company goals with market realities. A manager who understands the value of data can make better informed decisions and direct efforts where they bring the greatest value.

Another significant aspect was the attitude of companies towards data. Having a positive and proactive outlook on data, viewing it as a strategic tool, was correlated with better performance, especially in the areas of marketing and sales. This underlines the fact that beyond mere numbers and statistics, a company's attitude and culture towards data can have a major impact on its success.

On the other hand, while technical skills and resources allocated to data analysis were relevant, they were not necessarily paramount.

Thus, for Sibiu SMEs in the HoReCa sector and many other companies, data analytics can have a significant impact on performance and budget. To maximize this impact, companies need to cultivate a culture of learning, invest in developing managerial skills, and adopt a proactive and strategic attitude towards data. This will

not only boost performance, but will also help strengthen the company's position in the market, thus ensuring sustainable and durable growth.

In conclusion, data analysis is not just a technical tool or a momentary trend. It is a key contributor to the growth and competitiveness of companies in the modern business landscape. To maximize this value, organizations need to embed data analytics into their culture, invest in training, and commit to an approach that combines technology with strategy and human skills.

References

1. Bharadwaj, A.S.: A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS Q.* **24**(1), 169–196 (2000)
2. Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S.F., Childe, S.J., Hazen, B., Akter, S.: Big data and predictive analytics for supply chain and organizational performance. *J. Bus. Res.* **70**, 308–317 (2017)
3. Mikalef, P., Pappas, I.O., Krogstie, J., Giannakos, M.: Big data analytics capabilities: a systematic literature review and research agenda. *IseB* **16**(3), 547–578 (2017)
4. Grant, R.M.: The resource-based theory of competitive advantage: implications for strategy formulation. *Calif. Manage. Rev.* **33**(3), 114–135 (1991)
5. Gupta, M., George, J.F.: Toward the development of a big data analytics capability. *Inform. Manag.* **53**(8), 1049–1064 (2016)
6. Tallon, P.P., Mooney, J.G., Duddek, M.: Measuring the business value of IT. In: *Measuring the Business Value of Cloud Computing*, pp. 1–17. Springer, Cham (2020)
7. Tallon, P.P.: Corporate governance of big data: perspectives on value, risk, and cost. *Computer* **46**(6), 32–38 (2013)
8. Grover, V., Chiang, R.H., Liang, T.P., Zhang, D.: Creating strategic business value from big data analytics: a research framework. *J. Manag. Inf. Syst.* **35**(2), 388–423 (2018)
9. Tallon, P.P., Kraemer, K.L.: Fact or fiction? A sensemaking perspective on the reality behind executives' perceptions of IT business value. *J. Manag. Inf. Syst.* **24**(1), 13–54 (2007)
10. Le, S., Wu, J., Zhong, J.: Relationship quality and supply chain quality performance: The effect of supply chain integration in hotel industry. *Comput. Intell.* **37**(3), 1388–1404 (2021). <https://doi.org/10.1111/coin.12379>
11. Trkman, P., McCormack, K., de Oliveira, M.P.V., Ladeira, M.B.: The impact of business analytics on supply chain performance. *Decis. Support. Syst.* **49**(3), 318–327 (2010). <https://doi.org/10.1016/j.dss.2010.03.007>
12. Bicheno, J.: *The lean toolbox: the essential guide to lean transformation*. PICSIE Books (2008)
13. Kaplinsky, R.: *Globalization, poverty and inequality: between a rock and a hard place*. Polity, Cambridge (2005)
14. Mikalef, P., Boura, M., Lekakos, G., Krogstie, J.: Big data analytics and firm performance: findings from a mixed-method approach. *J. Bus. Res.* **98**, 261–276 (2019)
15. Creswell, J.W., Creswell, J.D.: *Research design: qualitative, quantitative, and mixed methods approaches*. SAGE, Los Angeles (2018)
16. Singleton, R., Straits, B.C.: *Approaches to social research*. Oxford University Press, New York (2018)
17. Mooney, J.G., Gurbaxani, V., Kraemer, K.L.: A process oriented framework for assessing the business value of information technology. *ACM SIGMIS Database* **27**(2), 68–81 (1996)

Chapter 3

Performance Evaluation of Data Vault and Dimensional Modeling: Insights from TPC-DS Dataset Analysis



Andreea Vines

Abstract The current paper presents a comparison study between the Data Vault and the Dimensional Modeling and focuses on their performance in data loading and storage efficiency using the TPC-DS dataset. Data Vault, known for its adaptability and scale, is analyzed against a star schema Dimensional Model, known for its simplicity and query optimization. Through empirical evaluations covering data load speed, storage utilization, query execution, and scalability metrics, the author investigates the strengths and weaknesses of each modeling technique. It leverages real scenarios represented by the TPC-DS dataset, exploring various data volumes, complexities, and schema evolution scenarios. By analyzing the compromises between complexity and performance, the research aims to provide valuable insights for organizations that seek to optimize their data storage strategy. The findings presented in this paper aim to guide researchers and practitioners in choosing a data modeling technique according to their needs.

3.1 Introduction

In the digital age, data penetrates all aspects of business and becomes indispensable for companies that are looking to unlock valuable insights. Businesses rely on data analysis as the cornerstone of strategic decision-making, whether it is for targeted marketing campaigns, optimizing supply chain logistics by real-time monitoring, or fine-tuning product offerings based on market trends. This dependence on data extends to various industries, from finance and health care to retail and manufacturing, driving competitiveness and simulating growth in today's business environment. As a result, companies are constantly exploring innovative ways to use their data and push the boundaries of analytical techniques to obtain deeper and more meaningful insights to progress in the digital race.

A. Vines (✉)
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: andreea.vines@csie.ase.com

Organizations face many challenges in terms of the use and the analysis of operational data that comes from various sources. The main objective of these efforts is to gain insights to facilitate decision-making processes. The central element of a decision-making system is a data warehouse. It can be built using a modeling approach that considers all aspects of the development phase, including data modeling, project and risk management, deployment, and many other key aspects [1]. The most known two approaches to data modeling in data warehouses were created by Inmon [2] and Kimball [3]. However, in recent years, a third approach, called Data Vault, has emerged, and is spreading at a high speed, being designed to improve the flexibility, scalability, auditability, and productivity of the data warehouses [4, 5].

The purpose of this paper is to provide a comprehensive comparison between Data Vault and Dimensional Modeling in terms of data ingestion process. To provide an accurate analysis in the context of large volumes of data, multiple input datasets using TPC-DS benchmark data. The current paper debuts with a short introduction between Data Vault and Dimensional Modeling, followed by some use cases where TPC data has been used to provide insights. The research work provides two versions of data architecture in the context of both modeling techniques, followed by the data ingestion results.

The research is based on the hypothesis that Data Vault modeling technique is more complex than Dimensional Model, but it offers some advantages, such as a better performance in terms of ingestion process, compute power, flexibility, and scalability in terms of changes in the data source model, traceability, and performance. The paper is organized into the following sections: an introduction, a comprehensive literature review study that covers data modeling technologies and use cases on how TPC benchmark data can be used in analysis, the methodology used and the architecture of both models, an analysis of the results obtained by ingesting TPC-DS data, and finally the conclusions with some key findings of the research and potential directions for future research.

3.2 Literature Review

In the fields of data storage and Business Intelligence, the selection of an appropriate data modeling technique is crucial to ensuring the efficacy, scalability, and flexibility of the underlying data structure. Among the numerous approaches available, two leading approaches are Dimensional Modeling and Data Vault modeling. The methods provide unique perspectives on how data structures and organizes within the data warehouse environment, each with its strengths and weaknesses.

Data Vault is a relatively new methodology introduced in the 2010s and currently used in more and more implementations, although there are not many specialized research papers related to this. Regarding this technique, Gribova Svetlana [6] provides an overview of the research studies published on this area, where most of them focus on providing theoretical details in terms of performance—parallel data loads, easy integration of multiple sources, auditability, and flexibility.

In contrast, according to the research of Gribova Svetlana [6], the least studied area is regarding using metadata in Data Vault projects. One of the little studies on metadata models was presented by Černjeka [7] who developed a metadata model that translated NoSQL document data into a Data Vault model, enabling data traceability, auditability, and trend discovery. The integration between data modeling and Gen AI was presented in a recent paper where the authors generated multiple data models using different Gen AI model versions combined with prompt engineering and ranked their capabilities in terms of correctness, completeness, integrity, flexibility, understandability, simplicity, integration, and implement ability [8].

When it comes to the comparison between Data Vault and Dimensional Modeling, multiple studies were conducted and several comparisons on various criteria, including ETL implementation, life-cycle management, traceability, and data modeling [1, 9, 10]. The paper presented by Naamane and Jovanovic [9] provides a detailed example of a customer experience in data mart. Summarizing the mentioned studies, the authors presented Data Vault as a more complex methodology that requires more steps in the initial implementation, but with a better life-cycle management in terms of modifying or integrating new sources. Regarding historical data, both models allow for some historical preservation, but the Dimensional Model only stores history of dimensions and not fact tables.

Based on the above studies, the topics of data ingestion performance and data volume were not addressed in terms of the complexity of Data Vault which requires writing the data into multiple tables, using the hubs, link, and satellite concepts. Considering that this topic wasn't mentioned in the previous research studies, this paper concentrates on understanding how both Data Vault and Dimensional Models perform in terms of these areas.

The Transaction Processing Council (TPC) is a non-profit corporation that was founded in August 1988, and it is the authority currently responsible for setting benchmarks for transactions processing and database applications. The organization was created to allow computer suppliers to compare products using standard tests, providing a standard for evaluating and comparing the performance of database management systems [11, 12]. TPC currently contains around 11 active benchmarks that serve different purposes, such as data integration, transactional systems, OLAP systems, or Big Data.

TPC-DS is used to model decision support functions for retail suppliers. The support system includes essential business information, such as customers, stores, items, and orders. It is a robust, rigorous, and complete benchmark used for evaluating database management systems. TPC-DS uses the snowflake schema and contains 17-dimensional tables that are associated with three sales channels—stores, catalogs, Internet, and additional return tables. The benchmark allows to generate different data volumes, depending on the scale factors.

These benchmarks are usually used to test the performance of various systems, such as PostgreSQL, MySQL, Microsoft SQL Server, or NoSQL databases [13, 14]. However, it is a useful source for other implementations, such as creating dimensional cubes models based on the input data, designing Cloud Analytics Platforms, or improving the existing models by adding new temporal dimensions [15–17].

3.3 Methodology

The current paper presents a comprehensive comparison between Data Vault and Dimensional Model, from the perspective of data ingestion and data storage. It uses a snowflake schema dataset from TPC-DS and translates it in both Data Vault and Dimensional Model star schema models. All transformations and orchestrations are done in Databricks SQL.

This research is based on the hypothesis that the Data Vault model adds more complexity in terms of implementation and storage required compared to a Dimensional Model as the structure is more complex, but those changes don't have a big impact on the overall performance of the model.

To address the hypothesis presented above, the following methodical steps were performed:

1. Initial data analysis on TPC-DS dataset to understand its structure, relationships, and cardinalities.
2. Design of the Dimensional Model to create the star schema, design dimension, and fact tables.
3. Implement the Dimensional Model to load data into dimension and fact tables using a Databricks workflow.
4. Design of the Data Vault model by creating the hubs, satellites, and link tables and identifying business keys.
5. Implement the Data Vault model to load data using a Databricks workflow.
6. Evaluate both models and compare based on performance, storage, and ease of maintenance.

This section covers the following—the initial dataset and how it was ingested, including data characteristics and volume, the Data Vault model architectures with its components and the Dimensional Model architecture.

3.3.1 TPC-DS Dataset

To ensure data quality and volume, the research paper uses TPC-DS dataset. The TPC-DS (Transaction Processing Performance Council Decision Support) is a benchmark that is designed to evaluate the performance of decision support systems, particularly data warehousing or analytics platforms. It offers a standardized dataset that represents real-world business scenarios and can be generated using different data volumes to simulate a mix of read-intensive and write-intensive operations.

The initial dataset consists in a set of 24 tables that cover the sales and returns process of an organization that uses three primary sales channels—stores, catalogs, and the Internet. For the current research, the dataset was restricted to only cover the sales of the organization through the store channel. The detailed dataset components are presented in Table 3.1.

Table 3.1 TPC-DS dataset description

Object name	Description
store_sales	This represents the transactions table which stores the information regarding the items that were sold
customer	This table contains a list of unique customer details
store	This table contains details of a unique list of stores (address, manager, number of employees, working hours)
item	This table presents unique products formulations (brand, manufacturer, size)
promotion	This table stores details of specific promotion products (advertising, sales, PR)
date_dim	This table contains calendar details, where each records stores a calendar day
customer_address	This table presents details regarding the customers' addresses
customer_demographics	This tables stores demographic information about customers, such as marital status, education level, gender, and credit rating

The snowflake schema of the initial TPC-DS model is presented in Fig. 3.1.

As the TPC-DS benchmark allows using multiple scale factors to test the performance, 3 different scale factors were used to generate data—1 GB, 100 GB and 1000

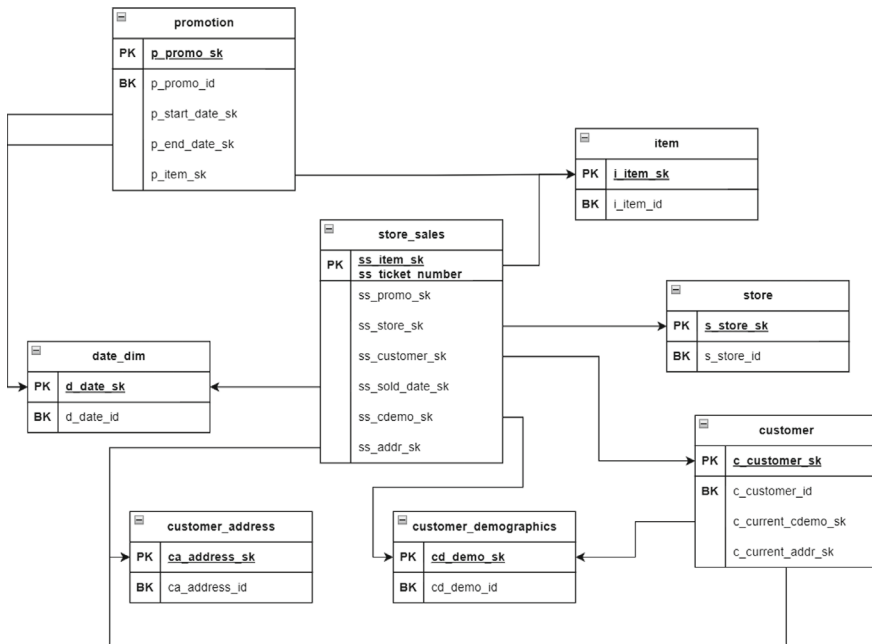


Fig. 3.1 TPC-DS dataset diagram

Table 3.2 TPC-DS dataset number of records for each scale factor

Object name	1 GB	100 GB	1000 GB
store_sales	488,870	49,025,613	493,225,008
customer	100,000	2,000,000	12,000,000
store	6	201	501
item	9,000	102,000	150,000
promotion	300	1000	1,500
date_dim	73,049	73,049	73,049
customer_address	50,000	1,000,000	6,000,000
customer_demographics	1,920,800	1,920,800	1,920,800

GB. Data was ingested for both Data Vault and Dimensional Modeling in batches, where each set of the three mentioned earlier was considered a batch, starting with the one of 1000 GB (Table 3.2).

3.3.2 Data Vault Architecture

The architecture diagram of the Data Vault model can be observed in Fig. 3.2. It consists of 11 tables that are split into hubs, links, and satellites. It was designed as a hub table for each business object (customer, store, promotion, and item) with a hash key generated using the business key. The attributes of the business objects are stored in the satellites table, which are used to also keep the history of the data and to track changes. As it can be observed in the initial dataset schema from Fig. 3.1, the sales table is linked with all business objects. To improve the performance of the Data Vault model, it created a transactional link table which contains the hash keys of all business objects that are linked with the sales table, instead of creating different link tables between each two objects.

From an ETL perspective, the Data Vault model doesn't have many dependencies on other tables, so they can be loaded into tables in parallel. The only dependencies on hub tables are for the two link tables which require reading the information from the hubs. Furthermore, the hubs and satellites are loaded in parallel, which offer a performance boost in the ingestion process.

All hash key columns are created using MD5 function and stored as binary columns.

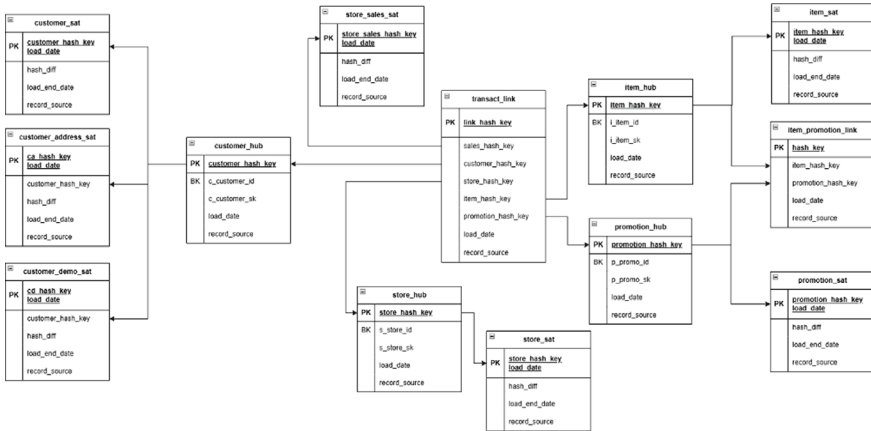


Fig. 3.2 Data vault model diagram

3.3.3 Dimensional Model Architecture

Dimensional modeling is concentrated in storing data in dimension and fact tables. The model usually contains a fact table which is surrounded by different tables or objects which house attributes, such as unique identifiers for each subject. Fact tables are used to define the performance KPIs of the organization.

The current model uses a fact table that stores the sales of the organization with foreign keys to the dimensional tables—customers, stores, items, promotions, and date. To keep the star schema of the model and avoid any possible loops, the model was defined with a bridge table that incorporated the information regarding promotions, sales, and items. Using a bridge table will avoid any possible many-to-many relationships or loops in the current model.

From an ETL process perspective, dimensional modeling requires loading dimensional tables first and then the fact table—which will use dimensional table’s primary keys as foreign keys in the fact table to connect them (Fig. 3.3).

3.4 Results and Discussion

This section describes a direct comparison of Data Vault and Dimensional Model results in terms of data ingestion process. Both models described in the previous section are filled using a Databricks workflow using a D8s_v3 cluster with five worker nodes. The cluster’s auto-scaling option has been disabled to obtain a more precise cost estimate.

The first step of the process requires the tables to be recreated and populated using the dataset generated for scale factor 1000 GB. The results obtained can be seen in

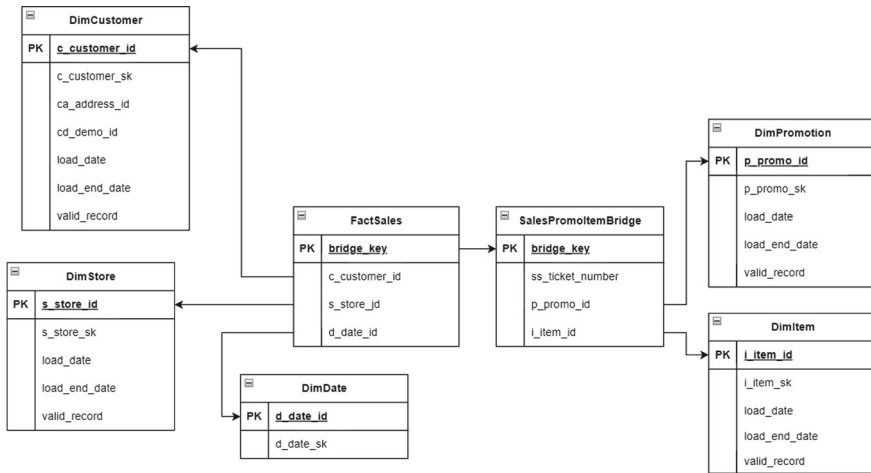


Fig. 3.3 Dimensional model diagram

Table 3.3—Data Vault models offer a better performance in terms of running time and can consume all the input data in less than 13 min. On the other hand, the same input data was ingested in the Dimensional Model in approximately 19 min as there were dependencies and the fact table requires information from dimensional tables.

In terms of storage, it can be observed that the Data Vault model requires more storage compared to the Dimensional Model. The first model has approximately 60% more storage than the second. The storage volume is higher because the first model requires to keep the historical data for the sales source as well, whereas in the Dimensional Model the historical data is only kept for dimensional tables, using slowly changing dimension’s type-2 approach.

The current results were extended and tested to analyze the performance of incremental load, considering that the tables were already populated with the previous batch of that which was described earlier. Both workflows were run for the second time with a new input dataset generated using the scale factor 100 GB, and the both models loaded data in approximately 10 min, having the same cost of 0.495\$ for the compute cluster which was described at the beginning of the section.

Third batch run had a smaller input dataset generated with the scale factor of 1 GB, where Data Vault tables were populated in approximately 7 min. The Dimensional

Table 3.3 Summary of full data load results (scale factor—1000 GB)

Data modeling approach	Number of tables	Total size of data (GB)	Running time	Total DBU	Cost (\$)
Data vault	11	94.54	12 m 28 s	1.95	0.585
Dimensional modeling	7	57.24	18 m 39 s	2.85	0.855

Table 3.4 Summary of incremental data load timings

Data modeling approach	Running time SF = 1 GB	Cost SF = 1 GB (\$)	Running time SF = 100 GB	Cost SF = 100 GB (\$)
Data vault	6 m 27 s	0.315	10 m 15 s	0.495
Dimensional modeling	9 m 45 s	0.45	10 m 9 s	0.495

Model ingestion process took close to 10 min, as it can be observed in Table 3.4. In the same table, it can be noted that the running time between the two input datasets is not significantly different as all the workflow runs require around two to three minutes for the job cluster to be spined up.

In conclusion, based on the results obtained in Tables 3.3 and 3.4, it can be confirmed that the hypothesis which agrees that Data Vault model is more complex compared to Dimensional Model and might require more storage or compute power, but without a big impact in terms of overall performance model. It was observed that the first model requires less compute power to ingest the data, hence the overall cost will be lower. However, it does require more storage compared to Dimensional Modeling as it will store more information, including the history of sales data for audit purposes. A Data Vault solution requires more storage overall, but this should not impact the overall project as the cloud storage is cheaper compared to compute power.

3.5 Conclusion

The present study provides valuable insights into the comparison between Data Vault and Dimensional Modeling using the TPC-DS dataset as inputs, reflecting a relatively unexplored area of data warehousing and business intelligence. Based on the results of the analysis presented, implementation in Data Vault has been observed to be advantageous in terms of ingestion, with data flow requiring less time than Dimensional Model, but with more tables that are being populated. This advantage is generated by the loading strategy, where most of the tables can be loaded simultaneously as there are only a few dependencies. In contrast, the models implemented in Data Vault have larger storage required than the dimensional approach, but this can be explained by the fact that they store the history of all input datasets, including sales data. This is not a concern as the costs for cloud storage are cheaper compared to the compute power, where it performs better.

Given the limited number of published papers on Data Vault, the current research contributes with additional information. While existing studies focus on theoretical comparisons between Data Vault and Dimensional Model in terms of integration, flexibility, life-cycle management, or auditability, this work focuses on evaluating the performance of the data ingestion and storage processes. Proving an example of implementation of a solution using both modeling techniques will showcase the

differences between the two approaches in terms of ingesting data and reviewing both full and incremental load batches.

The use of the TPC-DS dataset also adds value in the current research serving as a foundation for further exploration, as it can be leveraged by other researchers on exploring other directions. Furthermore, the study can be expanded to other areas, such as including new comparison factors, to form a comprehensive set of criteria for analyzing the two methodologies and providing additional value in decision-making.

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI—UEFISCDI, project number COFUND-CETP-SMART-LEM-1, within PNCDI IV.

References

1. Yessad, L., Labiod, A.: Comparative study of data warehouses modeling approaches: Inmon, Kimball, and Data Vault. In: 2016 International Conference on System Reliability and Science (ICSRS), Paris, France, pp. 95–99 (2016)
2. Inmon, B.: Building the warehouse, 2nd edn. John Wiley & Sons, New York (1996)
3. Kimball, R., Ross, M.: The data warehousing toolkit. John Wiley & Sons, New York (1996)
4. Mathiew, D.: Data Vault et BI. <http://fr.slideshare.net/dlinstedt/presentationdata-vault-et-biv-20120508>, last accessed 2024/03/27
5. Vines, A., Samoila, R.: An overview of data vault methodology and its benefits. *Informatica Economica* **5**, 99–110 (2023)
6. Gribova, S.: Literature review on data vaults—what is the state of the art of literature on data vaults. <https://www.fh-wedel.de>, last accessed 2024/03/27
7. Černjeka, K., Jakšić, D., Jovanovic, V.: NoSQL document store translation to data vault based EDW. In: 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, pp. 1197–1202 (2018)
8. Helskyaho, H., Ruotsalainen, L., Männistö, T.: Defining data model quality metrics for data vault 2.0 model evaluation. *Inventions* **9**(1), 21
9. Naamane, Z., Jovanovic, V.: Effectiveness of data vault compared to dimensional data marts on overall performance of a data warehouse system. *Intern. J. Comp. Sci.* **13**, 1 (2016)
10. Schnider, D., Martino, A., Eschermann, M.: Comparison of data modeling methods for a core data warehouse. https://danischnider.files.wordpress.com/2019/09/comparison_dwh_core_modeling.pdf, last accessed 2024/03/27
11. Barata, M., Bernardino, J., Furtado, P.: An overview of decision support benchmarks: TPC-DS, TPC-H and SSB. In: Rocha, A., Correia, A., Costanzo, S., Reis, L. (eds) *New contributions in information systems and technologies. Advances in Intelligent Systems and Computing*, vol. 353. Springer, Cham (2015)
12. Poess, M., Smith, B., Kollár, L., Larson, P.Å.: TPC-DS, taking decision support benchmarking to the next level. pp. 582–587. <https://doi.org/10.1145/564691.564759> (2002)
13. Vershinin, I.S., Mustafina, A.R.: Performance analysis of PostgreSQL, MySQL, microsoft SQL server systems based on TPC-H Tests. In: 2021 International Russian Automation Conference (RusAutoCon), Sochi, Russian Federation, pp. 683–687 (2021)
14. Zhao, H., Ye, X.: A practice of TPC-DS multidimensional implementation on NoSQL database systems, pp. 93–108, (2013)
15. Cuzzocrea, A., Moussa, R.: Multidimensional database design via schema transformation: turning TPC H into the TPC H*d multidimensional benchmark. In: *Proceedings of the 19th International Conference on Management of Data (COMAD '13)*. Computer Society of India, Mumbai, Maharashtra, IND, pp. 56–67 (2013)

16. Al-Kateb, M., Crolotte, A., Ghazal, A., Rose, L.: Adding a temporal dimension to the TPC-H Benchmark. In: Nambiar, R., Poess, M. (eds) Selected topics in performance evaluation and benchmarking. TPCTC 2012. Lecture notes in computer science, vol 7755, Springer, Berlin, Heidelberg (2013)
17. Chen, G., Johnson, T., Cilimdžić, M.: Quantifying cloud data analytic platform scalability with extended TPC-DS Benchmark, (2022)

Chapter 4

Revolutionizing Healthcare: Harnessing Natural Language Processing and Big Data for Predictive Disease Diagnosis



Diana Andreea Căuniac and Andreea Alexandra Cîrnaru

Abstract The current context of health systems is characterized by a rapid evolution of technology and an increasing demand for health services. There is an increasing emphasis on preventive care and the use of data and technology to improve the efficiency and effectiveness of healthcare delivery. Currently, there is an evolution of digital transformation and the integration of natural language processing and big data analysis has become very important to revolutionize various fields, including the field of health. This article explores the synergies between NLP and big data in the context of predictive disease diagnosis, aiming to elucidate how these technologies can improve early disease detection, treatment, and management. By exploring different NLP techniques, such as sentiment analysis and text classification, combined with big data analysis methodologies, the article presents how predictive models can be developed to anticipate the onset and progression of diseases. Through this paper, you will find the potential of combining natural language processing with big data to diagnose patients' diseases more quickly, which offers a path to personalized medicine and ultimately to a healthier society. Through this paper, various NLP and big data techniques are presented with brief descriptions. The dataset is described and preprocessed for analysis. Word cloud visualization highlights prevalent disease symptoms, while latent Dirichlet allocation categorizes symptoms by disease. Additionally, TF-IDF vectorization and K-means clustering are employed for symptom clustering. The silhouette score, calculated at 0.7567, underscores the quality of the analysis.

D. A. Căuniac (✉) · A. A. Cîrnaru
Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania
e-mail: diana.cauniac@csie.ase.ro

A. A. Cîrnaru
e-mail: alexandra.cirnar@csie.ase.ro

4.1 Introduction

In the realm of modern healthcare, the integration of cutting-edge technologies has become pivotal in enhancing diagnostic accuracy, treatment efficacy, and patient outcomes. Among these technologies, Natural Language Processing (NLP) and Big Data analytics stand out as transformative tools with the potential to revolutionize the landscape of disease diagnosis and management.

The convergence of NLP and big data offers unprecedented opportunities to extract valuable insights from the vast troves of medical data generated daily in healthcare settings. From electronic health records (EHRs) and medical imaging reports to patient feedback and clinical trial data, the immensity of information available holds immense potential for predictive analytics and personalized medicine [1].

This article delves into the intersection of NLP and big data in healthcare, focusing on their application in predictive disease diagnosis. By harnessing the power of advanced computational techniques, healthcare practitioners can leverage NLP algorithms to parse through unstructured medical text, extracting clinically relevant information and sentiments. Furthermore, the integration of big data analytics enables the identification of intricate patterns and correlations within extensive datasets, facilitating the prediction of disease onset, progression, and response to treatment.

Through a comprehensive exploration of data collection, preprocessing, model development, and validation methodologies, this article aims to elucidate the multifaceted process of leveraging NLP and big data for predictive disease diagnosis. Additionally, it seeks to highlight the potential impact of these technologies on enhancing clinical decision-making, improving patient outcomes, and advancing population health management strategies [2].

In essence, the amalgamation of NLP and big data represents a paradigm shift in healthcare delivery, empowering clinicians with the tools and insights necessary to anticipate, prevent, and effectively manage diseases. By embracing these innovative approaches, the healthcare industry can embark on a transformative journey toward precision medicine, ultimately leading to a healthier and more resilient society [3].

4.2 Collection and Processing of Medical Data

NLP involves computers understanding, analyzing, and generating human language using machine learning algorithms [4]. Conversely, big data focuses on managing and analyzing large, complex datasets from various sources. Integration of NLP and big data offers substantial benefits across sectors like healthcare, retail, and finance [5].

Techniques in NLP include tokenization, POS tagging, Named Entity Recognition, sentiment analysis, word embeddings, and topic modeling [6]. Big data analysis techniques encompass MapReduce, Spark, Hadoop, Apache Kafka, Apache Flink,

and Data Warehousing, facilitating distributed data processing and real-time analytics [7].

The collection and processing of medical data are fundamental steps in the development and implementation of predictive disease diagnosis systems. In this chapter, a description of the selected data and how this data is prepared for analysis and use in predictive models will be presented.

The medical dataset used in this analysis was obtained from Kaggle, a renowned platform for accessing and sharing datasets for various data science projects. The dataset encompasses a wide array of medical data, including patients' symptomatology, diverse diagnostic profiles, symptom descriptions, precautionary measures associated with symptoms, and numerical indicators of symptom severity, graded on a scale from 1 to 10. Each severity rating reflects the intensity of symptoms, with 1 indicating the mildest and 10 representing the most severe symptoms [8].

The construction of this dataset involved meticulous curation and compilation efforts, drawing from a range of sources such as medical records, patient interviews, and clinical databases. Each diagnosis in the dataset can potentially present up to 17 distinct symptoms, contributing to the dataset's complexity and richness. With a total of 5000 unique diagnoses, the dataset offers a comprehensive and diverse collection of medical scenarios, enabling in-depth analysis and exploration of symptom patterns, severity levels, and diagnostic correlations.

Within the dataset pertaining to symptom descriptions, one can identify prevalent diseases along with concise descriptions thereof. Conversely, the dataset concerning symptom precautions encompasses information regarding prevalent diseases, as well as the corresponding measures to be undertaken when a patient presents symptoms indicative of a specific ailment.

After data collection and the establishment of datasets for analysis, the subsequent phase involves the implementation of data processing methods. These methods encompass three key steps:

- **Data Cleaning:** This step involves the identification and elimination of errors, duplicates, and incomplete data. Its primary objective is to ensure data consistency and quality. In this stage, missing data, duplicate values, and certain unnecessary information were eliminated.
- **Data Normalization:** This step entails standardizing the format and measurement units of the data. It is crucial for facilitating data comparison and analysis across different datasets.
- **Data Transformation:** In this phase, raw data undergoes conversion into a suitable format for analysis. This process enables the transformation of data into a structured and analyzable form. This step was applied to the dataset comprising the severity of symptoms and was verified as data [9].

Prior to analysis, the dataset underwent a rigorous data processing pipeline to ensure its quality and suitability for analysis. The data cleaning phase involved several key steps. Firstly, missing values were identified and handled using appropriate techniques such as imputation, where missing values were replaced with estimated values based on the data's distribution or other statistical measures. Outliers, which could

skew the analysis results, were detected and addressed using statistical methods like Z-score analysis or interquartile range (IQR) method.

Text data preprocessing was also an integral part of the data processing pipeline. This involved tokenization, where symptom descriptions were split into individual words or tokens, and lemmatization or stemming to reduce words to their base or root form for standardization. Additionally, stopword removal was performed to eliminate commonly occurring words that do not carry significant meaning in the context of symptom analysis.

Furthermore, data normalization techniques were applied to standardize numerical indicators of symptom severity. This ensured that severity ratings across different symptoms were on a consistent scale, facilitating meaningful comparisons during analysis. Categorical variables, such as diagnostic profiles and precautionary measures, were encoded using techniques like one-hot encoding to convert them into a numerical format suitable for analysis.

Once the data was cleaned, preprocessed, and normalized, it was structured and organized into a coherent format for efficient analysis. This involved categorizing and indexing the various data attributes, such as symptoms and diagnoses, for easy retrieval and manipulation. The processed dataset provided a reliable foundation for exploring symptom patterns, severity levels, and diagnostic correlations in medical research and analysis.

The data was prepared for the future analysis by the methods mentioned above, to be further processed by the NLP techniques that will be detailed in the next chapter.

4.3 Medical Data Analysis Using NLP and Big Data

After the initial data has undergone verification and preparation for analysis, the subsequent phase involves an examination of the descriptive text associated with the symptoms of each disease. This analysis encompasses two main components. Firstly, the process of tokenization is employed, whereby symptom descriptions are segmented into individual words or phrases. Following this, text processing techniques including stopword elimination, lemmatization, or stemming are applied to normalize and streamline the text. Secondly, a sentiment analysis is conducted to evaluate the emotional context of the symptom descriptions. This aspect of the analysis is instrumental in identifying symptoms that may be perceived as more severe or distressing by patients.

Through Fig. 4.1 it is shown a word cloud generated by tokenizing and eliminating stopwords to analyze symptom descriptions. In this visualization, larger sized words indicate higher frequency of appearance within the symptom descriptions. Consequently, the visualization offers a swift overview of the most prevalent terms in the symptom descriptions across the examined dataset. Observing the word cloud, “infection” emerges as the most prevalent term in the description of patients’ symptoms, closely followed by “liver”. This suggests a notable occurrence of liver diseases or various infections among the analyzed patients.

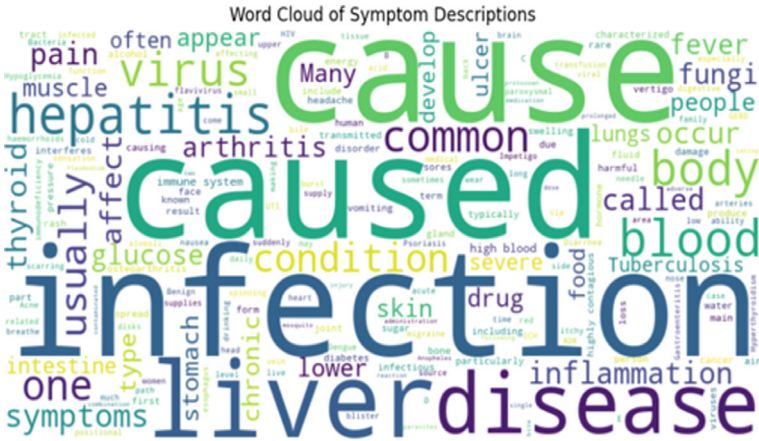


Fig. 4.1 Word cloud of symptom description

The symptom description dataset contains concise descriptions of the most prevalent diseases among patients. Figure 4.2 illustrates a graph depicting the most common words extracted from the symptom descriptions within the dataset. This visualization offers a rapid overview of the terms frequently utilized to depict symptoms, offering valuable insights into prevalent patterns observed in reported symptoms.

Identifying the most frequently occurring words provides insight into which symptoms are prevalent or most relevant among patients. This can be useful for doctors and

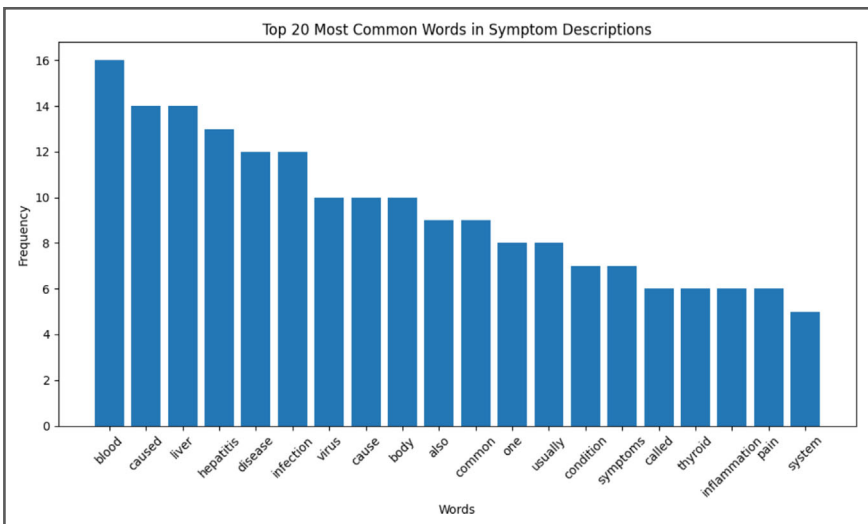


Fig. 4.2 Top 20 most common words in symptom descriptions


```

Topics for Varicose veins:
Topic 1: bruising prominent_veins_on_calf obesity cramps
Topic 2: swollen_blood_vessels bruising prominent_veins_on_calf obesity
Topic 3: prominent_veins_on_calf bruising obesity swollen_legs
Topic 4: fatigue obesity prominent_veins_on_calf bruising
Topic 5: cramps swollen_legs prominent_veins_on_calf bruising

Topics for Hypothyroidism:
Topic 1: dizziness enlarged_thyroid weight_gain fatigue
Topic 2: irritability puffy_face_and_eyes swollen_extremities fatigue
Topic 3: mood_swings lethargy fatigue puffy_face_and_eyes
Topic 4: brittle_nails cold_hands_and_feets fatigue swollen_extremities
Topic 5: abnormal_menstruation depression swollen_extremities irritability

```

Fig. 4.3 Most common symptoms across all diseases

researchers in managing and analyzing medical data, as well as identifying trends or correlations between certain symptoms and medical conditions. It can also provide clues for prioritizing or focusing certain aspects of healthcare within a health system.

Through Fig. 4.3 topic modeling was applied using latent Dirichlet allocation for each disease in order to identify common topics within the symptom descriptions for each disease, providing deeper insights into the nature of symptoms associated with different diseases. The CountVectorizer module was also used to transform the textual data (symptom descriptions) into a bag-of-words representation. The resulting matrix contains the frequency with which each term appears in symptom descriptions. The output provides a deeper understanding of latent patterns in symptom descriptions, which provide relevant information for disease diagnosis and management.

The chart exemplified in Fig. 4.4 represents an association between the diseases in the dataset and the clusters obtained by preprocessing the symptom descriptions, apply TF-IDF vectorization, and use k-means clustering to cluster similar symptom descriptions. Each row in the table corresponds to a disease and each column corresponds to a cluster. The values in the chart indicate the belonging of each symptom associated with the respective disease to a certain cluster. Through this representation, patterns or associations can be identified between the symptoms associated with different diseases and the clusters obtained, which can provide useful information in understanding the similarities and differences between the analyzed diseases.

For this illustration, the silhouette score was computed, yielding a value of 0.7567, indicating well-defined clusters. This suggests that the clustering algorithm successfully discerned notable patterns in the data.

Given the notably elevated concentration in cluster 7, it suggests a cluster of diseases sharing similar symptoms, predominantly grouped within cluster 7. This insight could be highly beneficial in unraveling the interplay between symptoms and diseases, offering valuable insights for diagnosis and treatment.

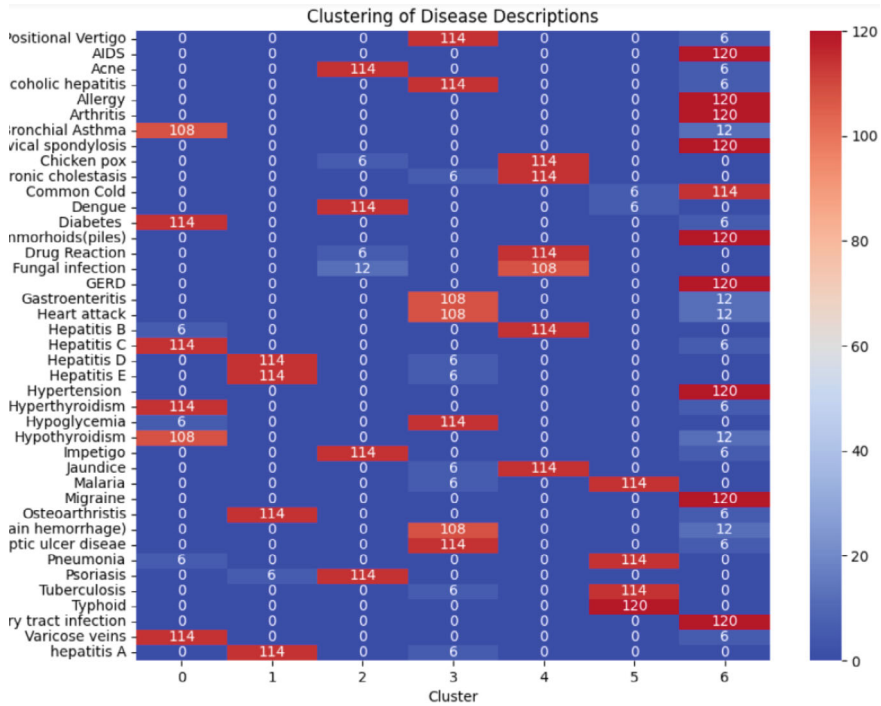


Fig. 4.4 Clustering of disease symptoms

Table 4.1 displays alphabetically the top 10 diseases from cluster 7 along with their specific symptoms. Some similarity can be observed among certain symptoms, which reinforces the integration of diseases into a common cluster.

In summary, clustering based on symptoms proves to be a valuable tool for unraveling the intricate connections between symptoms and corresponding diseases.

Big data assumes a pivotal role in the analysis and interpretation of extensive quantities of medical data, yielding substantial advantages across domains including medical research, healthcare provision, pharmaceutical innovation, and related spheres. Big data empowers researchers and healthcare practitioners to scrutinize and discern intricate patterns and trends within medical data. Big data can be used to anticipate health risks and take preventive measures before problems become severe. By analyzing historical medical data and other environmental and lifestyle factors, potential health conditions or risks can be identified, and steps can be taken to manage or prevent them.

By applying big data technologies to the analysis and interpretation of medical data, new and promising opportunities for improving health systems and providing more efficient and personalized care are opening up.

Table 4.1 Top 10 symptoms for the diseases in cluster 7

Disease	Symptom_1	Symptom_2	Symptom_3	Cluster
Acne	pus_filled_pimples	restlessness	scurring	7
AIDS	muscle_weakness	patches_in_throat	high_fever	7
Alcoholic hepatitis	yellowish_skin	abdominal_pain	swelling_of_stomach	7
Allergy	continuous_sneezing	shivering	breathlessness	7
Arthritis	muscle_weakness	stiff_neck	sewling_of_stomach	7
Bronchial asthma	cough	high_fever	breathlessness	7
Cervical spondylosis	back_pain	weakness_in_limbs	vomiting	7
Common cold	continuous_sneezing	chills	fatigue	7
Dengue	chills	high_fever	vomiting	7
Diabetes	weight_loss	restlessness	lethargy	7

4.4 The Impact and Implications in Medical Practice

By using NLP and big data technologies in the diagnosis and management of diseases, a series of significant benefits can be created that can contribute to an evolution of medicine due to several factors. First, NLP technologies allow detailed analysis and interpretation of information from patients' medical history, including reported symptoms and test results. This can lead to the early identification of diseases or medical conditions and their more accurate diagnosis.

Furthermore, through big data analysis, it is possible to keep track of each patient's medication and medical history. Consequently, treatments can be personalized based on each patient's needs and how they have responded to medications taken in the past.

By continuously monitoring health-related data as well as risk factors, big data technologies can help identify patients at high risk of developing chronic diseases and implement preventive measures or early interventions to help patients.

Another significant benefit is the optimization of resources and costs through the utilization of big data and NLP technologies. By accurately analyzing data, patients with urgent medical needs can be more easily identified, allowing medical staff to focus on patient care based on existing data [10].

Furthermore, the utilization of NLP and big data technologies can facilitate the interpretation of information from medical documents, potentially improving communication among healthcare professionals and resulting in better collaboration

between medical teams. The use of big data and NLP analyses in disease diagnosis and management contributes to the enhancement of the medical field [11].

4.5 Conclusion

This paper explores the intersection of NLP and big data in healthcare and its potential impact. Key concepts and techniques in the development and implementation of NLP solutions are presented, including Tokenization, Sentiment Analysis, Word Embeddings, and Big Data Analytics. The next chapter is dedicated to data collection and processing. Analysis was conducted on four datasets: the first consisting of diseases and their various symptoms, the second containing precautionary measures for specific symptoms, the third comprising concise disease descriptions, and the fourth comprising symptoms and their severity. The datasets were cleaned, normalized, and then transformed for further analysis.

In the initial phase of the analysis, the dataset underwent a tokenization process, followed by text processing techniques such as lemmatization and stopword elimination. The outcomes are represented through a word cloud, showcasing the most frequently encountered terms in the symptom descriptions of the analyzed dataset. Additionally, a bar chart displays the top 20 most common words in the symptom descriptions. This visualization aids doctors and researchers in analyzing medical data and focusing on the primary symptoms of patients.

Furthermore, the latent Dirichlet allocation algorithm was applied to each disease to identify the most common topics within each symptom. Additionally, the CountVectorizer module was utilized to transform the symptom data into a unified representation. TF-IDF vectorization was applied, and k-means clustering was employed to group the symptom descriptions. From the analysis, the silhouette score was calculated, yielding a value of 0.7567, indicating a well-defined clustering. In the final section, the impact and advantages of NLP and big data in medical practice were described.

Moreover, the fusion of big data with Natural Language Processing (NLP) techniques extends the horizon of possibilities even further. By leveraging NLP algorithms on vast medical datasets, healthcare professionals can extract valuable insights from unstructured textual data, such as patient records, clinical notes, and medical literature. This integration not only enhances the depth and breadth of analysis but also facilitates more nuanced understanding and interpretation of medical information, thereby fostering more informed decision-making and advancing the frontier of healthcare innovation.

Acknowledgements This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS-UEFISCDI, project number PN-IIIP4-PCE-2021-0334, within PNCDI III.

References

1. Anoop, E.B.T.W.S.D.R.D.H.H., Shah, D. : Natural language processing for disease phenotyping in UK primary care records for research: a pilot study in myocardial infarction and death. *J. Biomed. Sem.* **10**(2041–1480), 27–32
2. Denny, J.C.: Natural language processing improves phenotyping accuracy in electronic medical records. *Healthcare Inform Res* **30**, 12–21 (2011)
3. Kornelia Batkocorresponding, A.Ş. : The use of big data analytics in healthcare. *J. Big Data* **3**, 9–15 (2022)
4. Lane, H.H. : Natural language processing in action, Manning publications (2019)
5. Balac, N. : Big Data. in *Intelligent Internet of Things*, New York, Springer, pp. 315-356 (2020)
6. Anitha, R.T., Pillai, S. : Machine learning and deep learning in natural language processing. Boca Raton: CRC Press (2023)
7. Shi, Y. : Advances in big data analytics, Singapore (2022)
8. Kunal Nag, S.B.N.G. : Disease symptom prediction. Kaggle, 05 2020. <https://www.kaggle.com/datasets/itachi9604/disease-symptom-description-dataset/data>. [Accessed 04 2023]
9. Deepti Chopra, R.K. : Introduction to machine learning with Python. Singapore: Bentham Science (2023)
10. Anna Atefeh Farzindar, D.I.: Natural language processing for social media, Springer (2020)
11. Mazzei, M.: A machine learning platform for NLP in big data. in *SAI Intelligent Systems Conference* (2020)

Chapter 5

Enhancing the Involvement Level of Volunteer Students in the Academic Community: Efficient Usage of Digital Resources



Laurentiu-Gabrie Dinca and Vlad Diaconita 

Abstract This study analyzes the attitudes and behaviors of student volunteers within the academic community, focusing on the impact of the COVID-19 pandemic. Employing quantitative methods, including Chi-squared and t-tests, we explored how volunteer involvement correlates with academic progression and varies by gender with respect to the use of digital tools. We found that students' expectations for post-pandemic volunteering are significantly positive, indicating a strong commitment to continuing their volunteer activities. Satisfaction with remote volunteering during the pandemic was also positively linked to the continued use of digital platforms for volunteering. These findings point out the need to enhance digital platforms to improve volunteer interaction and engagement. Future research should incorporate qualitative methods to provide deeper insights into the motivations and experiences of student volunteers.

5.1 Introduction

Volunteering is defined as freely giving time to help others, traditionally through real-life “offline” activities [1]. Known as in-person volunteering, such actions involve dedicating effort and time resources to help organizations, or communities, such as family or neighborhood without the immediate use of digital tools and platforms. This kind of activity offers a unique set of benefits such as personal connection between peers who share the same ideology and cause and skill development [2, 3]. Unlike paid work, volunteers usually do not receive immediate financial compensation [4]. Yet,

L.-G. Dinca · V. Diaconita (✉)

Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, 010374 Bucharest, Romania

e-mail: diaconita.vlad@ie.ase.ro

L.-G. Dinca

e-mail: laurentiu.dinca@stud.ase.ro

their contributions significantly benefit others and even themselves in the long run, usually by enriching their cultural capital [5]. Students acknowledge various motivations and benefits from volunteering, yet instrumental and career-related motivations seem to be statistically indistinguishable from altruistic, value-driven motives and social or ego-defensive factors [6].

Online volunteering represents a new form of civic engagement emerging from the digital transformation of society. While traditional volunteering is experiencing a descending slope, even among younger demographics [7], online volunteering offers a potential approach to reinvigorate civic participation and possibly address societal inequalities. Online platforms can attract those not typically involved in offline volunteering, potentially democratizing civic engagement, and counteracting existing social stratifications. However, empirical evidence is somewhat limited whether online volunteering can truly mitigate societal inequalities by mobilizing new groups of participants [8].

As universities aim to create and develop a sense of community involvement among their students, it is essential to focus on encouraging their active participation beyond their classes and comfort zone. In this matter, volunteer associations encounter multiple challenges with today's education landscape. To address these challenges, organizations need to focus on creating meaningful, autonomous roles for volunteers, fostering supportive relationships, and maintaining regular appreciation communication.

Although involving yourself in different types of activities as a volunteer has been shown to be beneficial from both a personal and a professional perspective, constantly motivating students is a risk that must be properly assessed and monitored. There are a variety of factors, such as time constraints, lack of awareness or motivation, and competing priorities, that can interfere with the level of participation of a student volunteer in an academic community.

After 2 years since the COVID-19 pandemics ended, we can observe that it has completely changed the concept of volunteering in our academic community. Going back to 2019, our university center represented the home of nine student volunteering associations which were representing not only the students from the Bucharest University of Economic Studies, but from the entire community of students living in our capital city. In those days, the concept of an online or hybrid approach was not representing a feasible solution in terms of planning, scheduling, and carrying out the activities of a project.

In March 2020, the sudden transition from traditional classes to online learning, due to the COVID-19 pandemic, was facilitated by using digital resources and platforms, used to ensure academic continuity. At that time, extracurricular projects and activities were suspended, highlighting the unforeseen challenges faced by student volunteer associations. From this perspective, the coordinators and project managers, who could not imagine of a scenario in which the activities are not held on campus, were required to explore alternative methods of delivering qualitative workshops and development projects for students, in a virtual environment.

A response to these changes was made by trying to adapt all the activities that the volunteer organization was holding to the digital environment. An interesting

example would be Freshmen's Prom from 2020,¹ an event organized during the lockdown, exclusively held online, with the use of social media platforms, by applying the model of a live stream. Of course, such an action received both constructive and negative feedback. From a logistical approach, the conclusion was that such an event would not be reasonable to take place online for a second time.

This article examines the influence of digital resources on the participation of student volunteers within the academic community, especially in the context of the post-pandemic educational landscape. To guide our research, we formulate the following research questions based on specific aspects of student volunteer behavior and attitudes.

- **RQ1:** Is there a significant relationship between a student's year of study and their involvement in volunteering activities while attending the university?
- **RQ2:** Is there a significant association between gender identity and the belief that digital resources can improve student involvement in volunteering activities?
- **RQ3:** How do students envision their involvement in volunteering activities changing as we move forward from the pandemic?
- **RQ4:** How did the pandemic affect the level of involvement of student volunteers?
- **RQ5:** Is there a significant relationship between students' level of satisfaction with the remote/virtual volunteering activities during the pandemic and their intention to use a digital platform that connects student volunteers with different organizations on campus?

Following this introduction, each section builds on the initial discussions to extend our understanding of this issue. The literature review analyzes studies on the effectiveness of digital platforms and engagement strategies, pointing out how changes in volunteerism impact civic participation within universities. The Methods section details the research approach and techniques used to collect and analyze data from student volunteers at the Bucharest University of Economic Studies, emphasizing the integration and impact of digital tools. In the Results and Discussion section, we present and interpret the data, focusing on how digital resources influence volunteer participation across different student demographics and during distinct phases of their academic careers. The Conclusions and Future Work section summarizes the key findings and proposes directions for further research, particularly improvements in digital platform features and strategies to enhance interaction and cooperation among student volunteers. This organized exploration ensures a thorough investigation into the role of digital advancements in volunteer engagement, aiming to contribute to enhancing community involvement in higher education settings post-pandemic.

¹ Freshmen's PROM at Bucharest University of Economic Studies in 2020: <https://www.facebook.com/share/v/Ma5NcoXRj2X1cMLk/>.

5.2 Literature Review

The current section looks at strategies to improve participation among student volunteers within the academic community, focusing on the effectiveness of tools such as digital platforms. We have identified various perspectives and related studies in this area. The following review synthesizes the findings of relevant studies in student participation, the use of social networks in education, or service learning.

In the past 20 years, universities have increasingly recognized the value of student volunteerism, both for the impact it has on communities and the positive learning experiences it provides to student volunteers. However, encouraging a culture of volunteerism and ensuring sustained participation can be challenging [9].

Several studies have shown that volunteering experiences can cultivate a sense of civic responsibility, develop valuable skills such as teamwork and communication, and enhance student career prospects [10]. Therefore, by actively participating in volunteering, we can contribute to our general well-being by expanding our social networks, promoting self-confidence, and a sense of purpose.

Volunteering intentions increase when people are exposed to social information about similar groups (ingroups, “us”) compared to dissimilar groups (outgroups, “they”), an effect mediated by group identification and the anchoring effect [11, 12]. One study emerged from on a survey-based experiment in German universities in which 245 responses were collected and analyzed using ANOVA and t-tests [13]. Key controls included student status (whether participants were currently studying at the institution), stay at the university (continued enrollment in the following semester), gender (due to its potential influence on volunteering decisions), current volunteer status (as previous engagement predicts future contributions), and available free time (measured in hours per week). Senior citizens were presented as an outgroup for university students. The results of the study emphasize the need to strategically employ two aspects of social information: constructing reference group information to enhance group identification and willingness to volunteer, and providing information about group performance at an appropriate level to motivate potential volunteers.

Despite the documented benefits, several challenges can deter student volunteerism. Time restrictions, lack of awareness of opportunities, and difficulty finding placements that align with their interests can all contribute to low participation rates [10, 14]. Additionally, traditional methods for promoting volunteer opportunities, such as flyers and bulletin boards, may be less effective in reaching today’s digitally connected student population.

Digital resources have immense potential to overcome these challenges and improve student volunteer engagement. Online platforms can serve as a central hub for disseminating information about volunteer opportunities, allowing students to easily search and filter options according to their interests, availability, and skills [15]. Social networks can be used to create online communities for volunteers, fostering communication, collaboration, and a sense of belonging.

Furthermore, mobile applications can provide real-time updates, streamline volunteer management tasks, and offer gamification elements to increase motivation and engagement.

Although existing research acknowledges the potential of digital tools in education and volunteerism, there is a gap in understanding how these resources can be specifically used to optimize student volunteer participation within the academic community. This study aims to address this gap by investigating the specific features and functionalities of digital resources that can effectively attract, engage, and retain student volunteers within the university setting.

5.3 Materials and Methods

Stepping into the current section, we will concentrate in inspecting the factors influencing the level of participation of volunteer students from The Bucharest University of Economic Studies (ASE), with a specific focus on the efficient usage of digital resources. From this perspective, a quantitative survey was chosen to collect comprehensive and structured data from a diverse sample of volunteer students. Using such a methodology offers a practical approach to collect quantitative insights on student behavior, preferences, and perceptions related to the use of digital resources in the context of volunteering activities. For the purpose of this paper, the use of quantitative methodology offers us a reliable and efficient way to make a precise and objective statistical analysis to properly identify the possible patterns and trends in regard to the volunteering activities.

Moreover, based on the feedback received from the respondents, we will manage to propose the design and MVP² of a specialized digital platform, dedicated to student volunteers, orientated to efficient communication and sharing experiences.

Before launching the main survey, a pilot study was conducted, involving five volunteer students to pre-test the survey instrument. This pre-test phase allowed us to identify potential uncertainties or concerns in the survey questions, ensuring that the final instrument is valid and reliable for accurately capturing the intended data.

The research was carried out between 23 and 30 March 2024, within students of the Bucharest University of Economic Studies. During this period of time, 159 respondents have answered this survey. The educational level of the participants included all three existing levels of education in our community: bachelor's, master's, and doctoral degrees, while the age spectrum varied between 18 and 26 years old, reflecting a diverse range of respondents, having a different perception of volunteering and active involvement in the academic community.

The survey consisted of a series of 25 questions structured around the experience of volunteering before, during, and after college; between and after pandemics; the obstacles encountered during pandemics; the potential usability; and the impact of a dedicated social media platform for volunteers.

² Minimum Viable Product (MVP): What is it & Why it Matters (atlassian.com).

From this sample, approximately 8 in 10 respondents have done at least one volunteer activity, regardless of their type of environment. In the same manner, after screening those respondents, a percentage of 84% of the students were actually involved in volunteering activities in our academic community, and hence we can conclude that most of the students involved in activities as such have a necessary background to extend their expertise.

Asked whether digital resources may enhance student participation in volunteering activities, by using a Likert scale with five options, a percentage of 79.25% of respondents have positively appreciated the impact of this proposal, while 80.5% of respondents would be interested in using a digital platform that connects student volunteers within our academic community.

Furthermore, using a similar scale, respondents were asked to rate the importance of implementing a specific set of features for a digital platform: *Ability to search for diverse opportunities from our community, based on your personal interests and skills, progress tracking, communication tool to connect with your colleagues from other organizations (chat, video conferencing), open access to training materials and resources, calendar for upcoming activities, task management.*

At a close ratio, having the mean score of 4.43, and 4.40, respectively, out of the maximum score of 5.00, the volunteer students would find it very useful to have implemented a personalized search bar to filter out the events and projects in which they want to take part, based on their personal interests or skills, and to have open access to training materials. On the third place, the task management component was also appreciated as being relevant to the target audience.

Besides this ranking, all of the listed features have been evaluated, on average, with the minimum score of 4.00, or higher, but an interesting aspect is that the least preferred option was the progress tracking tool, designed to properly assess the actual level of involvement of the student volunteers in their activities, and the communication tool, an already existing component which might be easily associated with a basic social media application.

For RQ1, we employed a contingency table to explore relationships between the student's year of study and their participation in volunteering at the university, analyzed using the chi-squared test for independence. For RQ2, a similar method was used to examine the link between gender identity and perceptions about digital resources in volunteering. Additionally, for RQ3, we conducted a one-sample t-test to determine if students' expectations about their future involvement in volunteering activities post-pandemic were significantly different from a neutral standpoint (a score of 3 on a Likert scale). Similarly, for RQ4, we conducted a comparison between the level of involvement of students before and after the pandemics, in order to better understand the impact of the medical crisis on the behavior of young adults. For RQ5, we conducted an inferential analysis to assess the relationship between the satisfaction level of students while attending remote activities during the pandemic and the interest in using a digital platform.

5.4 Results and Discussions

Our statistical analysis provided insights into the volunteering behavior of university students across different demographics and periods. For **RQ1**, the chi-squared test showed a significant relationship between the year of study and volunteering involvement (chi-squared statistic: 9.121, $p = 0.0277$), indicating that students' participation in volunteering activities varies significantly with their academic progression.

When addressing **RQ2**, we found a significant association between gender and perceptions of digital tools for volunteering (chi-squared statistic: 16.195, $p = 0.0397$). This suggests that there may be gender-specific preferences or experiences with digital volunteer platforms, emphasizing the need for targeted digital resources that cater to diverse student groups.

For **RQ3**, the one-sample t-test showed that students' expectations about their post-pandemic volunteering involvement were significantly more positive than neutral (t-statistic = 8.017, $p < 0.001$). This finding implies a strong inclination toward increasing or maintaining their volunteering activities after the pandemic, possibly reflecting a renewed sense of community and social responsibility.

When analyzing the output for **RQ4**, it is shown in Fig. 5.1 that for first- and second-year students there is a decrease in their level of involvement. This could be explained by the fact that they were sophomores during the pandemics, so they had limited resources in terms of volunteering activities and personal development opportunities.

Nonetheless, we observed an increase of approximately one point in terms of involvement level for third-year students, the generation, who entered the university by using a virtual environment back in 2021, was eager to explore the benefits of leveraging digital platforms for both academic and extracurricular means. This positive shift in engagement emphasizes the enthusiasm of students accustomed to virtual learning environment to embrace technology as a tool for collaboration and community impact.

Another interesting aspect worth mentioning is that, in the case of our fourth-year students, exclusively coming from the Law Faculty, all of the respondents have graded their involvement level after the pandemic with the highest maximum score of 5.0 and this might be explained by the fact that they were actively involved in creating a new community of students volunteer within their specialty and giving birth to our newest volunteering association: ASFD.³

For **RQ5**, we identified a statistically significant positive relationship between the satisfaction of the students after performing remote activities during the pandemic and their interest in using a digital platform for volunteer activities ($r = 0.426$, $p = 0.0004$). This suggests that a higher level of participation in volunteering activities is associated with a greater interest in leveraging technological instruments for volunteer purposes.

These findings highlight the importance of considering academic progression and gender differences in designing and promoting volunteering opportunities. They also

³ ASFD | Law Faculty Association (Asociația Studentilor din Facultatea de Drept).

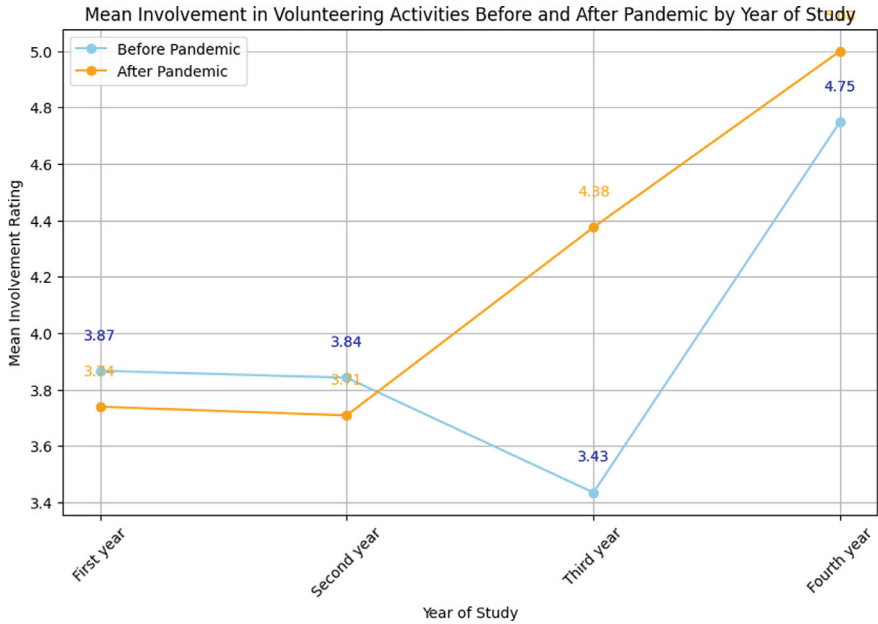


Fig. 5.1 Comparison between involvement level before and after the pandemic

suggested that after the pandemic, universities have opportunities to leverage positive sentiments toward volunteering by facilitating and improving digital platforms for student engagement.

5.5 Conclusions and Future Work

This article investigated the current perceptions and preferences of student volunteers active in the academic community, assessed the impact of the pandemic on their volunteer activities, and compared the behavior of student volunteers from 2020 to 2024. This is central not only to developing the sense of appreciation among students but also to leveraging their hard and soft skills.

The results of the quantitative research suggest that a dedicated digital platform would have a positive impact on volunteer students. This indicates that new generations are likely to adapt to a hybrid environment, where onsite activities are going to be planned and discussed using a certified digital solution. However, we must also consider the possibility that we might not observe significant improvement in terms of interactivity between volunteers from different groups, or associations.

While our study suggests that the platform would provide numerous benefits, it is important to assess the potential challenges. One significant concern is whether the digital solution will effectively enhance interactivity.

From the platform design perspective, the lack of interactivity might be due to limitations in the proposed platform's features, such as the absence of discussions forums, or project collaboration tools that encourage interaction among diverse groups. Limited adoption of the platform by student volunteers could also diminish the potential for interactivity. In such cases, efforts may be required to increase awareness and promote greater adoption of the platform among the target audience. Additionally, the digital platform may not address deeper underlying issues that hinder interactivity, such as a lack of incentives for intergroup collaboration or limited opportunities for joint projects.

Without improvements in this area, the platform's overall impact could be limited. To address these potential limitations, future research could explore several avenues. Conducting user experience research with student volunteers could prove valuable in identifying functionalities that would foster interactivity between different volunteer groups, thus optimizing platform design. Furthermore, developing strategies to promote the platform and incentivize student volunteers to actively use it could encourage broader adoption. Investigating the organizational structures and policies that might hinder intergroup collaboration among student volunteers could also provide insight.

Universities have the potential to play a pivotal role in addressing these challenges. The administration could support platform optimization by providing the necessary funding for the user experience research and collaborating with developers or architects to ensure the platform meets the needs of student volunteers. In addition, the university representatives could also launch awareness campaigns inside the campus to properly promote the platform, highlighting its benefits, and encouraging its use among the students. Moreover, by organizing workshops and training sessions with specialists from the academic community, including alumni, the universities can equip the students with the skills needed to effectively use the platform. Finally, universities can facilitate intergroup collaboration by creating internal policies that encourage joint projects and provide necessary space for students to meet and collaborate within the campus.

To conclude, while the introduction of a dedicated digital platform for students' volunteers shows a great opportunity in mitigating to a hybrid environment for planning and holding the volunteering activities, careful consideration must be given to its design and implementation. By addressing the potential challenges and leveraging the support of university administrations, we can create a more interactive and collaborative space for student volunteers. This will not only enhance their volunteering experience, but also build a stronger sense of community and shared purpose among diverse groups.

As we move forward, continuing the research from a qualitative approach and engaging with the student community will be essential in refining the platform to meet their evolving needs, ensuring its long-term success and impact on volunteerism in the academic community.

References

1. Wilson, J.: Volunteering. *Annu. Rev. Sociol.* **26**(1), 215–240 (2000). <https://doi.org/10.1146/annurev.soc.26.1.215>
2. Kameråde, D.: The same but different: a comparison between family volunteers, other formal volunteers and non-volunteers. *Voluntary Sector Review* **14**(1), 189–200 (2023). <https://doi.org/10.1332/204080521X16572238373614>
3. Lee, Y., Brudney, J.L.: Participation in formal and informal volunteering: implications for volunteer recruitment. *Nonprofit Manag. Leadersh.* **23**(2), 159–180 (2012). <https://doi.org/10.1002/nml.21060>
4. Snyder, M., Maki, A.: Volunteerism, Psychology of,” in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, pp. 268–272. <https://doi.org/10.1016/B978-0-08-097086-8.22021-7> (2015)
5. Sikora, J., Green, J.: Gifts as gains? The impact of volunteering on young people’s educational and occupational attainment in Australia. *Aust. J. Educ.* **64**(2), 177–194 (2020). <https://doi.org/10.1177/0004944120910813>
6. Smith, K.A., Holmes, K., Haski-Leventhal, D., Cnaan, R.A., Handy, F., Brudney, J.L.: Motivations and benefits of student volunteering: comparing regular, occasional, and non-volunteers in five countries. *Canadian J. Nonprofit Soc. Econ. Res.* **1**(1) (2010). <https://doi.org/10.22230/cjnser.2010v1n1a2>
7. Forner, V.W., et al.: Predictors of turnover amongst volunteers: A systematic review and meta-analysis. *J. Organ. Behav.* **45**(3), 434–458 (2024). <https://doi.org/10.1002/job.2729>
8. Ackermann, K., Manatschal, A.: Online volunteering as a means to overcome unequal participation? The profiles of online and offline volunteers compared. *New Media Soc.* **20**(12), 4453–4472 (2018). <https://doi.org/10.1177/1461444818775698>
9. Nichol, B., Wilson, R., Rodrigues, A., Haighton, C.: Exploring the effects of volunteering on the social, mental, and physical health and well-being of volunteers: an umbrella review. *Voluntas* **35**(1), 97–128 (2024). <https://doi.org/10.1007/s11266-023-00573-z>
10. Davis, T.C.: How college affects students (Vol. 2): a third decade of research (review). *J. Coll. Stud. Dev.* **47**(5), 589–592 (2006). <https://doi.org/10.1353/csd.2006.0055>
11. Epley, N., Gilovich, T.: The anchoring-and-adjustment heuristic. *Psychol. Sci.* **17**(4), 311–318 (2006). <https://doi.org/10.1111/j.1467-9280.2006.01704.x>
12. Hysenbelli, D., Rubaltelli, E., Rumiati, R.: Others’ opinions count, but not all of them: anchoring to ingroup versus outgroup members’ behavior in charitable giving. *Judgm. Decis. Mak.* **8**(6), 678–690 (2013). <https://doi.org/10.1017/S193029750000471X>
13. Henninger, M., Rottler, M., Helmig, B.: Under pressure: social information on reference groups and its influence on volunteering intention—a survey experiment. *Nonprofit Volunt. Sect. Q.* **53**(2), 510–535 (2024). <https://doi.org/10.1177/08997640231172209>
14. Thoits, P.A.: Motivations for peer-support volunteering: social identities and role-identities as sources of motivation. *Nonprofit Volunt. Sect. Q.* **50**(4), 797–815 (2021). <https://doi.org/10.1177/0899764020983897>
15. Thomson, A.M., Smith-Tolken, A.R., Naidoo, A.V., Bringle, R.G.: Service learning and community engagement: a comparison of three national contexts. *Voluntas* **22**(2), 214–237 (2011). <https://doi.org/10.1007/s11266-010-9133-9>

Chapter 6

Intelligent Models in Power Delivery Management for Economic Damage Computation



Mariana Dumitrescu

Abstract The presented article shows economic aspects of the power delivery systems due to loss of functional continuity. Complex power systems need to maintain their fault-tolerant operational stage. On this purpose intelligent models are used to compute the availability, taking in account the devices dependency (dependability). The intelligent models are meant to take in account the informatics/automation system action, designed to maintain the fault-tolerant stage. An algorithm is built for delivery system steady-state performance computation. The measurable results are presented for a real power delivery case study. Delivery power flow and energy losses are used for the economic study with the economic damage computation. The 20 kV power station was monitored and supervised for 1-year interval.

6.1 Introduction

The energy continuity must be analyzed both from the point of view of the distributor—through the component elements of the distribution network- and from the point of view of the receivers—as sources of disturbances. A good quality of the product and the service in the electrical delivery networks is obtained through the availability (dependability) of the system.

Availability is one of the key parameters in electricity distribution networks because continuity in the supply of consumers is a priority. The availability values are very high, 99.98% for a domestic consumer and 99.996% for an industrial consumer. These high levels of availability are achieved by configuration of electrical networks and appropriate delivery maintenance management procedures [1].

Continuity in supply results from the behavior of the entire energy system, which must satisfy adequate static and dynamic conditions, through the existence of sufficient sources in the system, to ensure the consumer's demand and an appropriate

M. Dumitrescu (✉)
Dunarea de Jos University, Galati, Romania
e-mail: Mariana.Dumitrescu@ugal.ro

response to the dynamic processes that occur, when the system passes from one state to another other. The evolution of an installation is characterized by a sequence of states of operation, failure, stoppages for unplanned repairs or planned repairs, the moments of occurrence as well as their durations being of a deterministic or random nature [2].

The assessment of safety in the operation of energy installations can only be done with the help of numerical indicators of a probabilistic nature, necessary to express different aspects of the system's behavior: the average duration of operation in a time interval $(0, t)$; the average failure duration in a time interval $(0, t)$; the average number of defects in a time interval $(0, t)$. The calculation of these quantities is carried out by the methods proposed by the reliability of the systems, capable of reproducing, as faithfully as possible, the states the system goes through in a time interval [3].

6.2 Intelligent Models in Power Systems Delivery Continuity Management

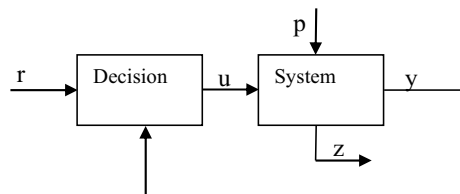
6.2.1 Neural Networks

The main specific operational features of the electricity production, transport, and delivery process are the voltage and frequency of the delivered energy which must be kept within certain precise limits close to the nominal values, as there is no possibility of storing electricity. The generators must constantly follow the power variations requested by the consumers; the location, separation, and remedy of the damages that have occurred must be done as quickly as possible to avoid their extinction.

The management of the systems expresses the requirement that the quality quantity follows a reference quantity r , specified a priori in the context of the disturbance p . This is achieved by elaborating a decision regarding the appropriate evolution of the command quantity u , elaboration resulting from the processing of the measured information y and eventually, when possible, of the disturbance p . The structure of such a regulator is shown in Fig. 6.1 [4].

Neural networks represent the object of study of the interdisciplinary field, located on the border between engineering, cognitive sciences, computer science, and fundamental sciences. The main topics are the presentation of the basic problems of the design of automatic regulation systems, the management carried out with the help

Fig. 6.1 A general control system



of neural networks, and some major problems that arise when implementing them within the power systems.

The use of neural networks (NN) for the optimal control is possible if a priori information about the process is known and it can be described deterministically. If estimated information is obtained, the designed regulator can reach an optimum. As the decision block accumulates more information about the unknown functions or parameters, the control law will be modified with the updated information in order to improve the system's performance [5, 6]. The online learning model for a process considered static is presented in Fig. 6.2. The design of power systems requires conditions of stability, robustness, and a response with good dynamics, conditions that cannot be obtained at the same time for a wide class of systems. The objective of the system is to obtain the desired output of the process r , provided that the input of the process that causes this output is known. For a dynamic process, the trained network must be provided with a sequence of recent outputs in addition to the current output of the process. This will allow the current state of the process to be reconstructed based on the most recent output history (Fig. 6.2) [7].

The neural system can have a neural regulator with 6-8-2 neurons, respectively, on the three layers, input, hidden, and output, and a modeling network with 12-14-2 neurons, respectively. The two subnets work with different durations during

Fig. 6.2 NN static (a) and dynamic architecture (b)

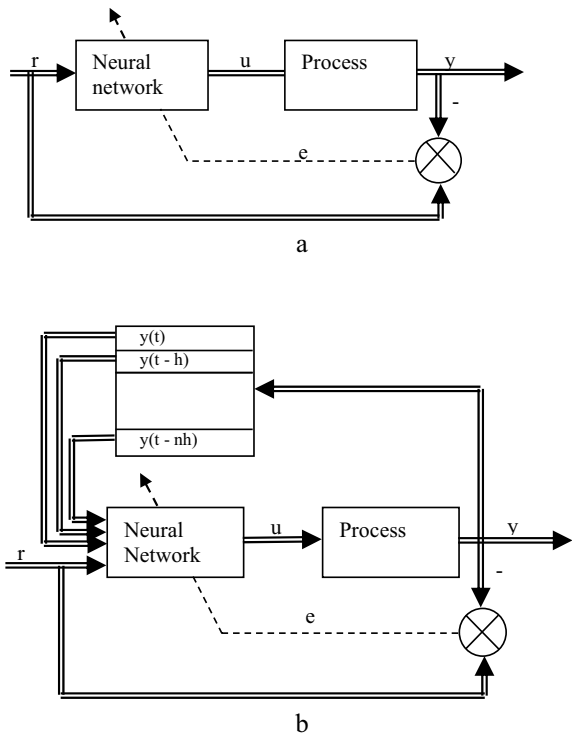
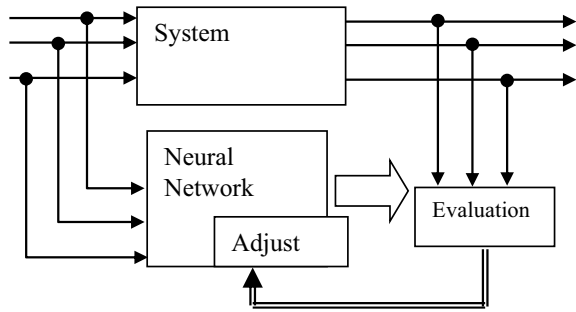


Fig. 6.3 NN dynamic process identification



a sampling interval. In the first moment, the modeling network works, using the error between the output of the process and its output. Then the weights in the neural controller are modified as adaptive parameters, based on the error between the desired output of the driven system and the output of the previously trained model network.

In this second moment, the propagation of errors for the training of the regulator network is done through the model network, after its training, keeping its weights obtained by training at the first moment.

Parallel processing has provided a new direction in control systems engineering. The role of computerized control and the use of the computer in the design process has increased, implicitly leading to total dependence on it. The general purpose should be oriented toward a distributed and reliable (fault tolerant) control such as, for example, testing at numerous points of a large system such as the power system [8].

The need to obtain a high performance for a control system designed to work in real time implies the use of a new architecture that meets the requirements. The control systems must be parallel because they work in a real world, where parallelism is a natural phenomenon of everyday life. A dynamic system is described by differential equations that often have a laborious, sometimes almost impossible solution. Although most neural networks are static, a dynamic process can be modeled in discrete time by providing the network with training examples consisting of past inputs and outputs of the process (Fig. 6.3) [9].

6.2.2 Decision Trees

One of the most well-known automatic learning methods is the decision trees method, which uses a restrictive representation of logical propositions specially built for learning, this method being very attractive for three reasons: offers hypotheses that give good results in new situations, is unused in the construction process; has a reduced calculation time, which is directly proportional to the number of examples and attributes that describe an example; and provides a representation that is easily

understood by human operators. An automatic learning system includes three stages: generating the database, extracting the information, and processing it [7].

Three databases which are used in an automatic learning system are:

- the primary database;
- the training base that contains the relevant data on the basis of which the automatic learning system is built;
- the knowledge base containing the input/output interdependence relations “learned” by the intelligent system.

Since MS SQL 2000 allows the definition and use of relational databases, the implementation of the three databases is very easy, as they can be incorporated into one. The link between the three modules is achieved by defining interdependence relationships and stored procedures, through which data from a certain module is modified based on changes in another module.

The training base contains data about different states of the power system. These are made up of training examples, each example being characterized by the state of the power system (states of power lines, values of active/reactive power generated/consumed, state of adjustment means, and so on.) and by the value of an indicator for voltage stability assessment [7, 9].

The knowledge base contains the decision tree built based on the training set (the examples from the training base). Each element of this database is described by the node number, the attached test, the related training set, and its two children. By defining the interdependence relationships between the three bases and using triggers and calculated columns, a change in an input value (for example, the voltage value in a node) will instantly generate a change in another value (for example, the voltage stability index). Thus, the use of an automatic learning system for the analysis of the voltage stability of the SEE is very simple and fast, lending itself to an online implementation [7, 8].

A decision tree is a binary or multipath tree with the property that its nodes contain many examples. The root of the tree contains the entire set of examples, and its leaves contain examples that are either only positive (“YES” node) or only negative (“NO” node). Information extraction is the essence of building a machine learning system. Starting from the information stored in the database, in this stage the intelligent system that can learn is built (a decision tree, a regression tree, an artificial neural network, and so on.) [4].

6.2.3 Discrete Models of Petri Nets

To evaluate the availability of systems, it is necessary to use appropriate methods and to know their modeling power. Within Markov models, the model of continuous-time Markov Chains (MC) is equivalent, in terms of modeling power, to the model of Generalized Stochastic Petri Nets (GSPN). The use of these models in the case of

complex systems is limited by the large size of the state space and the large volume of calculations.

Another limitation, specific to the MC model, is the difficulty in creating the state space, for the case of complex systems. This difficulty can be removed by using the model of generalized stochastic Petri nets, the accessibility tree associated with the GSPN being isomorphic to the MC system.

The model is specific to systems in whose evolution there are states with random duration, of exponential distribution, and transient states of negligible duration. They use stochastic, exponentially distributed, and instantaneous transitions (with zero timing). In system dependability assessment, they present two subnets: the event modeling subnet (SME), consisting of stochastic transitions (failure, restoration); the logical subnet (SL), giving the logical performance conditions, as well as the functional dependency consisting of instantaneous transitions.

The SL logical subnet creates several difficulties in using the model, as follows:

- the accessibility tree (AT) associated with PN presents several transient states (states that validate the execution of instantaneous transitions from SL). These states do not belong to the set of distinct steady states for the evolution of the system, they are additional states, which create an increased degree of complexity in the assessment of dependability (like protection/automation actions, which are mandatory to keep the power system fault tolerant);
- has a large number of places;
- for the modeling of the system with restoration, the additional construction of a logical subnet SL' is necessary, complementary to the logical subnet of the system without restoration SL (which models the logical restoration conditions of the system). For this reason, it is difficult to realize the PN for complex systems with restoration.

Figure 6.4 presents the delivery MV system with 13 transformer posts (PT) and the distances [m] of the cable sections between each two of the PT. To evaluate the system availability/dependability measures, with the help of PN, a calculation program was developed, which is structured as follows [1, 10]:

- the data input interface: the PN structure with locations and transitions, the initial state of the PN, the logical conditions for achieving the performance of the modeled system;
- generation of the AT associated with the PN model; elimination of transient states and generation of reduced AT isomorphic with the MC;
- generating the matrix of transition probabilities associated with MC and solving the system of algebraic equations to determine the vector of the steady-state probabilities;
- realization of the disjoint sets of states “Success”, “Partial Success”, “Refuse”, and the calculation of dependency measures.

With the help of the program, the following results are obtained [1]:

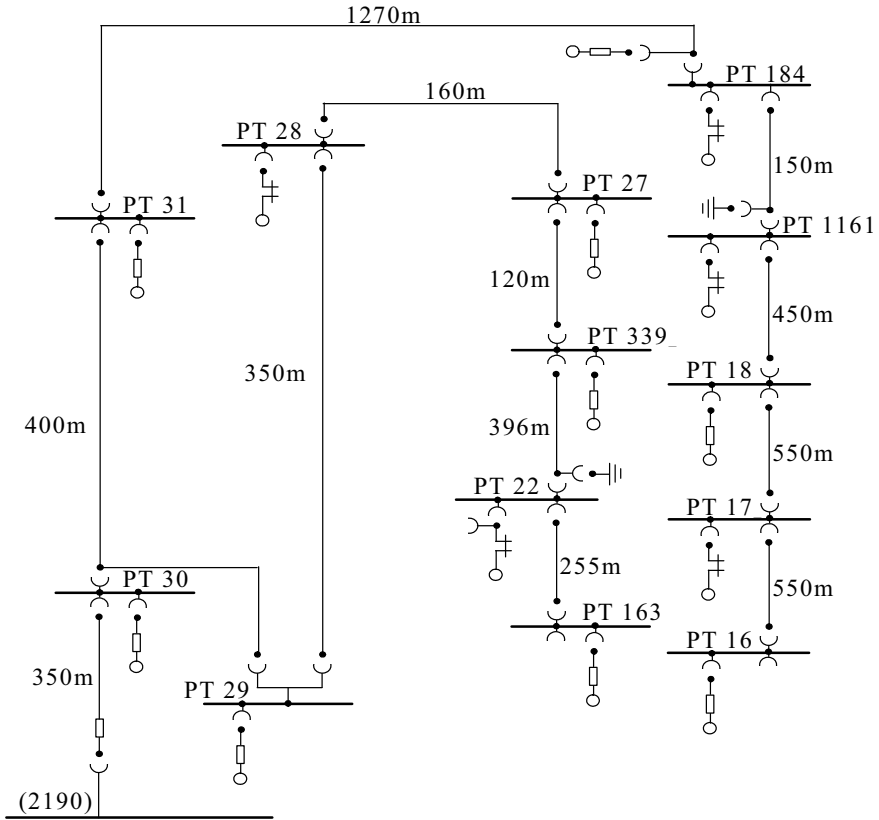


Fig. 6.4 Initial data for the delivery system

- the structure of the stationary states of the MC (reduced AT markings) with the links between them;
- coefficients of the system of algebraic equations associated with MC;
- the vector of state probabilities is solved by the Seidel–Gauss method;
- the disjoint sets of “Success”, “Partial Success”, and “Refuse” states corresponding to the success states of the system;
- the Partial success probability of some significant points in the system PSP (availability of some parts of the system);
- the Refuse probability of the system PR;
- the Success probability PS (availability/dependability of the system).

The results are presented in Fig. 6.5 and Table 6.1.

0	93475321298889	10	Succes	↑
290	1.32272013317687E-23	9.7	Refuz	
2	7.05968704909145E-03	9.4	Succes	
3	7.05968704909146E-03	9.4	Succes	
4	1.08438931889568E-03	9.2	Succes parțial 1	
5	1.08438931889568E-03	9.2	Succes	
6	4.68196565453226E-02	9.6	Succes	
287	3.58772565679233E-23	9.7	Refuz	
8	1.33394872626873E-05	9.4	Succes parțial 1	
9	5.33180100785619E-05	9.4	Succes	
10	5.28820644822426E-06	9.2	Succes parțial 1	
11	8.18980788127221E-06	9.3	Succes	↓

Rezultate finale		leșire	
Refuz	8.01005017915747E-13		
Succes parțial 1	1.16842519087006E-03		
Succes parțial 2	1.36841708521426E-06		
Succes	.998830206391244		
TOTAL M	1		
TOTAL P	1		

Fig. 6.5 Delivery system steady-state probability computation window

6.3 Using Intelligent Models in Power Delivery Economic Damage Computing

Using the results provided by the PN program package and the analysis of the continuity in the supply of electricity to the consumers for different types of distribution networks Radial (RDR), Ring (RDI_A and RDI_B), and Curled (RDB), for an operation time of 1 year, the Safety and Efficiency can be evaluated. Performance indicators can be computed as the following (Table 6.2).

The continuity indicators refer to long-term interruptions. They are computed using the proposed algorithm and the models and are presented below:

- System Average Interruption Frequency Index (SAIFI)—the average number of times that a customer is interrupted during the year;
- System Average Interruption Duration Index (SAIDI)—the sum of the customer interruptions, divided by the total number of customers served;
- Customer Average Interruption Frequency Index (CAIFI)—the average number of interruptions per customer interrupted per year;
- Customer Average Interruption Duration Index (CAIDI)—average interruption duration time for those customers that experience an interruption during the year.

The average frequency of short-term interruptions number takes the power delivery system short-term interruptions into account and needs to develop the computation algorithm for the MAIFI indicator (Momentary Average Interruption

Table 6.1 Availability/dependability computation–power delivery system example. Steady-state number: 48; transitory state number: 27; total number: 72

Steady-state probabilities	
Refuse probability 1.2985473085949e-03	Partial success 1 4.88768810567541e-03
State probability P(i)	State probability P(i)
P(65) = 5.16978540947285E-15	P(2)s = 3.63124688447074E-03
P(62) = 1.88984317368446E-14	P(4) = 1.11174568791528E-03
P(59) = 1.88984317368446E-14	P(16) = 1.09128519411818E-04
P(57) = 7.00791097852979E-14	P(18) = 3.34108823579135E-05
P(54) = 5.83774188084809E-12	P(25) = 1.63979424209188E-06
P(51) = 5.83774188084809E-12	P(27) = 5.02040774140705E-07
P(48) = 2.21476342702874E-11	P(35) = 1.09458323456786E-08
P(33) = 7.27309351053396E-09	P(37) = 3.35067108157916E-09
P(44) = 2.21476342702874E-11	
Partial success 2 2.41259171959008e-05	Success probability 0.998701478647951
State probability P(i)	State probability P(i)
P(8) = 1.37226965319673E-	P(0) = 0.960886507009012
P(9) = 4.20135264296932E-06	P(3) = 3.63124688447074E-03
P(11) = 4.20135264296932E-06	P(5) = 1.11174568791528E-03
P(13) = 1.28628976017256E-06	P(6) = 2.88771681975642E-02
P(53) = 3.87507557837895E-12	P(15) = 4.33917462599513E-04
P(50) = 1.26582423909109E-11	P(17) = 1.09128519411818E-04
P(29) = 4.12403105032493E-07	P(19) = 3.34108823579135E-05
P(30) = 1.26261692963721E-07	P(24) = 2.89787310033519E-06
P(31) = 1.26261692963721E-07	P(26) = 1.63979424209188E-06
P(32) = 3.86563896507217E-08	P(28) = 5.02040774140706E-07
P(47) = 1.26582423909109E-11	P(36) = 1.09458323456786E-08
P(39) = 6.19686266953948E-09	P(38) = 3.35067108157916E-09
P(40) = 1.89723673821917E-09	
P(41) = 1.89723673821917E-09	
P(42) = 5.80859639564858E-10	
P(46) = 4.13498657660655E-11	

Table 6.2 Power delivery continuity indicators

	M[v(T _P)]	M[α(T _P) × E+04	M[β(T _P)]	SAIFI	SAIDI min	CAIDI min	MAIFI
RDR	24.3509	7.970515	294.832	8,12	98,24	12,23	79,45
RDI_B	10.8902	7.989541	104.585	3,63	34,86	14,93	27,34
RDI_A	10.5931	7.986165	103.883	3,63	34,86	14,93	27,34
RDB	1.66225	7.998733	12.6627	0,55	4,33	8,56	4,92

Frequency Index). MAIFI is calculated by dividing the total number of momentary customer interruptions by the total number of customers served by the utility.

The system-specific availability metrics, presented in Table 6.2, are computed using the general steady-state performance results from Table 6.1. Mean operational time is

$$M [\alpha(T)] = P_S \cdot T \tag{6.1}$$

Mean failure time, during T period when power system will not supply the consumers and the system it is in an outage, is

$$M [\beta(T)] = P_F \cdot T \tag{6.2}$$

Mean failure interruptions number is

$$M [v(T)] = \left[\sum_{i \in S} \left(P_i \cdot \sum_{j \in R} q_{ij} \right) \right] \cdot T \tag{6.3}$$

where

- S is the success steady-state set;
- R is the refuse steady-state set;
- q_{ij} is the transition intensity from the success state i in the failure state j;
- P_i is the success i state probability.

The calculations are made for a 20 kV power station, for which active and reactive energy circulation measurements were made over a period of 1 year, both in summer and in winter (Tables 6.3 and 6.4).

The active and reactive power delivered on a circuit, for an apparent power of 1 KVA flow, taking in consideration the power factor $\cos\phi$ of the network, are given by the relations:

$$P = 1 \cos\phi = 0.855 \text{ [KW]}, Q = 1 \sin\phi = 0.517 \text{ [KVAR]} \tag{6.4}$$

Active Ensa energy and reactive Ensr energy which was not delivered to the system level (for $T_u = 80000$ h the time of the power station use) is

Table 6.3 Active and reactive energy not delivered because of the power interruptions

	SAIDI min	CAIDI min	Ensa KWh	Ensr KVARh	Enca KWh	Encr KVARh
RDR	98.24	12.23	11863.87	3822.792	1476.945	475.9033
RDI_B	34.86	14.93	4209.837	1356.5	1803.008	580.9679
RDI_A	34.86	14.93	4209.837	1356.5	1803.008	580.9679
RDB	4.33	8.56	522.9086	168.4924	1033.741	333.0934

Table 6.4 The damage because of the not delivered energy due to power interruptions

	Loss cost Euro/h	SAIDI min	CAIDI min	System annual damage Euro/year	Consumers annual damage Euro/year
RDR	3.800.000	98.24	12.23	6221866	774566
RDI_B		34.86	14.93	2207800	945566
RDI_A		34.86	14.93	2207800	945566
RDB		4.33	8.56	274233	542133

$$\text{Ensa} = \text{Ea} \cdot \text{SAIDI}/\text{Tu} \cdot 60 \quad (6.5)$$

$$\text{Ensr} = \text{Er} \cdot \text{SAIDI}/\text{Tu} \cdot 60 \quad (6.6)$$

where

- Ea is the active energy delivered on 1 year through the stations;
- Er is the reactive energy delivered/year through the station).

Active Enca and reactive Encr energy which was not delivered to the consumers are

$$\text{Enca} = \text{Ea} \cdot \text{CAIDI}/\text{Tu} \cdot 60 \quad (6.7)$$

$$\text{Encr} = \text{Er} \cdot \text{CAIDI}/\text{Tu} \cdot 60 \quad (6.8)$$

The costs of active energy and reactive energy transported for an apparent power of 1 KVA, in a period of 1 year, are given by the relations:

$$\text{CWp} = \text{Cp Tu} \cos\varphi = 1592.86 \text{ ron} \quad (6.9)$$

$$\text{CWq} = \text{Cq Tu} \sin\varphi = 135.24 \text{ ron} \quad (6.10)$$

where

- Cp = 0.3105 ron is the cost of 1 KWh delivered power;
- Cq = 0.0436 ron is the cost of 1 KVARh delivered power;

The costs of active energy and reactive energy for the analyzed 20 kV power delivery station to system level and consumer level are computed by (11, 12).

$$\text{Cnsa} = \text{Ensa} \cdot \text{Cp}, \text{Cnsr} = \text{Ensr} \cdot \text{Cq} \quad (6.11)$$

$$\text{Cnca} = \text{Enca} \cdot \text{Cp}, \text{Cncr} = \text{Encr} \quad (6.12)$$

The damage because of the not delivered active energy and reactive energy due to power system interruptions are shown in Table 6.4.

6.4 Conclusions

The proposed algorithm and computing program provide a powerful tool for analyzing delivery power system performance as well as availability/dependability measure computation. The program is extensible by including other reliability analysis applications.

The results obtained with this program system confirm:

- the correctness of the proposed GSPN models;
- the flexibility of the model construction;
- the correctness of the results evaluations;
- the operational accessibility for large systems;
- the possibility of evaluating the economic performance.

Due to the large amount of information, as well as the complexity of the processing, the execution time can be extended due to the operations of storing and retrieving the information on disk, during the processing stages. The execution time becomes acceptable (on the order of minutes) when analyzing complex systems with a large number of MC states (thousands of states), under the conditions of a large RAM memory and a maximum operating speed.

References

1. Dumitrescu, M.: Marine Industry Automation Systems Simulation. Scientific Bulletin Mircea cel Batran Naval Academy; Constanta vol. 22 (2019)
2. Kahlen, J.N., Muhlbeier, A., Andres, M., Unger, D., Kleinekort, K.: Synthetic data – a solution to train diagnostic systems for high-voltage equipment without fault-condition measurements. CIGRE Sci. Eng. vol. (2022)
3. Li, W.Q., Zhang, C.: Application of combination forecasting model in aircraft failure rate forecasting. Comput. Intell. Neurosci. (2022)
4. Ortiz, L., Aguila, A.: Modeling and control of an automated system in surgical microscopy based on petri nets. IEEE Lat. Am. Trans. **14**(1), 96–101 (2016)
5. Haidar, A., Mohamed, A., Jaalam, N., Khalidin, Z., Kamali, M.: A neuro-fuzzy application to power system. In: International Conference on Machine Learning and Computing Proceedings (2009)
6. Yang, M., Deng, S., Zhao, Z.: Online control method of small-and medium-sized electromechanical equipment based on deep neural network. J. Robot. (2022)
7. Kahlen, J.N., Andres, M., Moser, A.: Improving machine-learning diagnostics with model-based data augmentation showcased for a transformer fault. MDPI J. Energ. **14**(20) (2021)
8. Li, L.: Verification of the improvement effect of artificial intelligence computer control technology on electrical automation control. Adv. Multimed. (2022)

9. Mao, J., Yang, Z., Lu, J.: Research on signal analysis and diagnosis of chain drive system of the scraper conveyor guided by artificial intelligence algorithm. Elsevier Optik vol. 271 (2022)
10. Zhou, X., Wang, F.: Towards applicability evaluation of hazard analysis methods for autonomous ships. J. Ocean. Eng. Elsevier Ltd. **214** (2020)

Chapter 7

Smart City Parking Applications in Romania: An Analysis and Conclusions



Răzvan Daniel Zota and Mihai Adrian Lungu

Abstract The concept of smart cities (SC) has emerged as a focal point in contemporary urban discourse, representing a paradigm shift in urban development, characterized by the strategic deployment of digital technologies to improve efficiency, sustainability, and quality of life for residents. Our research is based on the hypothesis that the role of information technology applications in a smart city is becoming increasingly important. In this context, we analyze and compare four existing smart parking applications in Romania based on key characteristics. Furthermore, we proposed the integration of artificial intelligence concepts within these types of applications, noting that none of the analyzed applications currently incorporates such technology. Finally, we present the potential limitations of such applications and the possible improvements that could be made to them.

7.1 Introduction

The concept of SC has increasingly captured academic and practical attention in recent years, but a consensus on its precise delineation remains elusive. However, there is a consensus on harnessing and implementing information and communication technology (ICT) concepts to improve urban governance and operational efficiency. Also, an SC today appears in the foreground as a “modernist approach to facilitate the lives of citizens” [1]. Therefore, smart cities are built on information and communication technologies to enable a wide range of advanced services. As an overall idea, “technologies alone cannot make a city smart or smarter” [1], it is only the right way ICT services are used to serve all citizens of the city.

R. D. Zota · M. A. Lungu (✉)
Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania
e-mail: adrian.lungu@csie.ase.ro

R. D. Zota
e-mail: zota@ase.ro

One of the many problems of a large city area is the problem of parking spaces. Parking problem is very important because parking vehicles in densely populated areas poses numerous challenges, often resulting in stress and frustration for drivers who navigate congested streets. The process can become tedious and monotonous as drivers contend with limited parking availability, disorderly or poorly managed parking arrangements, and a lack of access to relevant parking information at their destination. All these factors contribute to delays in finding suitable parking spaces, resulting in wasted time, fuel, and energy, as well as increased environmental pollution. In conclusion, “the transformation of existing infrastructure into smart cities cannot ignore the smart management of parking systems” [2].

In this context, our study specifically focuses on examining existing smart parking applications (used in Romania) in the context of smart cities and delineates strategies for the improvement and refinement of such applications.

7.2 Literature Review

Creating new parking spaces is not always an effective solution to the crisis in many cities and local communities. A more viable alternative may be the adoption of smart parking systems, which optimize the use of existing spaces, improve security, and provide new ways for local authorities to generate revenue [3].

According to the study conducted by Smart City Romania [3], in 2021, the number of smart city projects had already increased by 42% to 860 separate initiatives in this area compared to other years. In 2024, the Romanian Association for Smart City (ARSC) organized an award ceremony for the most revolutionary initiatives in the field of urban development, sustainability, and digital transformation, counting more than 1300 projects and initiatives in this field [4], demonstrating that civic pressure is growing and bringing positive results in this direction.

The ARSC also tells us that at the ceremony, the 60 projects in the total number had an estimated value of around 2 million euros, representing 1% of Romania’s GDP.

According to a study conducted in 2023 [5], the Municipal Parking Company has invested semi-significantly in modernizing the parking service in Bucharest. This included the purchase of 18 electric vehicles and a wide range of technological equipment. The implementation of camera-equipped cars was a major step toward monitoring parking. The tenders for equipment such as parking meters and video monitoring solutions had a total value of approximately RON 12,741,100. It is also planned to optimize and expand digital payment options, including mobile apps and SMS, to offer users more convenient and fast payment methods.

Another study [6] informs us that Bucharest is a pioneer in the implementation of traffic management systems, advancing over time to the rest of the smaller cities. It also tells us that between 2007 and 2009, the Bucharest City Hall implemented an intelligent traffic management system, which at the time was the most modern in Europe. On 18 July 2017, the new Traffic Management Center was inaugurated,

which includes an integrated control for the public transport management system, an urban traffic control system, and a video surveillance system. It is, in fact, a first step toward making Bucharest a smart city.

7.3 Analysis of Four Smart Parking Solutions in Romania

7.3.1 *Amparcat*

The “Amparcat” application represents a significant change in parking management in Bucharest, marking an evolution toward smart city technologies. Launched on 1 March, the app implements new tariffs according to HCGMB no. 705 of 22.12.2022, offering a modern and efficient approach for parking users in the capital [7].

The main innovation brought about by Amparcat is the simplification of the charging system, with a single zone and a single hourly rate, eliminating the complexity of the previous three zones with different rates. The daily rate has also been unified and reduced, improving the user experience and making the payment process easier.

The app allows for saving cards, entering billing details, and listing all invoices to make it easier to keep track of money.

Another important aspect is the diversity of payment methods offered by the app, including cash, bank card, Apple Pay, Google Pay, Revolut Pay, Self-Pay, and SMS, increasing convenience and accessibility for users. If we don’t want to stay and pay by card every time, the app gives us the possibility to buy credits, through which we can make future payments (see Fig. 7.1).

This technological approach is an important step toward modernizing the parking infrastructure in Bucharest, aligning with global trends in smart cities, and improving the parking experience for residents and visitors.

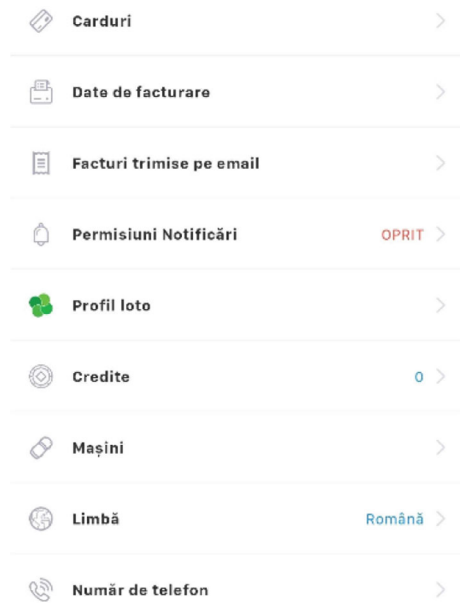
At the same time, the platform offers advanced parking management options, including the ability to purchase long-term parking passes and alert functions for the expiration of other vehicle-related documents, such as the vehicle registration card or the parking ticket.

With all these technological innovations, Amparcat is a prime example of how technology can improve and simplify the urban parking experience, helping to make Bucharest a smarter and more accessible city.

7.3.2 *TPARK*

TPARK is an innovative app that offers modern payment services for drivers in Romania and Hungary. With this app, users can pay for parking, purchase vignettes,

Fig. 7.1 Screenshot of Amparcat mobile application menu



RCA insurance (third-party driver's liability insurance), and pay parking fines, all directly from their mobile phone [8].

One of the distinctive features of the app is the ability to pay for parking in a fast and efficient way, eliminating the need to use parking meters or coins. Users can extend their parking period from the app without having to get in their car.

TPARK also offers the possibility to purchase vignettes for Romania and Hungary, saving time and avoiding queues at physical outlets or at the border. Another benefit of the app is the ability to pay for MTPL insurance directly from your mobile phone, without having to visit an insurance agency or broker (Fig. 7.2).

Users can also pay parking fines at any time of the day, eliminating the risk of higher penalties for not paying on time.

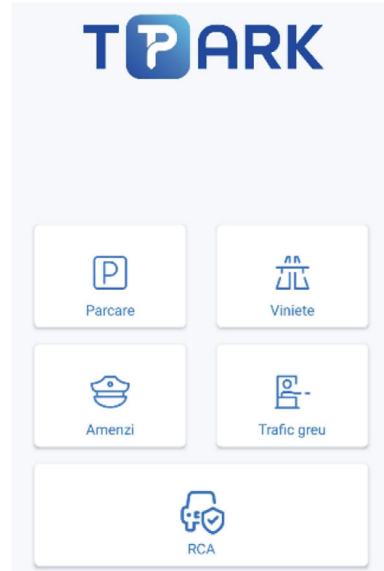
The TPARK app is easy to use and offers a wide range of functionality, such as keeping track of payments made, saving time, and eliminating the need to carry coins on hand.

With its modern and efficient services, TPARK is a complete solution for the car payment needs of drivers in Romania and Hungary, helping to simplify and improve their driving experience.

7.3.3 *Parking Bucuresti*

The Bucharest Parking app, developed by the Bucharest Parking Municipal Company S.A., is the ideal partner for parking management in the crowded city. Downloading

Fig. 7.2 Screenshot of TPARK mobile application menu



and installing is quick and easy, and the log-in process is simplified, allowing us to log in quickly through our Google or Apple account or email address [9]. The app's menu can be seen in Fig. 7.3.

The intuitive interface gives us quick access to essential functionality. With the “Quick Pay” option, we can pay for parking efficiently by selecting time periods or scheduling payments for the parking interval.

The “Search” page allows us to quickly find available parking via an interactive map, and the “History” page gives us a detailed view of previous parking and the ability to add time to active parking.

The “More” menu gives us the opportunity to customize our account, manage vehicles, add funds to our wallet, or adjust account settings.

According to a study conducted in 2022 [10], the Bucharest Municipal Parking Company aimed to transform the management of public parkings, with a target of 1 million lei per day revenue. By investing more than 26 million lei in 2022, including the barrier closure of 10 public car parks and the implementation of modern technological solutions, the aim was to improve services and generate significant revenue in a short time.

7.3.4 UPPark

The UPPARK mobile app stands out for its intuitive interface and advanced functionalities, allowing users to manage and locate available parking spaces in real time.

Fig. 7.3 Screenshot of Parking Bucuresti mobile application menu



Its nationwide deployment is a clear proof of its adaptability and relevance in the context of changing urban mobility [11].

The systems and equipment proposed by fastPark are clearly the result of a rigorous selection process, coming from world-renowned manufacturers such as Flowbird, HUB Parking Technology, Magnetic Control, and others [11]. These technologies cover the full spectrum of parking management needs, including payment process automation, video surveillance, and driver guidance systems.

FastPark's technical support services are widely distributed throughout the country, ensuring prompt response time and efficient resolution of technical issues. The company's team of specialists with expertise in various fields complements this solid support infrastructure, bringing added competence and innovation.

Close partnerships with equipment manufacturers and positive customer references attest to the quality and reliability of fastPark's solutions. This industry recognition demonstrates the company's commitment to excellence and supports its position as a leader in parking technology in Romania.

Parking systems on the street are an essential solution for efficient urban management. The equipment and services offered cover the full spectrum of needs, from automating the payment process to reducing the time it takes to find parking

spaces. These solutions not only generate additional revenue and stimulate economic development, but also improve the quality of life of drivers [12].

Off-street parking systems provide an increased level of safety and convenience for drivers. Working with leading suppliers such as HUB Parking Technology and Magnetic AutoControl Access, fastPark offers high-quality equipment and solutions tailored to specific customer needs. From ANPR cameras to automated payment machines and guidance systems, the company is committed to providing a seamless and efficient off-street parking management experience [13].

7.4 A Comparison of the Presented Parking Solutions

In recent decades, accelerated urbanization and the increase in the number of vehicles have created new challenges in the management of urban space and parking. Parking crisis and traffic congestion are becoming increasingly acute in large cities, requiring innovative and efficient solutions to manage these problems. In this context, smart parking applications have become increasingly popular and have started to play a crucial role in optimizing the use of available parking space and improving the parking experience for drivers.

This article proposes a comparative analysis of four available smart parking applications in Romania: Amparcat, TPARK, Parking Bucuresti, and UpPark. The aim of this analysis is to examine the features and functionalities offered by each application to identify its advantages and disadvantages in terms of urban parking management. By evaluating these applications in a comparative framework, the aim is to provide a more in-depth understanding of the diversity of solutions available and how they can contribute to solving urban parking and mobility problems.

In the following, we will analyze each application individually, highlighting the main features and functionalities offered, to provide a comprehensive overview of the diversity of options available to users. The analysis will be complemented by a comparative table highlighting the differences and similarities between the four applications, making it easier to understand and evaluate them in relation to the needs and preferences of the individual user (Table 7.1).

Comparative analysis of the four parking apps reveals several differences and similarities between them, highlighting the diversity of functionality and options available to users. One of the major features taken into account is the geographical coverage, where Amparcat and Parking București are available in less than 10 cities, while TPARK and UpPark are present in more than 10 cities, with TPARK extending its activity to Hungary. This difference in geographical spread may influence the choice of users, depending on their location and mobility needs.

Another important feature is the possibility of purchasing long-term parking passes, offered only by Amparcat and Parking București. This functionality can be essential for users who wish to have access to long-term parking without having to pay repeatedly for each use. Amparcat also stands out by providing alerts for expiration of car documents, including the RCA insurance and the road vignette, thus providing

Table 7.1 Parking applications comparison

	Amparcat	TPARK	Parking București	Up park
Number of cities covered	<10	>10	<10	>10
Countries of operations	România	România, Ungaria	România	România
Map	✓	X	✓	✓
Subscription > 1 day	✓	X	✓	X
Vehicle document alerts	✓	✓	X	X
Payment of fines	X	✓	X	X
On-Street parking	✓	✓	✓	✓
Off-Street parking	X	X	X	✓
Roadside assistance	✓	X	X	X
Rewards	✓	X	X	X
Payment methods	Card, SMS, Google Pay, Apple Pay, Revolut	Card, SMS	Card, SMS	Card, SMS

additional functionality for users that contributes to the efficient management of their vehicle.

In terms of paying parking fines, only TPARK offers this facility, which can be an important factor for users who wish to avoid penalties and penalties for late payment. Furthermore, all four applications offer parking management on the street, but only UpPark also offers parking management on the street, thus extending the range of options available to users in terms of available parking spaces.

Another important aspect is the roadside assistance and rewards offered by Amparcat, which add more value to the user experience. In addition, the variety of payment methods supported by Amparcat, including card, SMS, Google Pay, Apple Pay, and Revolut, gives users greater flexibility in how they pay for parking services.

Thus, the choice between these applications can be influenced by the user's individual preferences and specific parking and urban mobility needs. However, it is important to consider the diversity of options and functionalities offered by each application to find the solution that best fits your personal needs and preferences.

7.5 Conclusions

The comparative analysis of smart parking applications in Romania highlights the diverse functionalities and options available to users, showcasing the potential of technology in addressing urban parking challenges. However, a common thread

throughout the analysis is the absence of integration with artificial intelligence (AI) technologies in the existing applications. The conclusion drawn from this study emphasizes the pressing need for incorporating AI into parking systems to enhance efficiency, optimize resource utilization, and improve the overall parking experience for citizens.

Integrating AI into smart parking applications can revolutionize urban mobility by enabling predictive analytics for parking space availability, optimizing parking allocation based on demand patterns, and facilitating dynamic pricing strategies to manage congestion. Furthermore, AI-powered solutions can streamline parking enforcement processes through automated detection of violations, leading to more effective enforcement and revenue generation for municipal authorities.

Moreover, AI-driven insights can inform urban planning decisions by providing data-driven recommendations for infrastructure development and traffic management strategies. By leveraging AI algorithms to analyze traffic flow, parking utilization, and user behavior, city planners can design more sustainable and inclusive urban environments that prioritize accessibility, safety, and environmental sustainability.

The integration of artificial intelligence into parking management, as described in the second article, is a significant development in urban planning and technology. By using advanced number plate recognition and object detection systems, such as YOLO (You Only Look Once) technology, accurate and efficient operations in parking management can be achieved. This approach not only simplifies the process of identifying available parking spaces but also helps prevent accidents by continuously detecting and monitoring vehicle movement in parking lots [14].

The benefits of integrating artificial intelligence into parking management are multiple and tangible. By automating the processes of number plate recognition and vehicle movement monitoring, the time taken to find a parking space can be significantly reduced, helping to smooth traffic flow and reduce stress for drivers. YOLO technology also enables quick and accurate identification of available parking spaces, making the parking process easier and reducing the time spent searching for a suitable spot.

In addition to the benefits for users, the integration of artificial intelligence into parking management can also bring significant advantages for local authorities and parking operators. By implementing intelligent parking management systems, real-time data can be collected and analyzed, allowing informed decisions to be made regarding the development and optimization of urban infrastructure. Also, by continuously monitoring parking activity and detecting incidents or accidents, effective risk management and increased safety in the urban environment can be ensured.

In the future, we aim to build on this analysis of the current state of parking applications in Romania and focus on developing and integrating artificial intelligence into new parking systems to bring significant improvements in efficiency and user experience. A first goal would be to implement advanced data analytics and machine learning algorithms to perform predictive analytics on parking space availability. By collecting and analyzing parking usage data in real time, these applications could predict demand patterns and provide personalized suggestions to users,

leading to more efficient use of parking spaces and reducing time spent searching for an available space.

In conclusion, the integration of artificial intelligence into smart parking systems represents a crucial step toward realizing the full potential of smart city initiatives in Romania. By harnessing the power of AI technologies, parking applications can evolve from transactional tools to intelligent platforms that actively contribute to the creation of more livable, efficient, and resilient urban spaces. Therefore, policy-makers, urban planners, and technology developers must collaborate to drive innovation and ensure the seamless integration of AI into parking management systems, ultimately enhancing the quality of life for residents and visitors alike.

References

1. Ma, R., Lam, P.T.I., Leung, C.K.: Potential pitfalls of smart city development: a study on parking mobile applications (apps) in Hong Kong. *Telemat. Inform.* **35**(6), pp. 1580–1592, ISSN 0736–5853 (2018). <https://doi.org/10.1016/j.tele.2018.04.004>
2. Aditya, A., Anwarul, S., Tanwar, R., Koneru, S.K.V.: An IoT assisted intelligent parking system (IPS) for smart cities. *Procedia Comput. Sci.* **218**, pp. 1045–1054, ISSN 1877–0509 (2023) <https://doi.org/10.1016/j.procs.2023.01.084>
3. Smart City Romania. Smart Parking, soluția optimă pentru criza locurilor de parcare. <https://arsc.ro/smart-parking-solutia-optima-pentru-criza-locurilor-de-parcare/>. Accessed 06 Apr 2024 (2021)
4. Logistic Post. Proiecte de dezvoltare urbană însumând 2 miliarde de euro (1% din PIB), la Smart City Industry Awards. <https://logisticpost.ro/articole/proiecte-de-dezvoltare-urbana-insumand-2-miliarde-de-euro-1-din-pib-ul-romaniei-la-smart-city-industry-awards.html>. Accessed Apr 06 2024 (2024)
5. Asociația Română Smart City. Parking Experts începe colaborarea cu Primăria Capitalei. Soluții verificare, sancționare parcări și colectare date trafic digitale. <https://arsc.ro/parking-experts-incepe-colaborarea-cu-pmb/>. Accessed Apr 06 2024 (2023)
6. Săvulescu, C., Antonovici, C.G.: Alba Iulia – Landmark for smart cities in Romania. *Smart Cities International Conference (SCIC) Proceedings*. 5, Mar. 2023, 114–123 (2023)
7. Ion Dobreanu Aplicația Amparcat – soluția completă pentru parcare în București, https://www.economica.net/aplicatia-amparcat-solutia-completa-pentru-parcare-in-bucuresti_651593.html. Accessed Apr 06 2024 (2023)
8. TPARK <https://tpark.io/#despre>. Accessed Apr 06 2024 (2024)
9. Parking Bucuresti – Documentation. https://s3.eu-central-1.amazonaws.com/cdn.parkingbucuresti.ro/docs/Manual_utilizare_aplicatie_Parking_Bucuresti.pdf. Accessed Apr 06 2024
10. Catiușa Ivanov. Cum vrea compania care administrează parcările publice din București să ajungă la încasări de 1 milion de lei pe zi / Ce îi așteaptă pe bucureșteni. https://www.hotnews.ro/stiri-administratie_locala-25651991-cum-vrea-compania-municipala-parking-care-administreaza-parcarile-publice-din-bucuresti-ajunga-incasari-1-milion-lei-asteapta-bucuresteni.htm. Accessed Apr 06 2024 (2022)
11. Fastpark—About Us. <https://www.fastpark.ro/despre-noi/>. Accessed Apr 06 2024 (2024)
12. Fastpark—On-Street. <https://www.fastpark.ro/parcari-on-street/>. Accessed Apr 06 2024 (2024)
13. Fastpark—Off-Street. <https://www.fastpark.ro/parcari-off-street/>. Accessed Apr 06 2024 (2024)
14. Jung, I.H., Lee, J.M., Hwang, K.: Smart parking management system using AI. <https://doi.org/10.14704/WEB/V1911/WEB19307>. Accessed Apr 12 2024 (2022)

Chapter 8

Measurement and Verification Solution for Energy Performance Contracting in the Building Sector



George Suciu, Lucian Necula, Cosmina Stalidi, and Mădălin Silion

Abstract Facilitated by the global digitalization efforts and the proliferation of innovative technologies in the past decade, the cost of energy management solutions has decreased and they became widely available on the market. As Energy Performance Contracts (EPC) allows customers implementing energy efficiency measures to repay the project costs through the savings they achieve, a key aspect is to provide an instrument for a transparent and secure measurement and verification (M&V) of the benefit generated, both in terms of energy savings and cost. In this regard, FinSESCo provides a toolchain of innovative applications to generate and control the project data, manage investors, collect IoT data from the implementation premise and assess the savings achieved through the project. In this regard, the paper presents a M&V tool implemented within the FinSESCo toolchain for measuring the benefits achieved as a result of EPC implementation.

8.1 Introduction

As a result of permitting repayments in the form of energy savings, Energy Performance Contracts (EPC) have emerged as a crucial tool for promoting energy efficiency in the context of growing environmental consciousness and digital progress. An essential aspect of this method is the creation of a strong system for the clear

G. Suciu (✉) · L. Necula · C. Stalidi · M. Silion
Research & Development Department, Beia Consult International, Bucharest, Romania
e-mail: george@beia.ro

L. Necula
e-mail: lucian.necula@beia.ro

C. Stalidi
e-mail: cosmina.stalidi@beia.ro

M. Silion
e-mail: madalin.silion@beia.ro

and precise measurement and verification (M&V) of energy savings and cost advantages. FinSESCO project leads this drive by delivering a set of cutting-edge digital solutions designed for efficient data administration, investor involvement, and live monitoring using Internet of Things (IoT) technology. This paper presents the effectiveness and revolutionary influence of the digital solutions used by the FinSESCO project in improving energy performance contracting, with the ultimate objective of attaining sustainable energy management objectives.

8.2 Related Work

The FinSESCO project aims to digitalize the creation and implementation of Energy Performance and Energy Savings Performance Contracting (EPC/ESPC). This coincides with recent improvements in digital technologies that aim to improve energy efficiency and security. The incorporation of digital tools into energy systems, specifically through the establishment of data standardization (such as XML and dXML files), gamification, smart contracting, and meter-based repayment processes, presents a novel opportunity for enhancing energy management and conservation endeavors.

The second section of the article provides an analysis of current solutions and approaches to the implementation of energy efficiency measures through energy performance contracting.

Thus, recent research and projects offer a thorough analysis of how digital technology is being utilized to enhance energy efficiency and successfully handle energy performance contracts. A study examined the effective implementation of energy efficiency projects through Public–Private Partnerships (PPPs) by ensuring a fair distribution of benefits and risks between the public and private sectors. The study created a mathematical framework to determine the most advantageous way to distribute benefits in long-term agreements. It emphasized the significance of contractual arrangements that promote equitable partnerships and effective risk management in investments related to energy efficiency [1].

A further investigation sought to ascertain prospective areas of inquiry by examining scholarly articles on EPC. The statement highlighted the increasing interest in applying EPC projects, methods, business models, decision-making processes, the involvement of ESCOs (Energy Service Companies), and risk management in EPC projects. The analysis yielded useful insights into the prevailing research topics and potential future directions, emphasizing the crucial significance of case studies and descriptive statistical analyses in developing the field [2].

The digital transformation in the energy sector aims to enhance end-to-end workflows in order to redefine operational efficiency and support ambitious commercial goals. This method focuses on digital projects that make a major contribution to future goals, highlighting the necessity of small, practical activities to develop momentum and practical benefit from digital technologies. The text showcases instances in the energy industry where digital tools such as drones and predictive analytics have

been incorporated into operational workflows to improve efficiency and get valuable operational insights [3].

A research investigation studying the effects of digitalization on energy efficiency in Zhejiang Province, China, reveals that digitalization has a favorable impact on energy efficiency. The study investigates the correlation between digitalization and energy efficiency using empirical techniques, highlighting the significance of digital technology in overseeing energy systems, optimizing energy frameworks, and enabling enhancements in energy efficiency at both small-scale and large-scale levels. The results indicate that the implementation of digital technology, including Internet of Things (IoT) and cloud computing, can greatly improve the energy efficiency of systems and processes [4].

In order to substantiate the decision to implement photovoltaic solutions through EPCs, a vital role is played by key performance indicators (KPIs) for monitoring the operations and maintenance (O&M) of installations. A recent research investigation presents a set of KPIs classified for energy performance assessment and O&M services assessment, which were further weighted by importance and complexity using the stepwise weight assessment ratio analysis (SWARA) method [5].

A recent research study proposes an auction system for securing long-term contracts for solar power generation that aims to provide stable earnings for solar power producers. The researchers suggest a method based on mechanism design theory, where solar power providers bid both their operating costs and the amount of energy they can supply. The buyer then selects the winning bids and signs contracts for a set duration. This model takes into account the efficiency of the solar panels, ensuring that the contracts reflect the actual energy production. The proposed system is designed to be fair and encourage honest bidding, with the goal of maximizing the buyer's benefits over the contract term. A specific algorithm was created to run this auction. When tested with data from South Korea's solar power auctions, the new model performed comparably to existing uniform price auctions in terms of electricity costs and bid prices and was more cost-effective than the Vickrey auction. Although the new model slightly increased overall costs (by about 1%), it led to more electricity being bought and increased the buyer's profits [6].

Furthermore, the International Energy Agency (IEA) has emphasized the significance of digitization in advancing energy efficiency and mitigating emissions throughout the energy industry. The European Union and countries such as the United Kingdom are making significant advancements in the digitalization of their energy systems. This indicates a noticeable shift toward utilizing digital technology to enhance the efficiency, resilience, and sustainability of energy systems. The implementation of intelligent meters and interconnected devices, aided by digital technology, is considered a vital measure in the process of incorporating fluctuating renewable energy sources, improving the alignment of energy supply and demand, and facilitating the shift toward low-carbon energy alternatives [7].

The above findings indicate that the process of digitalizing EPC and ESPC projects, using technologies such as IoT, cloud computing, and machine learning, has the ability to bring about significant changes and improvements. Digital tools can enhance energy efficiency, optimize energy systems, and promote sustainable urban

development by facilitating precise data gathering, analysis, and real-time monitoring. The focus on Public–Private Partnerships (PPPs), current research patterns, and effective digital transformation techniques offers a clear plan for utilizing digitalization in energy performance contracting, in line with the goals of initiatives such as FinSESCO.

8.3 FINSESCO Concept

The FinSESCO project [8], aims to facilitate the digitalization of set-up and implementation of Energy Performance and Energy Savings Performance Contracting EPCo/ESPCo, through a toolchain of instruments. These interoperable instruments address the generation of standardized XML files from buildings’ passes and energy audits and dXML files describing the EPC-based project, the investment process through gamification, smart contracting, and meter-based repayment process from the savings achieved. FinSESCO is designed to provide interoperable and integrated tools that aid in building assessment, substantiation of energy performance measures, project creation, crowdfunding and investment, project implementation, M&V of savings, and digital controlling (Fig. 8.1).

To showcase the applicability and the benefits of these tools, pilot deployments have been strategically located across various European countries, including Austria, Germany, Romania, and Spain. The main focus of these pilots is to integrate innovative solutions in EPC for building renovation, renewable energy, and digital technologies, emphasizing energy efficiency and sustainable practices.

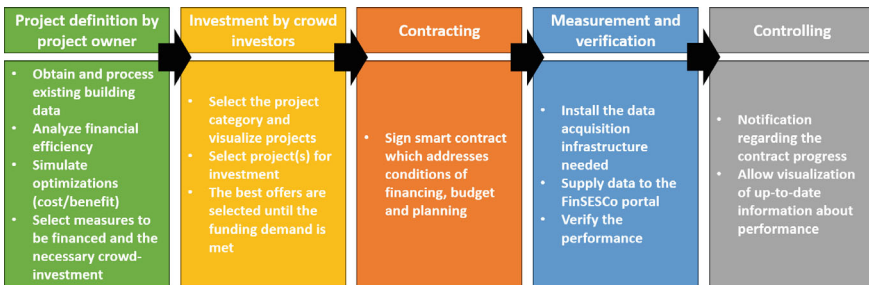


Fig. 8.1 Architecture of the measurement & verification tool

8.4 Measurement and Verification of Savings

The M&V (Measurement and Verification) of the results achieved through the implementation of an EPC (Energy Performance Contract) is addressed in the FinSESCO project through an instrument which is being implemented and validated as part of the Romanian pilot. It provides functionalities such as IoT (Internet of Things) data visualization, energy costs and savings monitoring, assessment of the investor KPIs (Key Performance Indicators) associated with their share of the investment, and project performance evaluation.

The architecture of the measurement and verification instrument integrated into the FINSESCO toolchain is presented in Fig. 8.2.

The data inputs required for measurement and verification are:

- dXML file providing investment data as the objective of the project, the investment value, the contract duration, the expected yield, and investor data (name, investment, and share in the total investment pool required for the action);
- energy inverters, providing insights into the total load, the consumption from the grid, and the consumption from photovoltaics;
- smart plugs, providing insights on the consumption of various cost centers within the building (e.g., data center or electric boiler);
- weather stations, which provide environmental data such as the ambient temperature, humidity, and wind speed;
- photovoltaic monitoring station, which is attached to the solar panels and provides valuable data as the temperature of the photovoltaic cell and the solar irradiance;
- external data sources (e.g., weather forecasts).

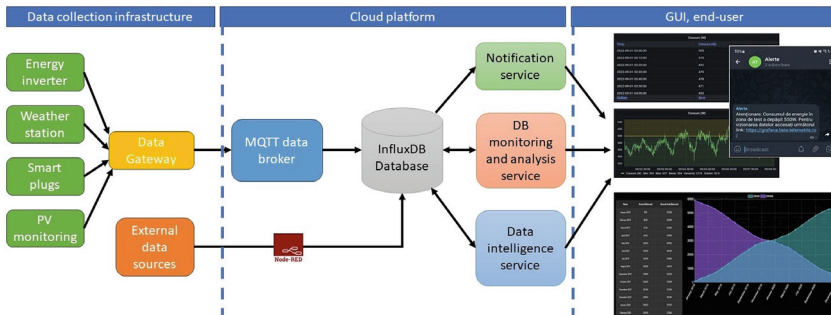


Fig. 8.2 Architecture of the measurement and verification instrument

8.4.1 Sensor Data Collection and Visualization

After being gathered by RTUs equipped with sensors, the data is sent to the data gateway (GSM, LoraWan, Zigbee, etc.) and then to the MQTT data broker in the Cloud. For the IoT integration [9], NodeRED is being used to set up custom dataflows. All the collected data is being stored in the InfluxDB database in BEIA's Cloud [10]. The notification service allows the customization of alerts and notifications in terms of communication channels (SMS, Telegraf, Discord, Skype, email, etc.), frequency of alerts, and thresholds (min or max value). The DB (database) monitoring and analysis service allows data structuring and processing, also handling user requests for sensor data visualization in different formats (graphs, tables, diagrams, custom reports, etc.).

For the monitoring and management of PV panels, a Libelium SmartDataSystem telemetry station has been used. It collects data regarding the solar irradiance and the temperature of the photovoltaic cell, providing an overview on the environmental data correlated with the energy produced in a specific timeframe. A dashboard model presenting data from the photovoltaic monitoring station is presented in Fig. 8.3.

To gather environmental data, an OTT HydroMet A764 telemetry station has been implemented on the Romanian pilot site. It is equipped with sensors to measure the ambient temperature, wind speed and direction, air relative humidity, solar irradiance, and rain volume. The solution integrates a datalogger and a GSM module which facilitates the transmission of pre-processed data to the MQTT data broker in the Cloud platform where the FinSESCo M&V instrument has been deployed. A dashboard model presenting data from the weather station is presented in Fig. 8.4.

The SolarEdge SE6K inverter provides data regarding the total load, the energy consumption from PV, and the energy consumption from the public grid. The data

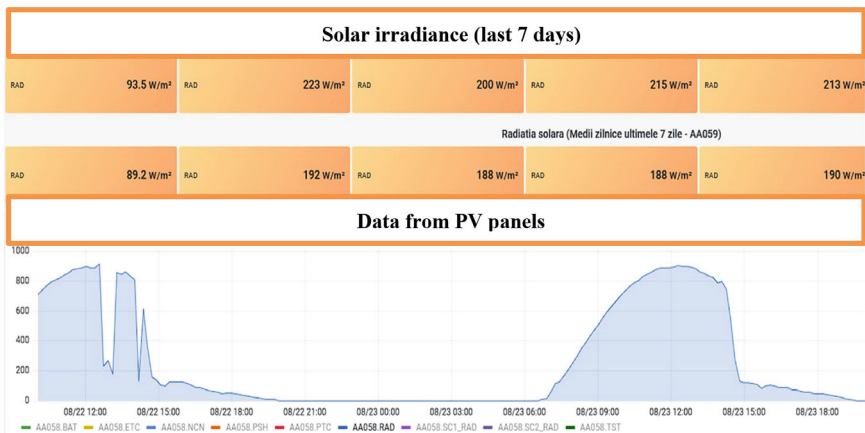


Fig. 8.3 Dashboard presenting data from the photovoltaic monitoring station

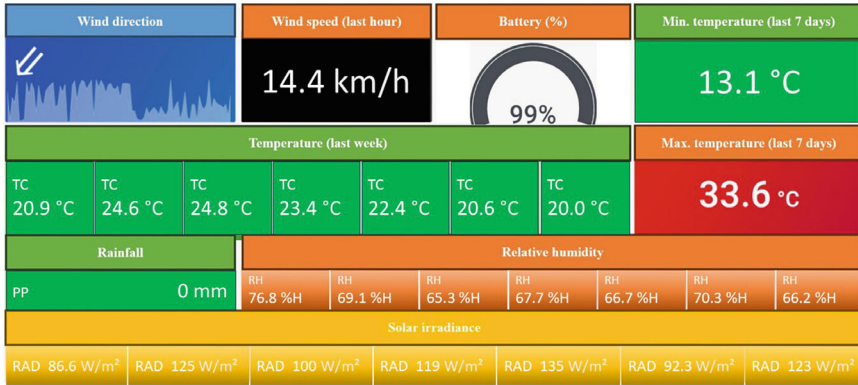


Fig. 8.4 Dashboard presenting data from the weather station

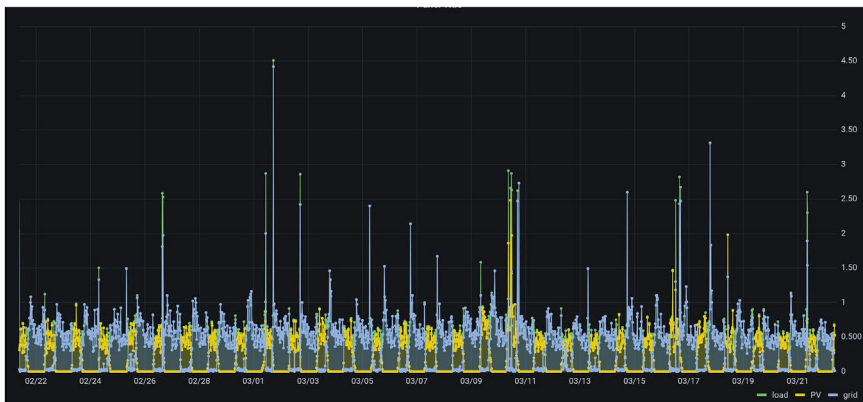


Fig. 8.5 Dashboard presenting data from the energy inverter

is being collected and sent to the FinSESCo data broker in the Cloud at 15-min intervals. A dashboard model presenting data from the energy inverter is presented in Fig. 8.5.

8.4.2 DXML Data

A dXML file is being used as an input to provide information regarding the project and the investment, as presented in Fig. 8.6. The dXML file is uploaded into the FINSESCO M&V portal where it is being parsed and information is extracted to generate or update the project data.

```

<?xml version="1.0" encoding="UTF-8" ?>
<dxml version="1.0">
<Project>
  <ID>00234</ID>
  <Type>Photovoltaics</Type>
  <EstimatedYield unit="KWh/Year">4000</EstimatedYield>
  <Investment unit="EUR">4000</Investment>
  <ContractDuration unit="months">60</ContractDuration>
  <Investors>
    <Investor1>
      <Name>S. C. Beia Consult International</Name>
      <Type>Business</Type>
      <ProjectRole>Owner</ProjectRole>
      <Investment unit="EUR">1800</Investment>
      <Share unit="%">45</Share>
    </Investor1>
  </Investors>
  <RenewableEnergies>
    <Photovoltaic>
      <System>
        <Size>
          <PeakPower unit="KW">6</PeakPower>
          <Area>33,68</Area>
        </Size>
        <Components>
          <Modules>
            <Manufacturer>JinkoSolar Holding</Manufacturer>
            <Model>JRM-335-60H Cheetah HC</Model>
            <Technology>Monocrystalline</Technology>
            <Number>20</Number>
            <Orientation>NE</Orientation>
            <Tilt>15,07</Tilt>
          </Modules>
          <Inverters>
            <Manufacturer>SolarEdge</Manufacturer>
            <Model>SE6K-RW0TEBNN4</Model>
            <Number>1</Number>
            <Power unit="VA">6000</Power>
            <Efficiency>0.98</Efficiency>
          </Inverters>
        </Components>
      </System>
    </Photovoltaic>
  </RenewableEnergies>
</Project>
  </dxml>
</?xml>

```

Fig. 8.6 Information included in the dXML file

8.4.3 M&V Data Analysis

The data intelligence service facilitates the analysis of project-specific data, such as savings, energy production and consumption forecasts, ROI (Return of Investment), NPV (Net Present Value), and IRR (Internal Rate of Return). The GUI (Graphical User Interface) provides data visualization and interaction with the user through a web application.

The energy data dashboard (Fig. 8.7) displays data related to the energy consumption, the associated cost, and the savings achieved. This dashboard provides a general image regarding the impact of the EPC measures on energy usage and savings.

The investor dashboard (Fig. 8.8) provides information regarding the amount invested by each party and the share they have in the overall investment pool, the amount returned from the investment, and specific KPIs describing the up-to-date and the expected performance.

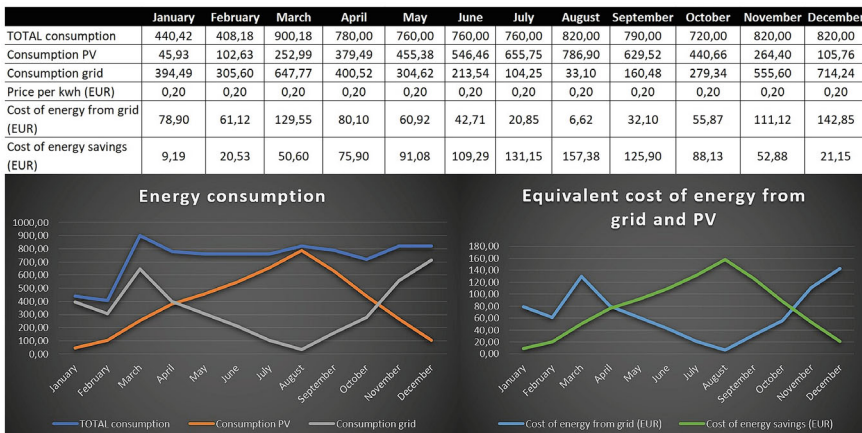


Fig. 8.7 Energy data dashboard

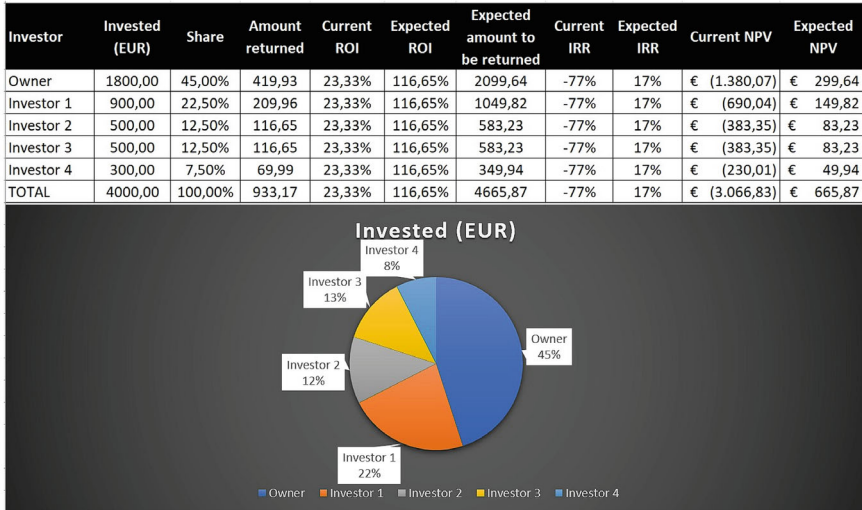


Fig. 8.8 Investor dashboard

While the investor dashboard provides data regarding investor-specific KPIs, which is useful for them to determine any risk associated with their individual investment, the project performance dashboard (Fig. 8.9) facilitates a more complex analysis allowing for timely risk management.

The project performance data can serve as a valuable reference for other projects in regions with a similar climate or aiming to implement a similar technical solution. This information can be used in the project design phase, assisting investors in substantiating their decision to invest in a particular project over another.

8.5 Conclusions

In conclusion, this paper presents the innovative implementation of the FINSESCO project which provides a toolchain facilitating the adoption of energy efficiency measures across Europe through Energy Performance Contracts, contributing to a wave decarbonization. FinSESCO project offers a progressive approach to energy management and efficiency, leveraging digital advances to encourage the use of environmentally friendly and cost-effective energy solutions. It facilitates the advancement of energy performance contracts, providing deep insights and methods for enhancing energy performance measurements. FinSESCO’s extensive digital toolkit highlights the crucial significance of digitization in revolutionizing the energy industry, signaling a new era of improved energy management tactics. As future work, FinSESCO aims to enhance the interoperability of these tools and finalize the integration and testing within the national pilots.



Fig. 8.9 Project performance dashboard

Acknowledgements This work has been supported by UEFISCDI through the FINSESCO project, funded in part by the European Union’s ERANET MICall on digital transformation for green energy transition, contract number 289/02.05.2022.

References

1. Martiniello, L., Morea, D., Paolone, F., Tiscini, R.: energy performance contracting and public-private partnership: how to share risks and balance benefits. In: *Energ. S.* **13**(14), 3625 (2020)
2. Zhang, W., Yuan, H.: Promoting energy performance contracting for achieving urban sustainability: what is the research trend? In: *Energ. S.* **12** (8), 1443 (2019)
3. McKinsey&Company page: Digital transformation in energy: Achieving escape velocity. <https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/digital-transformation-in-energy-achieving-escape-velocity>. Accessed 15 Mar 2024
4. Niu, Y., Lin, X., Luo, H., Zhang, J., Lian, Y.: Effects of digitalization on energy efficiency: evidence from Zhejiang province in China. In: *Front. Energy Res.* **10**, 847339 (2022)
5. Rediske, G., Michels, L., Siluk, J.C.M., Rigo, P.D., Rosa, C.B., Lima, A.C.: A proposed set of indicators for evaluating the performance of the operation and maintenance of photovoltaic plants. In: *Appl. Energy.* **354**, 122158 (2024)
6. Jung, J., Song, C.O., Lee, D.J., Yoon, K.: Optimal energy procurement with long-term photovoltaic energy contracts considering generation uncertainty: a two-dimensional auction approach. In: *Appl. Energy.* **356** 122383 (2024)
7. International energy agency page. <https://www.iea.org/energy-system/decarbonisation-enablers/digitalisation>. Accessed 02 Apr 2024
8. FinSESCO project page. <http://energycontracting.info/>. Accessed 04 Apr 2024

9. Suciu, G., Hussain, I., Badicu, A., Necula, L., Usurelu, T.: IoT services applied at the smart cities level. In: Trends and Innovations in Information Systems and Technologies, vol. 2 8, pp. 457–463, Springer International Publishing (2020)
10. Suciu, G., Necula, L., Iosu, R., Usurelu, T., Ceaparu, M.: IoT and cloud-based energy monitoring and simulation platform. In: 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE) pp. 1–4, IEEE (2019)

Part II
Digital Business and e-Transformation;
Cloud, Distributed and Parallel
Computing; IT Deployment in Cultural
Institutions

Chapter 9

Analyzing the Antecedents of Deceived Buying and Disinformation Risk



Luigia-Gabriela Sterie  and Daniel Mican 

Abstract The pervasiveness of social media advertising raises concerns about its potential to deceive consumers and exploit their susceptibility to misinformation. This study delves into the factors influencing deceived buying and leading to the propensity of believing and acting on disinformation in the context of advertised products on social media platforms. Employing Partial Least Squares—Structural Equation Modeling, the research analyzes data collected from 334 consumers. The findings reveal that perceived benefits, perceived risks of social media advertised products, and openness to change play an important role in deceived buying, while informativeness has no significant influence. Deceived buying, in turn, exhibits a strong positive association with the propensity of believing and acting on disinformation. The results hold significant implications for various stakeholders such as businesses, organizations, social media users, and policymakers. To combat deceptive social media ads, this study proposes a two-pronged approach: empowering platforms with algorithms specialized in deceived or false ad detection and creating educational programs for users to become more discerning consumers.

9.1 Introduction

The ubiquitous presence of social media demonstrably shapes the contemporary landscape of consumer buying decisions. Hence, the dissemination of deceptive content on social media fosters a concerning trend: deceived buying and an increased susceptibility to disinformation. Advertisement is a form of social media marketing that can influence purchasing behavior [4, 8], especially because it can be targeted to specific demographics or interests and can be very persuasive. Social media has

L.-G. Sterie · D. Mican (✉)
Faculty of Economics and Business Administration, Babeş-Bolyai University, Cluj-Napoca,
Romania
e-mail: daniel.mican@econ.ubbcluj.ro

L.-G. Sterie
e-mail: luigia.sterie@stud.ubbcluj.ro

made it easier for people to share information, including false or deceiving ones [9, 18, 24]. This can be a problem when it comes to purchasing behavior, as people may be misled into buying products or services that are not what they are advertised to be. Hence, deceived buying can occur when people are influenced by false information or do not have enough information to make an informed decision [10, 14, 21]. This can lead to people buying products they do not need or that are not what they expected, or that can even be harmful to health. False or misleading advertising is advertising that contains false or deceptive information [21]. This can include claims about the product's performance, ingredients, or benefits, and even testimonials from fake or paid actors that can have a devastating impact on a brand. When consumers are misled by false advertising, they are likely to lose trust in the brand [14] which can lead to lost sales, damaged reputation, and even legal action.

There are many examples of brands that have been hurt by false or misleading advertising. In 2015, the Federal Trade Commission (FTC) fined Herbalife \$200 million for making false claims about its products' weight loss benefits [1]. In 2017, Volkswagen was caught cheating on emissions tests for its diesel vehicles, costing the company billions of dollars in fines and lost sales [23]. Moreover, in the context of misinformation and misled purchase behavior, Truth in Advertising, a consumer protection organization, investigated health claims made by nutritional supplement companies within the Direct Selling Organization. A shocking 97% (60 out of 62 companies) were found making illegal claims, promising their products could treat or prevent diseases [1]. This raises serious concerns about misleading consumers for profit. The consequences of rampant misinformation extend beyond the realm of a single purchase. Consumers who are repeatedly exposed to deceptive marketing tactics on social media can become more susceptible to believing and acting upon disinformation in general.

Thus, this phenomenon necessitates a multifaceted approach to mitigating the risks associated with deceived buying and disinformation vulnerability. This study aims to analyze the impact of a series of potential drivers of deceived buying, which are perceived risk and benefits of social media promoted products, social media advertising informativeness and openness to change. The impact that the deceived buying has on propensity of believing and acting on disinformation was also analyzed. The results of this research highlighted that perceived benefits, risks, and openness to change significantly impact susceptibility to deceived buying, while the informativeness of the ad itself has no significant effect. Furthermore, deceived buying is strongly linked to a greater tendency to believe and act on disinformation.

The findings hold significant implications for researchers, businesses, social media companies and users, brands, organizations, governments and more, as social media is an omnipresent part of people's everyday lives. Hence, it is suggested for social media platforms to develop stronger algorithms that can identify deceptive content and fake ads. Second, educational programs can help users become savvier consumers of online information.

9.2 Theoretical Background and Hypothesis Development

Social media is an omnipresent part of our lives and it's becoming more and more important. In social media, influencers leverage their perceived authenticity to endorse products or services, blurring the lines between genuine recommendation and paid promotion [1, 7]. Hence, consumers can be swayed by the product choices of those they admire, which can lead to impulsive or uninformed purchases they may regret after [26]. Moreover, consumers are increasingly exposed to content that reinforces their existing beliefs [15], which can make them less likely to critically evaluate the veracity of information surrounding promoted products.

Considering the importance of this phenomenon and its effect on people's lives, and the effect that it holds on to branding trust and companies sales, analyzing what lead social media users to deceived buying, and to propensity of believing and acting on disinformation is essential. However, this is a wholly unexplored area of research. Thus, building upon the research model established by Kumar et al. [16], this study ventures into the critical domain of social media disinformation. Specifically, it investigates the phenomenon of deceived buying triggered by various factors and its link to the increased likelihood of the propensity of believing and acting on disinformation. Hence, various factors considered important in the context of impulsive buying behavior, social media advertisements or influencers endorsement were brought together in the proposed research model in the context of deceived buying, as potentially drivers of it. The following section provides a detailed analysis of each one, along with the developed hypotheses of this study.

The informativeness of social media ads, which refers to the amount of accurate and helpful details they provide about products or services, can be a double-edged sword. On the one hand, these ads can be highly informative, offering consumers in-depth product descriptions with features, benefits, and specifications [6]. They can integrate customer reviews and price comparisons, empowering users to make well-rounded purchase decisions [12]. However, the very same informativeness can be deceptive if the information itself is inaccurate or incomplete. Nam & Dân [19] showed that the quality of information positively influence customer buying intention. Arora et al. showed that informativeness have a positive and strong influence of attitude towards social media advertising, which lead to behavioral intention [3]. Hanaysha [13] also showed that informativeness has a direct and positive effect on purchase decision. Moreover, Nasir et al. [20] classified consumers under three segments, namely susceptible ones, dispassionate ones and impervious ones. They found that for consumers in the dispassionate segment, informativeness plays a major role in shaping their buying decisions, followed by ease of being persuaded, and social network proneness. Thus, considering the previous attention accorded to informativeness impact in buying behavior, the following hypothesis is developed:

H1: Informativeness positively influences the deceived buying.

When it comes to choose an advertised product to buy, perceived benefits and risks are very important [16]. They can be a deal maker or braker in the buying decision process. Kumar et al. [16] showed that the potential advantages individuals

can enjoy when using social media advertised products, namely the perceived benefits, are positively, direct, and strongly associated with purchase intentions. On the other hand, they showed that perceived risks of social media advertised products, that encompass the potential concerns and uncertainties that consumers associate with purchasing products promoted through social media has no effect on purchase intention. However, Martínez-López et al. [17] showed that a core element of buying behavior is perceived risk, which in their study have a negative, direct and very strong effect on intention to buy. Building on the previous results in the literature, we argue that consumer perception of both benefits and risks of social media advertised products can influence them to deceived buying. Hence, the following hypotheses were issued:

H2: Perceived benefits positively influence the deceived buying.

H3: Perceived risks negatively influence the deceived buying.

Openness to change is a personality factor influencing consumer behavior, characterized by a desire for novelty and independent decision-making, that makes consumers susceptible to the constant stream of new products advertised on social media [16]. Kumar et al. [16] previously highlighted a positive and strong indirect influence that openness to change has on purchase intention through perceived benefits moderator. Moreover, Tewari et al. [25] highlighted that openness to change has an indirect influence on purchase intention through attitude. However, until now, the influence that openness to change has on deceived buying was not studied. Hence, the next hypothesis was developed:

H4: Openness to change has a positive influence on deceived buying.

Deceived buying encompasses situations where consumers are influenced by inaccurate or incomplete information, hindering their ability to make informed decisions [10, 14, 21]. Previously, Kumar et al. [16] proposed a novel perspective regarding purchasing intentions in the context of fake news: consumers with a strong intention to buy a product might become more susceptible to believing and acting on fake news about that product. Their revealed that propensity of believing and acting on fake news had a strong, direct, and positive effect of purchase intentions. Awal et al. [5] also investigated how customer product purchase intentions influence their propensity of believing and acting on scam/fraud news related to online shopping. They revealed that male customers with online purchase intentions display lower tendencies to believe and act upon scam news regarding online shopping compared to female customers. Looking from another perspective, if consumers are convinced by social media advertised products into deceived buying, they can be more susceptible to propensity of believing and acting on disinformation on social media. Thus, the final hypothesis was issued:

H5: Deceived buying has a positive influence on propensity of believing and acting on disinformation.

9.3 Research Methodology, Results, and Discussion

9.3.1 Survey Instrument and Data Collection

This study employed an online survey with variables previously established in literature, adopted, or adapted on the current context, ensuring its relevance and comprehensiveness. Regarding informativeness construct, the measurement scale was adopted from Nasir et al. [20]. Perceived benefits, perceived risks, and openness to change measurement scales were adopted from Kumar et al. [16]. The deceived buying construct is a new construct proposed in this study, based on the Apuke and Omar research [2]. The propensity of believing and acting on disinformation scale was adapted from Kumar et al. [16]. A seven-point Likert scale anchored by “strongly disagree” (1) and “strongly agree” (7) was used. The survey was distributed to individuals that had bought social media advertised products. A total of 334 consumers participated in the study, with a gender breakdown of 61.98% female and 38.02% male. The age range spanned from 19 to 51 years old.

9.3.2 Data Analysis

To analyze the collected data, we utilized Partial Least Squares—Structural Equation Modeling (PLS-SEM) through SmartPLS software [22]. Given the exploratory nature of the current research, the PLS-SEM methodology is particularly suitable due to its ability to assess both direct and indirect relationships between variables while focusing on model quality evaluation. Moreover, it offers flexibility in analyzing complex models with multiple regressions, which has predictive power rather than simply explaining observed phenomena at the population level. Additionally, as is common in most social science research, the collected data exhibited a non-normal distribution, and the PLS-SEM technique performs well with such data. Hence, we evaluated the model’s reliability, internal consistency, convergent validity, and collinearity issues. The assessment for all these measures met established standards [11] confirming the model’s validation. The structural model visual representation is presented in Fig. 9.1.

Table 9.1 showcases the results of our analysis concerning construct reliability and validity. These measures, including indicator loadings, Cronbach’s alpha, composite reliability, and average variance extracted, fall within acceptable ranges [11]. This indicates a lack of any significant concerns regarding the model’s validity. Hence, the measures used have proven internally consistent and effectively capture the intended concepts they represent. This strengthens our confidence in the model’s overall reliability and validity.

Table 9.2 presents the results of the discriminant validity analysis using the heterotrait-monotrait ratio of correlations (HTMT). All HTMT values fall below the threshold of 0.85 [11], indicating that our constructs are distinct and capture unique

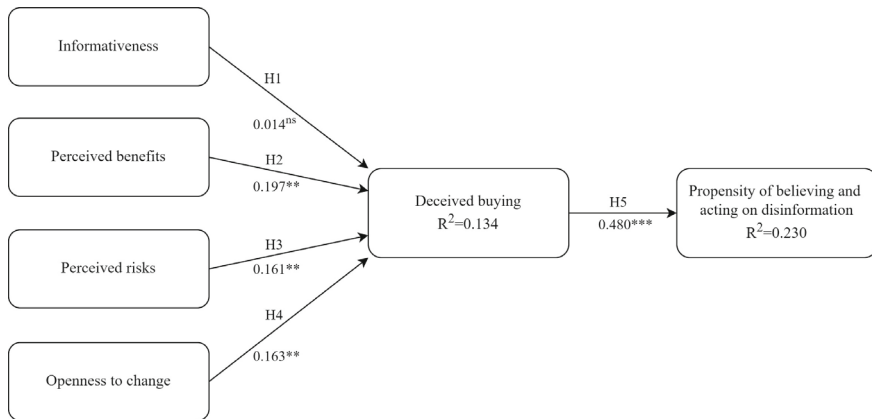


Fig. 9.1 Structural model relationships and obtained results

Table 9.1 Evaluation of convergent validity and internal consistency

Construct	Items	Factor loadings	Cronbach's alpha	Composite reliability	Average variance extracted
Deceived Buying (DB)	DB1	0.795	0.884	0.911	0.632
	DB2	0.837			
	DB3	0.785			
	DB4	0.766			
	DB5	0.774			
	DB6	0.809			
Informativeness (INF)	INF1	0.858	0.821	0.882	0.653
	INF2	0.840			
	INF3	0.822			
	INF5	0.705			
Openness to change (OC)	OC1	0.932	0.806	0.911	0.836
	OC2	0.896			
Perceived benefits (PB)	PB1	0.819	0.758	0.861	0.674
	PB2	0.799			
	PB3	0.844			
Propensity of believing and acting on disinformation (PBAD)	PBAD1	0.871	0.876	0.923	0.800
	PBAD2	0.891			
	PBAD3	0.883			
	PBAD4	0.828			
Perceived risks (PR)	PR1	0.863	0.892	0.925	0.755
	PR2	0.922			
	PR3	0.896			

Table 9.2 Discriminant validity (HTMT criterion)

	DB	INF	OC	PB	PR	PBAD
Deceived Buying (DB)						
Informativeness (INF)	0.213					
Openness to change (OC)	0.309	0.438				
Perceived benefits (PB)	0.344	0.692	0.393			
Perceived risks (PR)	0.226	0.074	0.221	0.16		
Propensity of believing and acting on disinformation (PBAD)	0.528	0.062	0.11	0.155	0.449	

Table 9.3 Outcomes of hypothesis testing

H #	Hypothesis	β	t	P	Support
H1	Informativeness → Deceived Buying	0.014	0.231	0.818	No
H2	Perceived benefits → Deceived Buying	0.197	2.786	0.006	Yes
H3	Perceived risks → Deceived Buying	0.161	2.843	0.005	Yes
H4	Openness to change → Deceived Buying	0.163	2.725	0.007	Yes
H5	Deceived Buying → Propensity of believing and acting on disinformation	0.480	10.315	0.000	Yes

concepts within the study. This further strengthens our confidence in the model’s ability to accurately assess the relationships between these constructs.

In the case of structural model assessment, VIF values fall below the recommended threshold, resulting in the absence of collinearity issues. The results of the bootstrapping procedure can be seen in Fig. 9.1 and Table 9.3. The R² values indicate that in the case of deceived buying, 13.4% of the variance of the final dependent variable is explained by its predictors included in the model, and 23.0% in the case of propensity of believing and acting on disinformation.

9.3.3 Discussion, Limitations, and Future Research

This study empirically analyzed four variables that might drive people into deceived buying, and how deceived buying can lead to propensity of believing and acting on disinformation. These variables are informativeness, perceived benefits and risks, and openness to change.

Concerning hypothesis H1, the findings showed that informativeness of social media ads has no effect on deceived buying ($\beta = 0.014$; $p > 0.05$) which means that this hypothesis was not confirmed. The results differ from previous findings in the context of customer buying intention [19], behavioral intention [3], and purchase decision [13, 20]. There are two possible explanations for this result: either the persuasion

in the advertisements overwhelms the information presented, or the informativeness does not focus on the essential, truthful details that would influence the deceived buying.

On the other hand, hypothesis H2 and H3 were confirmed, as both perceived benefits ($\beta = 0.197$; $p < 0.01$) and perceived risks of social media promoted products ($\beta = 0.161$; $p < 0.01$) have a direct, positive, and strong impact on deceived buying. The results regarding perceived benefits are similar with those obtained by [16], while the findings regarding perceived risks are not in line with those of [17]. Hence, in this case, a higher level of perceived risks can lead consumers to the deceived buying.

Hypothesis H4 was also confirmed, the results showing that openness to change has a direct, positive, and strong impact on deceived buying ($\beta = 0.163$; $p < 0.01$). This finding is in line with previous research by Kumar et al. and Tewari et al. [16, 25].

Finally, hypothesis H5 was also confirmed, the findings highlighting a direct, positive, and very strong impact that deceived buying has on the propensity of believing and acting on disinformation ($\beta = 0.480$; $p < 0.001$). The result is in line with previous finding of Kumar et al. [16] in the fake news context, and Awal et al. [5] in the context of scam/fraud news related to online shopping.

There are certain limitations inherent in current research. For example, this study utilized an Eastern European sample, potentially limiting the generalizability of the findings to other cultures and countries. Hence, future research on an international scale could leverage this validated model to explore potential cultural disparities or variations in motivations across different populations. For future investigations, we recommend that scholars in this field employ the research model in their inquiries, augmenting it with diverse variables to observe shifts in the behavior of social media users. In addition, this research model can be used for analyzing the deceived buying behavior of different brands or companies' customers and consumers.

9.4 Conclusions and Implications

This research investigates the factors influencing deceived buying of social media promoted products, that leads further to the propensity of believing and acting on disinformation. The findings demonstrate that perceived benefits associated with social media advertised products, alongside the perceived risks of such products, and an individual's openness to change, all exert a strong influence on the likelihood of engaging in deceived buying. The informativeness of the advertisement itself does not have a statistically significant impact on deceived buying. The research reveals a strong positive correlation between deceived buying and the propensity of believing and acting on disinformation.

The results provide significant contributions both theoretically and practically. From a theoretical perspective, this research establishes a novel link between deceived buying and the tendency to believe and act on disinformation suggesting

that falling victim to deceptive marketing practices leads individuals more susceptible to believing and spreading false information online. From a practical perspective, this study offers valuable insights for various stakeholders. Policymakers can leverage the findings to craft regulations safeguarding consumers from deceptive online marketing. Social media platforms can utilize this knowledge to refine strategies and implement algorithms for identifying and eliminating misleading advertisements. Consumer protection organizations are empowered to develop educational tools based on this research, fostering informed online purchasing decisions. Businesses, too, can benefit by adopting more ethical and transparent marketing practices that minimize the risk of deceiving consumers. Ultimately, this research equips stakeholders with a deeper understanding of factors influencing susceptibility to deceived buying, paving the way for effective strategies to combat this issue.

Acknowledgements This work was funded by a grant of the Ministry of Research, Innovation and Digitization through Programme 1—Development of the national research-development system, Subprogramme 1.2—Institutional performance—Projects to fund excellence in RDI, contract no. 21PFE/30.12.2021 code/ID PFE-550-UBB.

References

1. AlFarraj, O., et al.: Examining the impact of influencers' credibility dimensions: attractiveness, trustworthiness and expertise on the purchase intention in the aesthetic dermatology industry. *Rev. Intern. Bus. Strat.* **31**(3), 355–374 (2021). <https://doi.org/10.1108/RIBS-07-2020-0089>
2. Apuke, O.D., Omar, B.: Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telemat. Inform.* **56**, 101475 (2021). <https://doi.org/10.1016/J.TELE.2020.101475>
3. Arora, T. et al.: Impact of social media advertising on millennials buying behaviour. *Intern. J. Intell. Enterp.* **7**, 4, 481 (2020). <https://doi.org/10.1504/IJIE.2020.110795>
4. Aslam, H. et al.: Impact of personalized social media advertising on online impulse buying behavior. *SEISENSE Bus. Rev.* **1**, 3, 12–25 (2021). <https://doi.org/10.33215/sbr.v1i3.660>
5. Awal, M.R., et al.: Application of the SOBC model to study customers' online purchase intentions in an emerging economy during COVID-19: does gender matter? *PSU Research Review*. (2023). <https://doi.org/10.1108/PRR-02-2022-0014>
6. Van den Broeck, E., et al.: How do users evaluate personalized Facebook advertising? An analysis of consumer- and advertiser controlled factors. *Qual. Mark. Res.* **23**(2), 309–327 (2020). <https://doi.org/10.1108/QMR-10-2018-0125>
7. Cheung, M.L. et al.: "I follow what you post!": The role of social media influencers' content characteristics in consumers' online brand-related activities (COBRAs). *J. Retail. Cons. Serv* **66**, February, 102940 (2022). <https://doi.org/10.1016/j.jretconser.2022.102940>
8. Sitar-Tăut, D.-A. et al.: Customer behavior in the prior purchase stage—information search versus recommender systems. *Econ. Comput. Econ. Cybern. Stud. Res.* **54**, 3/2020, 59–76 (2020). <https://doi.org/10.24818/18423264/54.3.20.04>
9. Domenico, G. Di et al.: Fake news, social media and marketing: A systematic review. *J. Bus. Res.* **124**, December 2020, 329–341 (2021). <https://doi.org/10.1016/j.jbusres.2020.11.037>
10. Gaber, H. R., Labib, A. A., Salem, K.O.: The effect of marketing deception on consumer buying decision on facebook. an empirical study on university students in Libya. *Eur. J. Bus. Innov. Res.* **6**(3), 12–18 (2018)

11. Hair, J.F. et al.: A primer on partial least squares structural equation modeling (PLS-SEM). Sage (2022)
12. Hamouda, M.: Understanding social media advertising effect on consumers' responses: An empirical investigation of tourism advertising on Facebook. *J. Enterp. Inf. Manag.* **31**(3), 426–445 (2018). <https://doi.org/10.1108/JEIM-07-2017-0101>
13. Hanaysha, J.R.: Impact of social media marketing features on consumer's purchase decision in the fast-food industry: Brand trust as a mediator. *Intern. J. Inform. Manag. Data Insights.* **2**, 2, 100102 (2022). <https://doi.org/10.1016/j.jjime.2022.100102>
14. Held, Johanna, Germelmann, C.: Deceived or not deceived? : How food consumers perceive deception. *Advances in consumer research. Assoc. Cons. Res. (U.S.)*. **42**, 313–317 (2014)
15. Jacobson, J. et al.: Social media marketing: Who is watching the watchers? *J. Retail. Cons. Serv.* **53**, September 2018, 101774 (2020). <https://doi.org/10.1016/j.jretconser.2019.03.001>
16. Kumar, S., et al.: Purchasing natural personal care products in the era of fake news? The moderation effect of brand trust. *J. Retail. Consum. Serv.* **63**, 102668 (2021). <https://doi.org/10.1016/j.jretconser.2021.102668>
17. Martínez-López, F.J., et al.: Buying through social platforms. *J. Organ. End User Comp.* **33**(4), 70–93 (2021). <https://doi.org/10.4018/JOEUC.20210701.oa4>
18. Mican, D. et al.: User perceptions of fake news sharing behavior on social media through social networking sites. *Bus. Manag. Rev.* **13**, 01, 11–12 (2022). <https://doi.org/10.24052/BMR/V13 NU01/ART-14>
19. Nam, L.G., Dân, H.T.: Impact of social media Influencer marketing on consumer at Ho Chi Minh City. *Intern. J. Soc. Sci. Human. Invent.* **5**, 5, 4710–4714 (2018). <https://doi.org/10.18535/ijsshi/v5i5.10>
20. Nasir, V.A., et al.: Segmenting consumers based on social media advertising perceptions: How does purchase intention differ across segments? *Telematics Inform.* **64**, 101687 (2021). <https://doi.org/10.1016/j.tele.2021.101687>
21. Nisa, N. un et al.: The consequences of upholding deceptive marketing practices and its effect on consumers' related attitudes including consumer loyalty and consumer trust: Evidence from the fashion industry in the UAE. *J. Adv. Res. Soc. Sci. Human.* **7**, 2, (2022). <https://doi.org/10.26500/JARSSH-07-2022-0203>
22. Ringle, C.M. et al.: SmartPLS 4, (2022)
23. Siano, A., et al.: "More than words": Expanding the taxonomy of greenwashing after the Volkswagen scandal. *J. Bus. Res.* **71**, 27–37 (2017). <https://doi.org/10.1016/j.jbusres.2016.11.002>
24. Sterie, L.-G. et al.: Analyzing the antecedents of fake news sharing in online social networks. *Smart Innovation, Systems and Technologies*, vol 367. Springer, Singapore (2024). https://doi.org/10.1007/978-981-99-6529-8_13
25. Tewari, A., et al.: Examining the role of receptivity to green communication, altruism and openness to change on young consumers' intention to purchase green apparel: A multi-analytical approach. *J. Retail. Consum. Serv.* **66**, 102938 (2022). <https://doi.org/10.1016/j.jretconser.2022.102938>
26. Wasike, B.: When the influencer says jump! How influencer signaling affects engagement with COVID-19 misinformation. *Soc Sci Med* **315**, 115497 (2022). <https://doi.org/10.1016/j.socsci med.2022.115497>

Chapter 10

Driving Factors of Social Commerce Intention: The Role of Social Commerce Constructs, Social Influence, and Trust



Diana-Ştefania Ghenie , Virgil-Vasile Şoncutean ,
and Dan-Andrei Sitar-Tăut 

Abstract The emergence of social commerce has revolutionized how consumers engage in purchasing experiences over social media platforms due to its novel technical features and interaction possibilities in online communities or forums. This research aims to assess how social commerce constructs, social influence, and trust, as a main driver, influence usage intention. Two conceptual frameworks were selected and combined to emphasize the structures, social support (informational and emotional), and perceived value (hedonic and utilitarian), resulting in a final strong relationship between trust and usage intention. Based on a sample of 299 social commerce users, the proposed research model was tested using the PLS-SEM methodology. The study's findings demonstrated that almost all hypotheses were confirmed, which fulfilled our research gap and postulated that social commerce constructs and social influence have a positive effect on usage intention, mediated by trust. However, one relationship was not confirmed, regarding the effect of the utilitarian value on usage intention.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-981-96-0161-5_10.

D.-Ş. Ghenie (✉) · V.-V. Şoncutean · D.-A. Sitar-Tăut
Business Information Systems Department, Faculty of Economics and Business Administration,
Babeş-Bolyai University, Cluj-Napoca, Romania
e-mail: diana.ghenie@econ.ubbcluj.ro

V.-V. Şoncutean
e-mail: virgil.soncutean@econ.ubbcluj.ro

D.-A. Sitar-Tăut
e-mail: dan.sitar@econ.ubbcluj.ro

10.1 Introduction

Social commerce (or s-commerce) encompasses the use of social media platforms for online sales [1, 2] and Web 2.0 technologies in order to support online consumer interactions, and internet purchases before, during and after acquisitions [3]. Social commerce constructs (or SCCs) allow users to create content, assess goods and services and share their experiences, serving as a source of online social support [4]. This social support consists of both informational (cognitive) and emotional (affective) states [5, 6]. These are in line with other research showing that users' information generated and shared on social media influences other customers' trust when engaging in s-commerce activities [4]. Perceived value consists of hedonic value (affective state) and utilitarian value (cognitive state), distinct dimensions [7]. Researchers have also discovered that the tendency of consumers to utilize social commerce is positively correlated with trust [8–13].

The main purpose of this study is to assess the influence of SCCs and social influence, mediated by constructor trust on usage intention. A quantitative study has been conducted with the support of two conceptual frameworks adapted and expanded upon from two specific studies [6, 14]. Particularly, we evaluated the social constructs that explain trust in the context of online communities, facilitated by s-commerce platforms and conversely, this research assessed a secondary research structure, which incorporated facets of social influence and its impact on the intention to utilize s-commerce for purchases. The relationship among SCCs, social influence, trust, and usage intention has not been fully investigated in prior research. The fusion of two research frameworks resulted in a final strong relationship between trust and usage intention, which fulfills the gap of our research purpose. The analysis is drawn on the stimulus–organism–response (SOR) model, where **stimulus** incorporates SCCs and social influence, **organism** encompasses social support (emotional and informational) [10] and perceived value (hedonic and utilitarian value) [15], and the **response** is represented by trust and usage intention, both variables being investigated as stimulants for using these services [16]. The results demonstrated that the hypotheses were confirmed, except one, the relationship between utilitarian value and usage intention.

10.2 Theoretical Foundation

SCCs, in the form of forums, communities, ratings, reviews, recommendations, and referrals, create a more friendly environment where users discuss their experiences [11]. This fosters a more trustworthy environment for s-commerce platforms and increases social support [17]. Previous studies [18, 19] suggest that social influence is a crucial factor that influences people's attitudes and intentions toward a particular behavior.

Social support, in the context of s-commerce refers to the emotional and informational support that users of social networking sites (or SNSs) and online communities receive [17, 20, 21]. Emotional support involves affective and relational characteristics, which may strengthen social bond between peers, generating trust [22]. On the other side, informational support focuses on offering brief guidance or answers, which can assist customers in directly learning about a good or service [21]. Customers usually purchase commercial goods and services and consume them for two main motives: utilitarian advantages or functional and hedonic awareness, derived from sensory features. Higher perceived value users (such as those with hedonic and utilitarian values) are more likely to experience satisfaction and to find s-commerce platforms as reliable, which in turn generate purchase intentions [23].

Users' decisions to adopt and believe in different technologies are mostly determined by their experiences and interactions, which in turn shape their perception of trust [19, 24]. This aligns with the perspective of other studies that have claimed that SCCs encourage and enhance trust, contributing to a boost in customers' intention to make purchases [6, 25]. Research on recommendations and usage habits on social media platforms has demonstrated how social interaction affects a user's perception toward a particular social media platform, which in turn affects recommendations and intention to use [26].

10.3 Research Model and Hypotheses Development

Consumers create interactions through SCCs that in turn enhance the provision of social support [5]. This will help users in receiving social support within s-commerce platforms. As a result, we think that applying SCCs will enhance users' social support [8, 27]. The following research hypotheses are proposed:

H1a: Social commerce constructs have a positive influence on emotional support

H1b: Social commerce constructs have a positive influence on informational support

Social influence may impact usage intention by influencing users' perception of value. According to the social influence hypothesis, the volume of SNSs users will have an impact on the perceived value [14]. This affirmation suggests that social interactions have an impact on an individual's behavior [18]. If a s-commerce platform makes it possible for more individuals to interact and share information with them, perceived value can rise [14]. The following hypotheses are presented in this study:

H2a: Social influence has a positive influence on hedonic value

H2b: Social influence has a positive influence on utilitarian value

Trust is developed and purchase decisions are subsequently guided by the availability of social support, which includes both emotional and informational support [19]. When consumers receive accurate, timely, and credible information

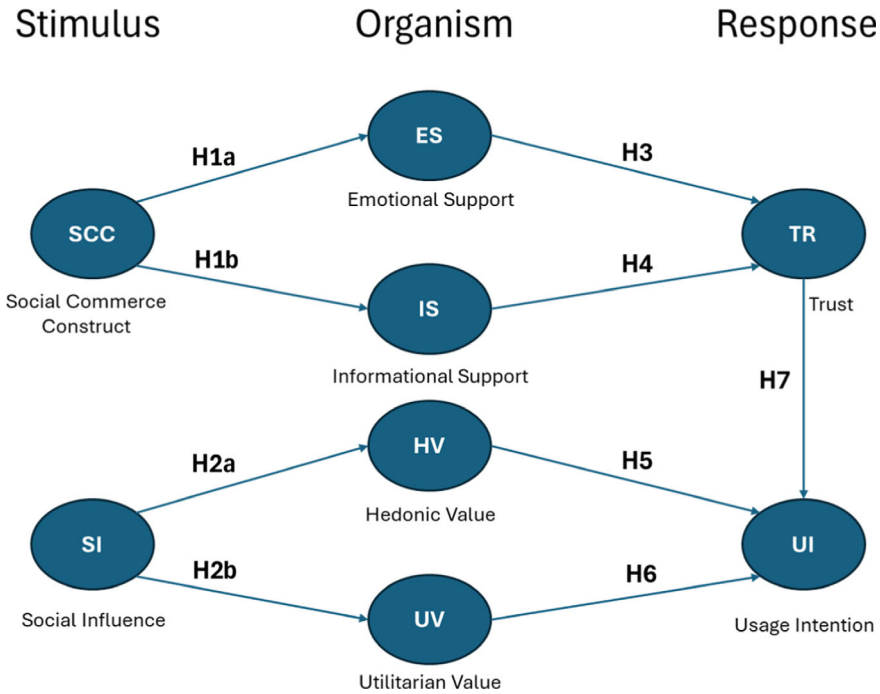


Fig. 10.1 Conceptual model

and resources from s-commerce networks, their level of trust increases [24, 28]. As a result, we propose the following hypotheses:

H3: Emotional support has a positive influence on trust

H4: Informational support has a positive influence on trust

Users are more likely to develop contentment and intend to utilize the s-commerce platforms for buying when they have a positive experience and perceive interest and enjoyment from the shopping process. Further evidence provides that the reliability of the information has a favorable impact on users' perceived value, which we infer is leading to increased users' intentions to use the social commerce system [12, 23].

H5: Hedonic value has a positive influence on usage intention

H6: Utilitarian value has a positive influence on usage intention

SCCs are applied in the context of s-commerce to analyze variations in usage intention [4, 17]. Additionally, different SCCs have different effects on risk and trust, which ultimately influence users' intentions for using specific platforms [16]. The following hypothesis is proposed:

H7: Trust influences usage intention positively (Fig. 10.1).

10.4 Research Methodology

10.4.1 Survey and Scales

The model proposed in this study is built on two models which are joined by two variables we consider appropriate connectors in this research environment. The scales we adopted and adapted have been previously validated. Thus, the scales for SCC, ES, IS, and TR were taken from Shanmugam et al. [6], since those for SI, UV, HV, and UI from Y.-S. Yen [14]. All the items were scored using a 7-point Likert scale.

10.4.2 Sampling and Data Collection

The study was conducted concerning the confidentiality and anonymity of the responders, subjected to an informed consent. Data were collected in 2023 using Google Forms. The eligibility criterion was the existence of experience with s-commerce platforms. The data cleaning process consisted of incomplete responses, duplicates, and straightliners removal. Finally, a number of 299 responses were retained, being distributed by gender, age, and graduation level, as follows. There were 66.56% females, 32.44% males, 1.00% diverse. By age, the sample distribution consists of 27.76% s-commerce consumers under 21 years, almost half (47.49%) are between 21 and 22 years old, since 24.75% are older than 22 years. They declare the higher graduation level is high school (82.61%), bachelor's (12.71%), and 4.68% graduated from a master's program or above.

The PLS-SEM methodology was successfully used in survey-related studies from various fields, such as behavior in social media [29], e-commerce tackling [30], or s-commerce analysis [31]. The methodology we have previously mentioned is used to confirm and assemble complex reflective, and formative models, as well. PLS-SEM works well with small samples without prerequisite tests regarding distribution's normality. To assess the launched hypotheses, a set of preliminary tests is run, and quality indicators are computed, as depicted in Hair et al. [32]. The data analysis was accomplished by employing SmartPLS v4.1.0.0 software [33].

10.5 Data Analysis, Results

Both, outer and inner model assessments are driven by steps and procedures outlined by Hair et al.'s [32] work.

Table 10.1 Discriminant validity assessment

Construct	ES	HV	IS	SCC	SI	TR	UI	UV
ES								
HV	0.493							
IS	0.859	0.393						
SCC	0.705	0.434	0.665					
SI	0.223	0.658	0.199	0.209				
TR	0.585	0.566	0.59	0.535	0.348			
UI	0.529	0.633	0.496	0.445	0.533	0.777		
UV	0.463	0.784	0.467	0.484	0.568	0.607	0.568	

Note Values represent heterotrait–monotrait ratios of correlations (HTMT values)

10.5.1 Outer Model Assessment

The outer or measurement assessment in reflective models consists of performing indicator reliability and internal consistency reliability tests, followed by convergent and discriminant validity assessments. The lowest outer loading value is 0.713 and belongs to SI1. Thus, the theoretical reference of 0.7 is accomplished for all indicators. Composite reliability (CR) and Cronbach's Alpha (CA) indicators are considered and calculated to evaluate the internal consistency reliability and then compared with the lower accepted threshold of 0.7. Since this test is passed, there are no internal consistency issues identified. Furthermore, Average Variance Extracted (AVE) computation is essential to determine convergent validity legitimacy. Since the lower AVE value is 0.670, for SI, the lower boundary of 0.5 is passed by all the constructs. Consequently, the convergent validity is validated. Finally, at this measurement level assessment, heterotrait–monotrait (HTMT) is involved in discriminant validity evaluation [34]. This work stipulates a superior threshold of 0.85, which is passed by all construct pairs (see Table 10.1), except the IS–ES ratio.

The related value of 0.859 slightly exceeds the theoretical limit. Even so, additional tests are required [35]. According to this, we had to compute the percentile bootstrap confidence interval for IS-ES. Discovering that its HTMT value deviates significantly from 1, we can affirm that no discriminant validity issues were revealed in our data.

10.5.2 Inner Model Assessment

The inner or structural model assessment investigates collinearity presence issues, relationships' significance and strength, and determining the explanatory power. Because the highest value of the inner VIF is 2.818, which is below the most severe threshold of 3, then we can argue that our model is free of any collinearity problems.

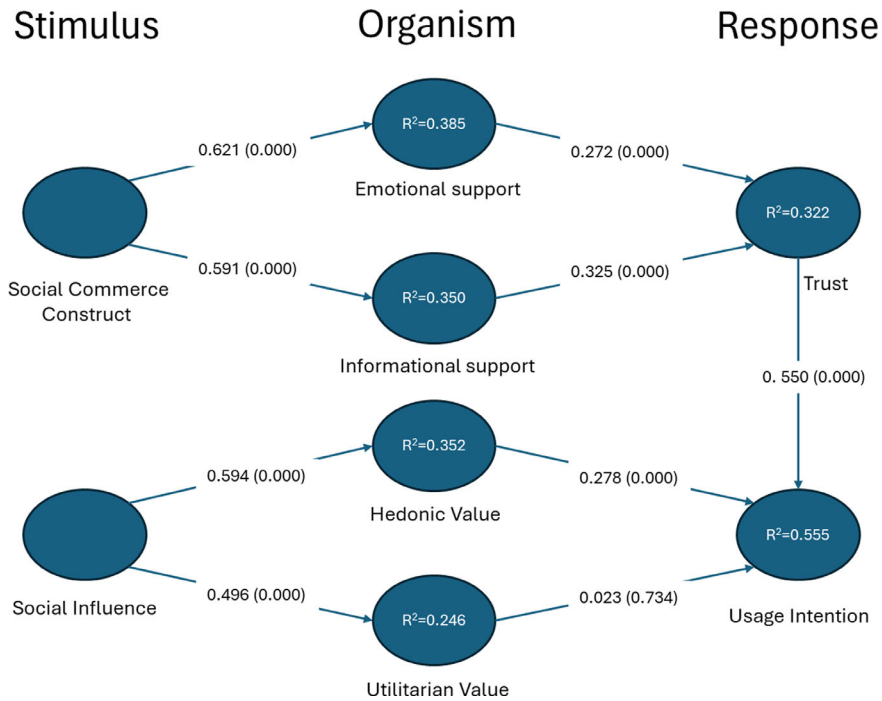


Fig. 10.2 Research model with path coefficients and R2 values. Note *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$

Furthermore, the explanatory power is checked, being given the R2 values of endogenous factors. The final target construct, usage intention, reveals a high predictive accuracy ($R^2 = 0.555$). Moreover, the two immediate predictors, HV and TR ($R^2 = 0.352$ and $R^2 = 0.322$) experience moderate predictive accuracies (Fig. 10.2).

Finally, we explored the significance and the path coefficients associated to the hypotheses we formulated. The bootstrapping involves iterations on 10,000 subsamples. A number of seven out of eight hypotheses were supported. Their magnitude and significance are presented in Table 10.2 and Fig. 10.2.

The most important relationships are between SCC and ES ($\beta = 0.621$), followed by SI and HV ($\beta = 0.594$), being very close to the third, between SCC and IS ($\beta = 0.591$). Oppositely, the most fragile associations are among HV and UI ($\beta = 0.278$), and between ES and TR ($\beta = 0.272$). Even if these are the weakest relationships, their strength is still moderate. However, the link between UV and UI was not confirmed.

Survey questions used to validate the hypotheses in the research model, Table S1, are included in the Supplementary material.

Table 10.2 Structural model assessment—direct effects

#	H#	Path	Effect	t statistics	p values
Direct effects					
1	H1a	SCC → ES	0.621	15.443	0.000
2	H1b	SCC → IS	0.591	13.656	0.000
3	H2a	SI → HV	0.594	15.917	0.000
4	H2b	SI → UV	0.496	9.230	0.000
5	H3	ES → TR	0.272	4.037	0.000
6	H4	IS → TR	0.325	4.751	0.000
7	H5	HV → UI	0.278	4.076	0.000
8	H6	UV → UI	0.023	0.340	0.734
9	H7	TR → UI	0.550	10.707	0.000

10.6 Discussion, Implications, Limitations, Conclusion

The results indicate that our two core variables, social constructs, and social influence significantly influence usage intention mediated by the trust variable, the relationship between the two revealing a quite strong relationship of $\beta = 0.550$ and our final target construct, usage intention, indicating a high predictive accuracy ($R^2 = 0.555$).

Firstly, our research demonstrates that user involvement and communication via SCCs builds an emotional and informational sense of support among social network users. These findings support also earlier research that demonstrated these results [5, 6]. Likewise, informational and emotional support have a positive impact on trust, which is in line with further studies [10, 36]. One interesting aspect that emerged from the analysis is that the relation between utilitarian value and usage intention was not confirmed. Although some studies confirm that utilitarian value has an impact on purchase intention, satisfaction, or user behavior [23, 37], hedonic value had a greater impact than utilitarian value on the final construct. Thus, our results are consistent with the study conducted by Y.-S. Yen. [14]. Hedonistic purchasing is connected to happiness and can satisfy customers' experience demands. We can presume that with the evolution of social media and its innovative features, like promoting short videos ads, s-commerce has transformed the way to promote products and services in an enjoyable, funny, and satisfying way. Rather than practicality or helpful information, social network interactions and entertainment are the consumers' primary social incentives. Our final target relationship between trust and usage intention was also positively confirmed. This result is in line with the literature mentioned above [4, 16, 17]. From the theoretical point of view, our research adopted the S-O-R framework and tried to fill the gap through the fusion of the two research models, which demonstrated that our investigation gathered further essential factors that can influence the adoption of s-commerce platforms for online purchases. Trust can be considered an essential incentive that increases the likelihood of usage intention of these platforms, when a sense of community and hedonistic experiences are available. Thus,

we argue that our addition to the current literature, by linking two research models and discovering a strong relationship between the final constructs, lies in building a new conceptual structure that has not been earlier provided. Our results suggest insight into how managers or vendors could employ SCCs and social support factors to develop trust among online participants in an online community. When users can engage and produce information on social media platforms through SCCs, it significantly boosts trust. Vendors should focus on promoting a feeling of community among customers increasing their desire to engage in s-commerce.

Some limitations should be considered. First, the empirical results are not fully generalizable. Our work took into consideration s-commerce consumers or users on s-commerce platforms; therefore, the study might not be applicable to some specific social commerce contexts. Further investigations should take into consideration different s-commerce platforms. Also, the representativeness of the sample is constrained. In order to improve the study's accuracy, larger data from different countries, geographical areas, or corporate employees should be taken into consideration. Third, considering that utilitarian value has no effect on usage intention, future investigations should be conducted in order to test and examine this relationship. In order to evaluate trust-building in s-commerce platforms, we chose to conduct an empirical analysis in order to verify our conceptual structure and data were gathered from s-commerce users. The findings indicate that both SCCs and social influence have a beneficial impact on usage intention, mediated by trust, but it was discovered that utilitarian value has no effect on usage intention.

References

1. Liang, T.-P., Turban, E.: Introduction to the special issue social commerce: a research framework for social commerce. *Int. J. Electron. Commer.* **16**(2), 5–14 (2011). <https://doi.org/10.2753/JEC1086-4415160201>
2. Lin, X., Li, Y., Wang, X.: Social commerce research: definition, research themes and the trends. *Int. J. Inf. Manage.* **37**(3), 190–201 (2017). <https://doi.org/10.1016/j.ijinfomgt.2016.06.006>
3. Meilatinova, N.: Social commerce: factors affecting customer repurchase and word-of-mouth intentions. *Int. J. Inf. Manage.* **57**, 102300 (2021). <https://doi.org/10.1016/j.ijinfomgt.2020.102300>
4. Hajli, N.: Social commerce constructs and consumer's intention to buy. *Int. J. Inf. Manage.* **35**(2), 183–191 (2015). <https://doi.org/10.1016/j.ijinfomgt.2014.12.005>
5. Hajli, N., Sims, J.: Social commerce: the transfer of power from sellers to buyers. *Technol. Forecast. Soc. Change* **94**, 350–358 (2015). <https://doi.org/10.1016/j.techfore.2015.01.012>
6. Shanmugam, M., Sun, S., Amidi, A., Khani, F., Khani, F.: The applications of social commerce constructs. *Int. J. Inf. Manage.* **36**(3), 425–432 (2016). <https://doi.org/10.1016/j.ijinfomgt.2016.01.007>
7. Zheng, X., Men, J., Yang, F., Gong, X.: Understanding impulse buying in mobile commerce: an investigation into hedonic and utilitarian browsing. *Int. J. Inf. Manage.* **48**, 151–160 (2019). <https://doi.org/10.1016/j.ijinfomgt.2019.02.010>
8. Li, C.-Y.: How social commerce constructs influence customers' social shopping intention? An empirical study of a social commerce website. *Technol. Forecast. Soc. Change* **144**, 282–294 (2019). <https://doi.org/10.1016/j.techfore.2017.11.026>

9. Zhao, W., Hu, F., Wang, J., Shu, T., Xu, Y.: A systematic literature review on social commerce: assessing the past and guiding the future. *Electron. Commer. Res. Appl.Commer. Res. Appl.* **57**, 101219 (2023). <https://doi.org/10.1016/j.elerap.2022.101219>
10. Alalwan, A.A., et al.: Examining the impact of social commerce dimensions on customers' value cocreation: the mediating effect of social trust. *J. Consum. Behav.* **18**(6), 431–446 (2019). <https://doi.org/10.1002/cb.1782>
11. Kim, S., Park, H.: Effects of various characteristics of social commerce (s-commerce) on consumers' trust and trust performance. *Int. J. Inf. Manage.* **33**(2), 318–332 (2013). <https://doi.org/10.1016/j.ijinfomgt.2012.11.006>
12. Hajli, N., Sims, J., Zadeh, A.H., Richard, M.-O.: A social commerce investigation of the role of trust in a social networking site on purchase intentions. *J. Bus. Res.* **71**, 133–141 (2017). <https://doi.org/10.1016/j.jbusres.2016.10.004>
13. Lu, B., Fan, W., Zhou, M.: Social presence, trust, and social commerce purchase intention: An empirical research. *Comput. Human Behav.* **56**, 225–237 (2016). <https://doi.org/10.1016/j.chb.2015.11.057>
14. Yen, Y.-S.: The relationship among social influence, perceived value, and usage intention in social networking sites, in *2013 3rd International Conference on Consumer Electronics, Communications and Networks*, IEEE, pp. 699–702 (2013). <https://doi.org/10.1109/CECNet.2013.6703427>
15. Hsiao, K.-L., Chen, C.-C.: What drives in-app purchase intention for mobile games? An examination of perceived values and loyalty. *Electron. Commer. Res. Appl.Commer. Res. Appl.* **16**, 18–29 (2016). <https://doi.org/10.1016/j.elerap.2016.01.001>
16. Bugshan, H., Attar, R.W.: Social commerce information sharing and their impact on consumers. *Technol. Forecast. Soc. Change* **153**, 119875 (2020). <https://doi.org/10.1016/j.techfore.2019.119875>
17. Sheikh, Z., Yezheng, L., Islam, T., Hameed, Z., Khan, I.U.: Impact of social commerce constructs and social support on social commerce intentions. *Inf. Technol. People* **32**(1), 68–93 (2019). <https://doi.org/10.1108/ITP-04-2018-0195>
18. Zhao, K., Stylianou, A.C., Zheng, Y.: Sources and impacts of social influence from online anonymous user reviews. *Inf. Manage.* **55**(1), 16–30 (2018). <https://doi.org/10.1016/j.im.2017.03.006>
19. Asanprakit, S., Kraiwant, T.: Causal factors influencing the use of social commerce platforms. *J. Open Innov. Technol. Mark. Complex.* **9**(4), 100172 (2023). <https://doi.org/10.1016/j.joitmc.2023.100172>
20. Hu, X., Chen, Z., Davison, R.M., Liu, Y.: Charting consumers' continued social commerce intention. *Internet Res.* **32**(1), 120–149 (2022). <https://doi.org/10.1108/INTR-07-2020-0397>
21. Wang, Y.: Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). *Comput. Human Behav.* **110**, 106373 (2020). <https://doi.org/10.1016/j.chb.2020.106373>
22. Hu, X., Chen, X., Davison, R.M.: Social support, source credibility, social influence, and impulsive purchase behavior in social commerce. *Int. J. Electron. Commer.Commer.* **23**(3), 297–327 (2019). <https://doi.org/10.1080/10864415.2019.1619905>
23. Gan, C., Wang, W.: The influence of perceived value on purchase intention in social commerce context. *Internet Res.* **27**(4), 772–785 (2017). <https://doi.org/10.1108/IntR-06-2016-0164>
24. Yeon, J., Park, I., Lee, D.: What creates trust and who gets loyalty in social commerce? *J. Retail. Consum. Serv.* **50**, 138–144 (2019). <https://doi.org/10.1016/j.jretconser.2019.05.009>
25. Al-Tit, A.A., Omri, A., Hadj, T.B.: The driving factors of the social commerce intention of Saudi Arabia's online communities. *Int. J. Eng. Bus. Manag.* **12**, 184797901989974 (2020). <https://doi.org/10.1177/1847979019899746>
26. Curras-Perez, R., Ruiz-Mafe, C., Sanz-Blas, S.: Determinants of user behaviour and recommendation in social networks. *Ind. Manag. Data Syst.Manag. Data Syst.* **114**(9), 1477–1498 (2014). <https://doi.org/10.1108/IMDS-07-2014-0219>

27. Riaz, M.U., Guang, L.X., Zafar, M., Shahzad, F., Shahbaz, M., Lateef, M.: Consumers' purchase intention and decision-making process through social networking sites: a social commerce construct. *Behav. Inf. Technol.* **40**(1), 99–115 (2021). <https://doi.org/10.1080/0144929X.2020.1846790>
28. Yang, X.: Exchanging social support in social commerce: the role of peer relations. *Comput. Human Behav.* **124**, 106911 (2021). <https://doi.org/10.1016/j.chb.2021.106911>
29. Sitar-Taut, D.-A., Micán, D., Frömbing, L., Sarstedt, M.: Digital socialligators? Social media-induced perceived support during the transition to the COVID-19 lockdown. *Soc. Sci. Comput. Rev.Comput. Rev.* **41**(3), 748–767 (2023). <https://doi.org/10.1177/08944393211065872>
30. Micán, D., Sitar-Taut, D.-A.: The effect of perceived usefulness of recommender systems and information sources on purchase intention. *Kybernetes* (2023). <https://doi.org/10.1108/K-08-2022-1145>
31. Mousavi, S., Roper, S.: Enhancing relationships through online brand communities: comparing posters and Lurkers. *Int. J. Electron. Commer.Commer.* **27**(1), 66–99 (2023). <https://doi.org/10.1080/10864415.2022.2158596>
32. Hair, J. F., Hult, G.T.M., Ringle, C.M., Sarstedt, M.: A primer on partial least squares structural equation modeling (PLS-SEM), 3rd ed. Sage (2022)
33. Ringle, C.M., Wende, S., Becker, J.-M.: SmartPLS 4. SmartPLS GmbH (2022)
34. Henseler, J., Ringle, C.M., Sarstedt, M.: A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **43**(1), 115–135 (2015). <https://doi.org/10.1007/s11747-014-0403-8>
35. Franke, G., Sarstedt, M.: Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. *Internet Res.* **29**(3), 430–447 (2019). <https://doi.org/10.1108/IntR-12-2017-0515>
36. Zhang, M., Liu, Y., Wang, Y., Zhao, L.: How to retain customers: Understanding the role of trust in live streaming commerce with a socio-technical perspective. *Comput. Human Behav.* **127**, 107052 (2022). <https://doi.org/10.1016/j.chb.2021.107052>
37. Pang, H.: Identifying associations between mobile social media users' perceived values, attitude, satisfaction, and eWOM engagement: the moderating role of affective factors. *Telemat. Inform.* **59**, 101561 (2021). <https://doi.org/10.1016/j.tele.2020.101561>

Chapter 11

Strategic Collaboration in International University Networks—Case Study



Margarita Bogdanova , Evelina Parashkevova-Velikova ,
and Yuriy Kuznetsov 

Abstract The purpose of the research is to explore the conditions for transforming project partnerships into a strategic network, identify the key challenges facing individual participants, and outline research intentions for analyzing the transformation from operational to strategic interaction within the network. A case study of a university alliance of the non-spatial collaboration type is presented. The study of network connections through social network analysis is in its early stages but patterns of interaction are already emerging—subgroups (clusters) for which data are being generated for future research. GEPHI is applied for visualization. Due to the early stage of development of the Alliance, the study is still in a conceptual stage, outlining the main risks to interactive relationships within the network. These include managing subgroups, the collision between intra-organizational and network decision-making mechanisms, especially when there are institutional and cultural differences among participants, the balance between institutional and personal connections, and, broadly speaking, between formal and informal relationships within networks. A research intention is outlined based on the specifics of the case study.

11.1 Introduction

International university networks are widely used for implementing both large and small projects that require joint participation from partners in multiple countries. Typically, they are externally funded, and in this sense, the funding donor defines the conditions under which the partnership arises and develops. These types of networks

M. Bogdanova (✉) · E. Parashkevova-Velikova · Y. Kuznetsov
Tsenov Academy of Economics, 5250 Em. Chakarov 2, Svishtov, Bulgaria
e-mail: m.bogdanova@uni-svishtov.bg

E. Parashkevova-Velikova
e-mail: e.parashkevova@uni-svishtov.bg

Y. Kuznetsov
e-mail: y.kuznetsov@uni-svishtov.bg

have clearly defined boundaries and a framework for collaboration since the participating universities are known. However, managing the network raises numerous questions for which the literature has not yet provided clear answers. These questions are related to sustainable relationship management and the transformation of the network from project operational to strategic.

This study aims to investigate the conditions for transforming project partnerships into a strategic network and to identify the key challenges facing individual participants. Additionally, it intends to present research intentions for analyzing networks and their transformation.

Relevant literature has been explored, and qualitative methods have been applied for empirical research through observation, interpretation of documents, and intensive self-reflection [1], with the authors' participation in the process of creating and growing a university network. A partnership network for non-spatial collaboration [2] between universities [3] has been examined, based on cognitive proximity among actively collaborating universities.

The study of network connections was conducted through social network analysis. Patterns of interaction are already emerging—subgroups (clusters) that are visualized through GEPHI. The authors have developed a research program for future examination tailored to the observed specificities in the network's development.

11.2 Literature Review

Interorganizational networks emerge and exist to address specific challenges that no single organization could handle alone [4, 5]. In the scientific literature, the development of networks is assessed from two perspectives [6, 7]: (1) The resource-based perspective, which emphasizes the importance of network participants who possess key resources and share them within the network [8]; (2) The theory of social capital [9], which suggests that collective actions in society depend on the relationships within the community.

Based on these perspectives, two approaches to network formation are observed. In the first approach, the initiative stems from an active entity that actively recruits potential members to join the consortium [4]. This approach is intentional, with the process being designed by the initiator who has identified an opportunity but requires more resources to realize it.

In the second approach, collaborative processes arise almost spontaneously, triggered by changes in the environment and a shared interest and similar perspectives among potential members. These types of networks are also known as multi-actor networks [10]. Most often, they are characterized by a single overarching contractual agreement, shared management, and the pursuit of a common objective requiring each partner to interact within the alliance [10]. These two approaches, due to their nature, imply differences in partner network management strategies.

According to Baraldi et al., the two main types of networks—intentionally created networks and emergent networks—are characterized by different management approaches [11]. In intentionally created networks, management can be exercised either by a single participant (the initiator of the network who holds authority) or by establishing a new organizational structure tasked with managing the entire network [12].

Meanwhile, in emergent networks, management is exercised by all participants, as each of them seeks to manage their own position within the network [13], depending on their individual goals, strategies, and challenges [14].

Different are also the mechanisms of management. While in intentionally created networks, the participants do not know each other, do not have a common history of joint projects, formal mechanisms for regulating relationships [15] can be applied, most often by the central participant or by a separate group of participants who are in immediate proximity to him. Formal mechanisms may include role allocation, establishment of management committees empowered to make decisions, development and adoption of contracts, rules for resource allocation, reporting and control, introduction of deadlines, centralized determination of participants in project initiatives, etc. Integrating them, for example, between roles, resources, and deadlines, may require the parties involved to adapt management processes to the network's requirements, i.e., to demonstrate flexibility in favor of the network's common goals.

Spontaneously emerging networks are characterized by informal mechanisms (at least initially), based on mutual trust, responsibility, and reciprocity, established in prior contacts [16].

According to Vivien Lowndes and Chris Skelcher, the key challenge for partnerships lies in the necessity to apply various approaches and methods of management within the network, which may interact with each other. In some cases, this will generate competition, while in others, cooperation.

As a rule, the objectives of inter-university partnerships in project implementation are externally defined by public entities, which are the main source of funding for participating academic and private institutions and organizations [11]. Therefore, outcome indicators are determined by the governing body that provides the funding. However, sometimes these indicators are not sufficient for tracking sustainable network development.

Among the notable strategies for managing inter-university partnership networks are: exploring content, connecting, arranging, and process agreements [5].

Mechanisms for management frame inter-organizational relationships for when and how individual participants interact with each other; for example, the presence of contracts and routine procedures reduces risk. Encouraging free collaboration leads to the formation of dyads, triads, and other subgroups. They will be further incentivized if bilateral, trilateral, and other agreements are applied within the network. This, in turn, can lead to new ideas and diversify the sources of funding for initiatives within the network. However, it is also possible to become a risk under certain conditions.

The set of management tools applied by network participants also depends on the internal organizational mechanisms within individual universities. Private universities place much greater emphasis on efficiency and competitiveness, so return on

investment is highly sought after there. State-funded universities, on the other hand, tend to lean toward bureaucratic approaches, and hierarchy continues to be a leading mechanism or at least applied to some extent in decision-making.

Furthermore, there is the issue of the clash between the two approaches—centralized and decentralized. Typically, partnerships with external counterparts (other universities or industries) result from faculty members' professional connections and are owed to their specialized skills. At the same time, the official counterpart in a partnership network is the university itself, meaning other representatives of the university who are part of the management board [18].

This poses the risk of intersecting and reducing motivation for external engagement due to conflicts of interest between academics and administrators. Furthermore, changes in priorities or management views regarding the university's development or the departure of key academic staff may freeze or even eliminate their participation in the network. This is defined as "organized anarchy" [19, 20] and represents an extreme understanding of academic autonomy.

The bureaucratic style of management is difficult to transfer into networks that require broad coordination among diverse partners. This introduces additional diversity in management practices within the consortium.

11.3 Approaches to Network Analysis

Analyzing interaction approaches and the degree of collaboration between network actors is a complex process that requires an assessment of the intensity of participation and the exchange of resources. Effective collaboration can be a key factor in the success of projects, especially in complex or interdisciplinary fields.

Network interaction assessment methods can also be used in assessing the intra-institutional patterns of organizational governance and communication observed within the individual participating organizations.

Measuring the level of collaboration is a multidimensional process that requires combining quantitative and qualitative methods for the most accurate assessment. Success in collaboration often depends on the culture of organizations, relationships between teams, working patterns practiced throughout the project cycle, and effective project management.

The following basic evaluation methods and metrics can be considered when analyzing communications:

- Measuring and evaluating communication and interaction. Common platforms used for communication and project management and ways of integrating and using these platforms are evaluated. For example, integrating tools like Slack, Microsoft Teams, or Asana for collaborative work can improve coordination and efficiency.

- Assessment of volume and manner of sharing resources and knowledge. Investing in shared resources and infrastructure can show commitment to the project and mutual trust, which is a good foundation for partnership and network sustainability
- Evaluation of interaction in carrying out joint activities. The publication of common research results can be an indicator of successful collaboration and integration of knowledge with long-term effects.
- Evaluation of goal setting and performance reporting. Measuring project success against predetermined indicators and goals can reflect the level of effectiveness of collaborative efforts.
- Evaluation of data collection methods and analysis tools used. The use of social network analysis (SNA) can identify the structure of interactions in the project [21].

Social network analysis (SNA) is a specialized approach to network analysis that focuses on studying social structures through the use of networks and graphs. It differs from “normal” network analysis (which can refer to a wider range of network structures, including telecommunications networks, the Internet, transport networks, etc.) in several key respects, focusing specifically on the interactions and relationships between people or groups.

SNA is distinguished by its focus on social interactions and the importance of network structure for understanding social processes. It provides specific methods and algorithms that are used to analyze the complex and dynamic social structures that are formed in the real and virtual world.

Social network analysis uses several groups of metrics to measure and evaluate different network parameters. The indicators for centrality, betweenness, and closeness are most often applied, as well as for structural holes and network density. They analyze network connections to identify nodes that act as brokers between otherwise unconnected groups, offering opportunities for information control and innovation.

There is relatively little research on dynamic social networks (dynamic network analysis), the use of agent-based simulation models (ABM), and the variable structure of models (temporal network models). The latter group analyzes how changes in the structure of ties over time affect information dissemination, community formation, and social dynamics.

11.4 Case Description—Alliance of EU Universities

This article examines a specific type of network structure—an Alliance of European Universities (Alliance). The European alliances are created on the basis of a leading initiative of the European strategy for the development of universities, which aims to create 60 alliances uniting over 500 institutions by the middle of 2024. Due to the need to respect the confidentiality of the partnership, the study will not comment on specific universities.

Each EU Alliance is a kind of project-based network that develops and implements an integrated long-term joint development strategy, i.e., the idea is to make the network sustainable in the long term.

The creation of a joint network management structure within the Alliance is a prerequisite for its funding, as it is a key factor in ensuring shared governance and sustainable development. Every single Alliance of European Universities is based on the principle of sharing, not only of management but also of resources—human, informational, financial, administrative, infrastructural, intellectual, etc.

As a result, the network should provide access to a wide range of stakeholders to high-quality and inclusive education, research and innovation, and mobility on European inter-university campuses. The fundamental aspects include the application of interdisciplinary and cross-sector approaches, personalized support for students and researchers, the development of various forms of education including lifelong learning, the provision of microcredits, internships, and career development opportunities. The networks are based on the principles of publicity and transparency, not only in terms of their governance and funding but also in terms of the need to ensure access, namely: open science, open education, and open data practices.

11.5 Research Methodology

The research followed a qualitative and explorative approach, adopting the case study methodology [22]. Furthermore, the lack of sufficient studies that examine the transition from operational to strategic interaction in research networks highlights the need for case studies to gain in-depth knowledge on this topic.

Qualitative methods of empirical research were applied through observation, interpretation of documents, and intensive self-reflection [1] with the participation of the authors in the process of creating and growing a university network. The study of network connections through social network analysis is still in its early stages. Patterns of interaction are already emerging—subgroups (clusters) for which data are being generated for future research. The SNA is made by GEPHI. In perspective, the authors' research program includes the application of Python and R software tools with which to study the network topology in the distinct stages of alliance development [23]. Particularly useful would be the stage in which communication with the various groups of stakeholders should be deployed, especially related industries that expand sectoral and functional cooperation in the network. In this way, network dynamics [24] will be analyzed, which will give a new perspective to research in the subject area.

11.6 Preliminary Analysis Results and Discussion

In the present study, the network topology in the first two stages of alliance formation is presented. The initial network is composed of 7 universities from 7 countries. During this first stage, when the project proposal for funding itself was prepared, dozens of virtual and 7 common physical meetings were held—in each country of the consortium. Bilateral contracts were signed and individual mobilities were carried out under them, which started the implementation of informal relations. The need to expand contacts even before the official funding led to a second stage of the network, in which 4 new universities from 4 new countries were joined.

The increase in the number of partners also increased the potential undirected contacts in the network from 21 during the first stage to 55 in the second stage. In other words, the network density ($n \times (n - 1)/2$) increased by more than 2.5 times. However, the network of universities, i.e., organizations, is represented by people who work in those organizations. The participation of an average of three people from each university multiplies the possible number of contacts. Each work package development team exchanged information with the others and with the central unit, which summarizes and consolidates the information, causing communication to become sporadic and clustered depending on the nature of the tasks.

The measurement of relationships (connections between universities) in such a type of network should be considered from at least two perspectives. On the one hand, these are the official relationships at the university management level. They are certified by signing memoranda—common to the alliance and between individual universities.

But at the next level, the network begins to thicken not from official documents, but from contacts between individual researchers. It is here that one of the essential risks for network management can be pointed out. Common interest, of course, moves the network forward through the project cycle, but it is the efficiency of interpersonal contacts that determines the speed and success (quality) of some of the results.

This is largely related to the culture of the participating universities, which respects contractual relationships but cannot take place without close contacts on a personal level.

The emergence of the Alliance can be defined as a combination of orchestrated network and spontaneous emergence. In other words, it is created deliberately [11], but its goals are not related to the benefit of the initiator, but of all members—this is the mechanism of this type of university alliance. That is why governance is shared and carried out by all participants, but it is also formalized. This combination of formal and informal contacts is considered a major factor in favor of the sustainability of the alliance.

The nature of European alliances is such that the key performance indicators (KPIs) are determined by the managing authority that provided the funding. This pretty much solves the issue of network performance metrics. But it does not answer the question of how, with what tools the results should be achieved. This question is decided by the participants. This is where the next risk in the management of

the Alliance is rooted—how to correctly observe the key indicators, i.e., to satisfy short- and medium-term program requirements while avoiding conflict with long-term strategies.

The question of the cohesion of the organizational strategies within the Alliance was not discussed in depth at the beginning—during the formation of the Alliance, but its neglect can have serious consequences at subsequent stages.

This was particularly evident during the second stage of the alliance's development, when four more universities joined it.

It was here that differences in organizational management strategies came to the fore. Some cultural and institutional differences made it necessary during the negotiations to adapt internal organizational processes in some universities to the requirements of the network. But with the help of accepted and established cultural norms of behavior, disagreements were contained.

When associated organizations are included in the network during the third stage (on average 4–5 for each university), the network is about to grow even more. Investigation of network metrics such as centrality, modularity, and density are forthcoming at that stage, as well as a more in-depth study of structural holes and communities (clusters) and how they reveal important network characteristics.

The alliance of universities in the case study is conceived as a polycentric terrain [17] for generating innovation. But the relationships between the participants are not equally intense. Therefore, a study of the strong and weak ties in the network is to be done [25]. This is important because the already formed strong ties form inner social circles and even clusters. At the same time, the presence of weak ties opens up access to new possibilities as participants seek to access more knowledge and information. In this sense, the analysis of relationships will form a forecast for the development of the network and, respectively, will suggest strategies and management approaches.

In view of the possibility of directly observing the development of the network from its inception, the authors are also interested in the application of dynamic network analysis (DNA) and tools such as ORA (Organizational Risk Analyzer) to analyze networks that change over time. Python and R software tools could be implemented due to their flexibility and extensive library collection for data manipulation and network analysis. Such flexible tools are particularly useful in distributed networks, as they cover not only local partners but also those that are geographically very distant.

A typical social network analysis focusing on centrality indicators (such as the number of inward or outward links) provides a basis for analysis, but is not sufficient to guarantee the achievement of interaction density. In other words, social network analysis does not illuminate all the ways in which two system components without a direct relationship can influence each other [26].

To improve interaction, the network is to be analyzed through additional techniques, such as cognitive maps, gap analysis, historical lessons learned analysis, functional analysis of network participation, etc.

The development of a network management model must consider the answers to several questions: purpose (the why), process (the how), partnership (the who), and

product (the what) [26, 27] and what are the attitudes of the participants in the short and long term.

When the network sets out to build a strategic partnership, the participants must make decisions about what business model the network will follow and what will be the contribution of everyone involved in the decision-making. The strategic nature focuses not on the direct consequences of decisions for individual actors in the short term, but on the consequences at the systemic level—for the entire network. This means overcoming institutional short-sightedness [24] in the name of achieving strategic advantages.

11.7 Conclusion

This article adds knowledge about the conditions and risks for strategic collaboration in university networks, which is a poorly developed niche in the literature.

Due to the early stage of development of the Alliance, the research is still in the conceptual stage, but the main risks to interactive relationships in the network have been outlined, such as the management of dyads, triads, and other subgroups (clusters); the collision between intra-organizational and network decision-making mechanisms, especially when there are institutional and cultural differences between participants; the balance between institutional and personal ties, and in a broad sense—between formal and informal relationships. Last but not least, the strategic orientation depends on the quality of the interactive relationships in the network.

The topic is broad and interdisciplinary, as it binds on the one hand the network theories, respectively, the factors for the emergence and development of networks, and on the other—the approaches and tools for strategic management, but applied in distributed networks for non-spatial collaboration with the aim of cross-enrichment of ideas between participants. This compilation of perspectives calls for deepening research and shaping future research intentions that combine network metrics at each stage of network development and the relationship of network topology statistics to strategic network analysis.

Acknowledgements The paper is part of activities under project № КП-06-H65/10 “Formal and Informal Innovation Networks,” funded by the Research Fund of the Ministry of Education and Science in Bulgaria.

References

1. Janesick, V.: The dance of qualitative research design: metaphor, methodolatry, and meaning, in handbook of qualitative research, N. K. a. L. Y. S. (Denzin, Ed., Thousand Oaks, CA, Sage Publications, pp 209–219, (1994)
2. Bernini, C., Galli, F.: Networking and spatial interactions: what contributes most to increasing museums’ attractiveness? *Pap. Reg. Sci.* **102**(6), 1215–1233 (2003)

3. Santos, E.G., et al.: Spatial and non-spatial proximity in university–industry collaboration: mutual reinforcement and decreasing effects. *Reg. Sci. Policy Pract.* **13**(4), 1249–1262 (2021)
4. Doz, Y.L., Olk, P.M., Ring, P.S.: Formation processes of R&D consortia: which path to take? where does it lead. *Strat. Manag. J.* **21**(3), 239–266 (2000)
5. Klijn et al.: The impact of network management strategies on the outcomes in governance networks. *Public Adm.* **88**(4), 1063–1082 (2010)
6. Hagedoorn, J., Link, A., Vonortas, N.: Research partnerships. *Res. Policy.* **29**, 567–586 (2000)
7. Arroyabe, M.F et al.: R&D partnerships: An exploratory approach to the role of structural variables in joint project performance. *Technol. Forecast. Soc. Chang.* **90**, 623–634 (2015)
8. Fritsch, M., Lukas, R.: Who co-operates on R&D? *Res. Policy* **30**, 297–312 (2001)
9. Burt, R.: Structural holes and good ideas” *Am. J. Sociol.* **110**, 349–399 (2004)
10. Partanen, J.: How to build a strategic network: a practitioner-oriented process model for the ICT sector. *Ind. Mark. Manage.* **41**(3), 481–494 (2012)
11. Baraldi et al.: Managing interorganizational interactions for social impact: a study of two antibiotics R&D networks. *J. Bus. Res.* **141**, 264–278 (2022)
12. Kenis, P.: Modes of network governance: structure, management, and effectiveness. *J. Public Adm. Res. Theory.* **18**(2) (2008)
13. Ritter, T., Wilkinson, I. F., Johnston, W. J.: Managing in complex business networks. *Ind. Mark. Manag.* **33**, 175–183 (2004)
14. Ford, D., Mouzas, S.: The theory and practice of business networking. *Ind. Mark. Manage.* **42**(3), 433–442 (2013)
15. Sarjito, A.: Assessing R&D investment policy in the Indonesian defense sector: towards long-term sustainability. *Int. J. Appl. Adv. Multidiscip. Res.* **2**(1), 47–66 (2024)
16. Grøn, A.B et al.: Can leadership improve interorganizational collaboration? field-experimental evidence from a team-based leadership training intervention. *Am. Rev. Public Adm.* (2024)
17. Lowndes and Skelcher: The dynamics of multi-organizational partnerships: an analysis of changing modes of governance. *Public Administration* **76**(Summer), 313–333 (1998)
18. Broström, A., Feldmann, A., Kaulio, M.: Structured relations between higher education institutions and external organisations: opportunity or bureaucratization. *High Educ.* **78**, 575–591 (2019)
19. Cohen et al.: A garbage can model of organizational choice. *Adm. Sci. Q.* 1–25 (1972)
20. Blaschke, St., Frost, J., Hattke F.: Towards a micro foundation of leadership, governance, and management in universities. *High. Educ.* **68**(5), 711–732 (2014)
21. Majeed, S., Uzair, M., Qamar, Q., Farooq, A.: Social network analysis visualization tools: a comparative review. In: *IEEE 23rd international multi topic conference* (2020)
22. Yin, R.K.: *Case study research and applications: Design and methods.* Sage Publications, Thousand Oaks, CA (2017)
23. Apostolato, I.A.: An overview of software applications for social network analysis. *Int. Rev. Soc. Res.* **3**(3) (2013)
24. Forliano, C., De Bernardi, P., Bertello, A., Temperini, V.: Innovating business processes in public administrations: towards a systemic approach. *Bus. Process. Manag. J.* **26**(5), 1203–1224 (2020)
25. Granovetter, M.: The impact of social structure on economic outcomes. *J. Econ. Perspect.* **19**(1), 33–50 (2005)
26. Hassannezhad, M., Gogarty, M., Connor, C.H.O., Cox, J., Meier, P.S., Purshouse, R.C.: A cybernetic participatory approach for policy system of systems mapping: case study of inclusive economies, futures, vol 152 (2023)
27. Gray, S et al.: Purpose, processes, partnerships, and products. *Four Ecol. Appl.* **28**(1), 46–61 (2018)

Chapter 12

Comparative Analysis of Natural Language Query Responses on BPMN Model Serializations: RDF Graphs Versus BPMN XML



Damaris Naomi Dolha and Robert Andrei Buchmann

Abstract Leveraging the power of symbolic knowledge representation, this study compares outcomes of natural language interaction with BPMN model serializations, examining semantic graphs derived from RDF export of BPMN provided by Bee-Up, in contrast to the conventional BPMN XML export from the process modeler of SAP Signavio. By prompt engineering, we investigate the proficiency of certain GPT services of OpenAI in navigating the semantic intricacies of RDF and the structural hierarchy of XML, ultimately illuminating implications for knowledge retrieval. The findings delve into the complexities of querying BPMN representations using natural language, revealing the transformative capabilities of RDF, but also the value of BPMN employed as a schema for procedural knowledge graphs—i.e., shifting away from their traditional role as diagrams or automation configurators. As per the experimental results, the RDF export showcases superior richness for natural language queries as the graph-like structure of visual diagrams is closer to semantic networks than to XML tag structures, carrying implications for Business Process Management.

12.1 Introduction

In the practice of business process modeling, the choice between serialization formats significantly influences how organizations perceive, analyze, process, and implement their business processes. XML and RDF, two prominent content structuring

D. N. Dolha (✉) · R. A. Buchmann
Faculty of Economics and Business Administration, Babeş-Bolyai University, Cluj-Napoca, Romania
e-mail: damaris.dolha@econ.ubbcluj.ro

R. A. Buchmann
e-mail: robert.buchmann@econ.ubbcluj.ro

approaches, offer distinct ways of encoding BPMN diagrams, each carrying advantages and implications for process management and automation. While XML has long been the standard for the interchange of BPMN models, the emergence of RDF export capabilities, particularly by Bee-Up [1] and other experimental tools [2], introduces new possibilities for semantic process enrichment. This study explores how Large Language Models (LLMs)—for now specifically those from OpenAI (to be expanded in future work to others)—benefit from this richness compared to the traditional formats. Although multimodal image interpretation capabilities have also been recently introduced, the typical APIs of Business Process Management Systems (BPMS) and automation platforms will deliver serializations (rather than images) to client apps or partnering services, therefore, we focus this stage of our research on such serializations. Moreover, some process knowledge is not found on the visual canvas—modeling tools often allow to store data attributes in model elements (e.g., costs, execution times of process tasks) or non-visual links (e.g., subprocesses, individual RACI roles linked to specific tasks).

The current study explores the interplay between natural language queries and BPMN models in RDF representation compared to conventional XML export, in a setting showcasing human–LLM interaction over serialized business process diagrams. We test various prompts equivalent to process querying different patterns found in BPMN diagrams, although acknowledging some comparability limitations: while BPMN XML is a universally adopted standard, RDF is only standardized as a knowledge structuring method and lacks standardization as a process description vocabulary—these vocabularies are emerging in experimental implementations such as [2, 3] and public tools such as Bee-Up. Nevertheless, this investigation offers valuable insights into how diagram serialization can enhance diagrammatic model interpretability for LLMs. Although our current results are preliminary and limited in experimentation diversity, we plan to expand our study to include a range of LLM services and multimodal capabilities. The current exploration underscores that the choice between RDF and XML export extends beyond technical preference; it signifies a shift in conceptualizing BPMN models as semantic networks rather than hierarchical XML tags decomposing a control flow. Through this analysis, we suggest a transformative potential of RDF diagram exports within the BPMN ecosystem, which has been advocated for a while in domain-specific, semantically supported Business Process Management (BPM) [3, 4].

This paper is structured as follows: Sect. 12.2 outlines challenges in BPM and discusses the benefits of integrating LLMs. In Sect. 12.3, we provide a comprehensive review of past solutions proposed for process queries. Immediately afterward, Sect. 12.4 details the tools used in our study, ensuring transparent documentation of our methodology for reproducibility and meaningful comparisons with future research. The outcomes of our experiments, presenting empirical evidence regarding the comparative performance of RDF and XML exports in responding to natural language queries, are detailed in Sect. 12.5. Finally, Sect. 12.6 synthesizes findings, discusses limitations, and provides an outlook for future exploration within the field.

12.2 Problem Statement and Background

The potential integration of LLMs into BPM presents both opportunities and challenges [5]. During the first phase of the BPM lifecycle, namely *process identification phase*, LLMs can aid in extracting key information from unstructured procedural documentation. Additionally, they may contribute to process prioritization by semantically matching with strategic statements. During the *process discovery*, Generative AI can enhance traditional process mining methods by handling XML formats that currently capture process logs and serializations. Process querying, traditionally XML-based, could benefit from the diagram understanding capabilities of multimodal models like GPT-4 [6]. In the *analysis phase*, LLMs can assist in issue discovery, by summarizing large volumes of text from various platforms where complaints may arise and identifying patterns. For *process redesign*, LLMs may employ mechanisms akin to those used in code generation. In the *implementation phase*, LLMs provide textual explanations of BPMN models, aid in process execution and enhance interaction between users and models. Processes can serve as orchestrators of Artificial Intelligence and chatbots. For *process monitoring*, visualization support or interpretation through dashboards are conceivable.

To navigate this diverse landscape of use cases, it is crucial to explore the extent to which LLM can offer alternatives to traditional process queries, since traditionally, process queries have relied on formal languages and visual grammars [7] or they have repurposed graph database queries [2, 3].

12.3 Related Works

An article that explores the potential of LLMs in accomplishing text-related BPM tasks is [8], where, through tailored prompts, the study evaluates the performance of an LLM in the context of mining declarative process models from natural language descriptions. There is interest in leveraging the structured nature of business process modeling languages for knowledge graph integration. Notably, a recent paper [2] introduced a converter translating BPMN XML format to Neo4J labeled property graphs, ensuring conceptual alignment with BPMN 2.0 standards. Similarly, another work [3] presents BPMN2KG, a tool converting BPMN 2.0 process models into knowledge graphs, facilitating process-centric data integration and SPARQL queries. The authors demonstrate the utility of the tool through a real-world industrial use case in quality management. The evolving landscape of BPMS underscores the need for semantic agility, especially as knowledge-driven applications become more context-aware. Paper [4] addresses this need, integrating process knowledge with contextual data. The proficiency of ChatGPT in generating and interpreting conceptual models

like ER diagrams and UML Class diagrams is demonstrated in [9], while paper [10] discusses leveraging pre-trained language models for BPM research, noting that the application of GPT-3 in BPM often requires fine-tuning and significant training data. Integrating GPT-3.5 Turbo into Lean Six Sigma 4.0 methodologies offers transformative capabilities for real-time data-driven strategies [11]. Study [12] explores integrating GPT with Robotic Process Automation (RPA), noting potential security and legal risks. Researchers have proposed enhancing LLMs with knowledge graphs to overcome these challenges, as discussed in article [13], complemented by our study on applying knowledge graph treatment to BPMN models.

12.4 Experimental Setup

In order to conduct our investigation, we harnessed several tools. Firstly, Bee-Up 1.7 is the core tool within the OMiLAB Digital Innovation environment [14]. The RDF export functionality in Bee-Up was based on patterns elucidated in the paper referenced at [15]. We employed Ontotext GraphDB 10.5 [16], a scalable and robust graph database with RDF and SPARQL support. This allowed us to apply various types of SPARQL queries, including “magic predicates” to redirect inquiries to gpt-3.5-turbo model [17]—on the RDF serialization exported from Bee-Up 1.7, facilitating in-depth analysis and exploration of the semantic relationships within the BPMN models. Secondly, SAP Signavio Process Transformation Suite, Academic Edition 17.13.0 [18] was part of our experimental setup for exporting BPMN into standard XML serialization, subsequently parsed and analyzed by ChatGPT 3.5 from OpenAI. While using a local solution with an OpenAI API key or a private chatbot for secure document handling can be important, ChatGPT also ensures safety and privacy through several robust measures. OpenAI employs stringent security practices, including SOC 2 Type 2 compliance, verified through rigorous audits to ensure high data protection standards [19]. Moreover, data encryption, both at rest and in transit using AES-256 and TLS 1.2+, further secures data handling. Additionally, the OpenAI Privacy Request Portal [20] allows users to manage their data and privacy settings efficiently. Users can opt out of sharing their data for training purposes in the ChatGPT settings, before starting the conversation.

The BPMN exemplar used for the discussed prompting cases is shown in Fig. 12.1.

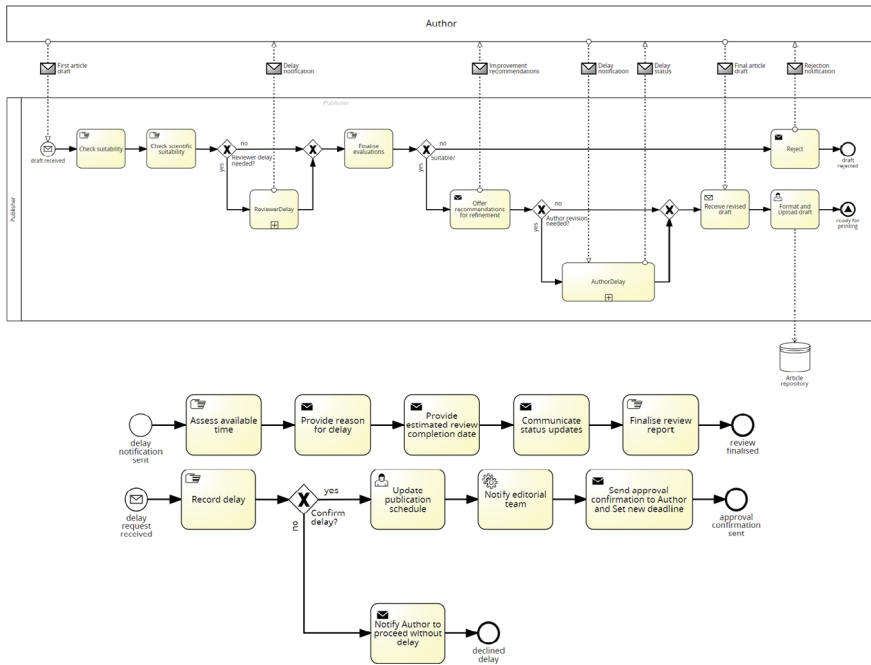


Fig. 12.1 Main BPMN process diagram (top) and the two embedded subprocesses

12.5 Commented Prompting Strategies

The responses generated by the OpenAI model for each prompt will be compared for process descriptions exported from Bee-Up 1.7 in RDF format (noted as CASE I) and from SAP Signavio Process Transformation Suite, Academic Edition 17.13.0 in XML format (represented by CASE II). It is crucial to note that the emphasis of this study lies not on uncovering limitations inherent to any particular tool, but rather on the broader exploration of the comparative analysis between the aforementioned representation formats. This approach ensures a holistic examination of the effectiveness and implications of different representation formats on the interpretation of natural language queries within the BPMN context. The repertoire of prompt types encompasses a range of methodologies aimed at discerning differences and nuances in the model output, based on the examined BPMN process and subprocesses exported from the aforementioned tools.

Scenario 1. Open-ended prompt. “Can you summarize the overall workflow of the BPMN process?”

Case I (GraphDB queries redirected together with RDF export to OpenAI gpt-3.5-turbo)	Case II (ChatGPT 3.5 on XML export)
<p>The BPMN process generally involves several tasks or activities that are interconnected and flow in a sequence to achieve a specific goal. In the provided RDF data example, we have tasks such as “Check suitability,” “Check scientific suitability,” “Finalize evaluations,” “Reject,” “Offer recommendations for refinement,” “Receive revised draft,” and “Format and upload draft.” These tasks are part of the BPMN process flow and typically require input, processing, and output to progress from one task to another. The workflow of the BPMN process might involve the following steps based on the tasks mentioned: 1. Check suitability and check scientific suitability tasks are performed to assess the suitability and scientific aspects of a particular process or project. 2. Based on the evaluations and assessments, a decision is made to either finalize evaluations or reject the proposal. 3. If the evaluations need refinement, recommendations are offered for further improvement. 4. The revised draft is received and processed through the “Receive revised draft” task. 5. Lastly, the formatted and reviewed draft is uploaded as the final output. This workflow represents a sequential flow of tasks within the BPMN process, where each task contributes to the overall goal or objective of the process</p>	<p>The overall workflow of the BPMN process can be summarized as follows: 1. Article Submission: The process begins with the Author submitting a draft article for publication. 2. Review and Feedback: Upon submission, the Publisher initiates the review process. The Publisher and other reviewers provide feedback on the submitted article. If revisions are needed, the Author is notified and required to make the necessary changes. 3. Revision and Resubmission: The Author revises the article based on the feedback received. Once revisions are completed, the Author resubmits the revised article for further review. 4. Review Process Management: The Publisher manages the review process, ensuring timely feedback and resolution of any issues that arise. Notifications regarding delays or required actions are communicated between the Author and Publisher as needed. 5. Final Approval and Publication: Once the article meets the required standards and receives approval, the Publisher coordinates its publication. The finalized article is published, completing the workflow. Throughout the process, there is a continuous exchange of information and feedback between the Author and Publisher, ensuring the quality and timely publication of articles</p>

While both cases capture the essence of the provided BPMN process, CASE I offers a more granular breakdown of the process steps—with each task explicitly named and interconnected—whereas the approach within CASE II is in a rather higher-level approach, with a sole focus on the key stages of the workflow.

Scenario 2. Clarification prompt. “What is the last task within the current BPMN process?”

Case I	Case II
<p>Based on the provided RDF triples, the last task within the current BPMN process is labeled as “Format and Upload draft” with the URI: https://www.omilab.org/#Task_BPMN-16315-Format_and_Upload_draft</p>	<p>The last task within the current BPMN process is the “Update publication schedule” user task in the “AuthorDelay” subprocess</p>

In CASE I, the last task identified within the current BPMN process is “Format and Upload draft”, which clearly pinpoints the final action within the main process flow. However, in CASE II, the focus shifts to a task within the “AuthorDelay” subprocess, namely “Update publication schedule”. While this task is part of the overall BPMN process, it pertains to a subprocess rather than the main process itself.

“What is the type of task ‘Finalise evaluations’?”

Case I	Case II
<p>The type of task ‘Finalise evaluations’ is “Manual”</p>	<p>The type of task “Finalise evaluations” in the BPMN (Business Process Model and Notation) diagram is a manual task</p>

Both descriptions aptly detect the task type (explicitly available in both serializations) of “Finalise evaluations”, albeit with slight differences in terminology. Neither description performs noticeably better than the other in terms of accuracy or clarity.

“Could you clarify how author delays are managed in the BPMN process and who is responsible for each activity?”

Case I	Case II
<p>In the provided RDF data, the responsibilities for managing author delays in the BPMN process are outlined for each activity. Here is a breakdown of the tasks and their responsible performer: 1. Task: Record delay—Responsible Performer: Mary J.—Activity: Recording the delay 2. Task: Update publication schedule—Responsible Performer: Mary J.—Activity: Updating the publication schedule based on the delay recorded 3. Task: Notify editorial team—Responsible Performer: Mary J.—Activity: Informing the editorial team about the delay and any schedule changes 4. Task: Send approval confirmation to Author and Set new deadline—Responsible Performer: Mary J.—Activity: Sending confirmation of approval to the author and setting a new deadline for the revised schedule 5. Task: Notify Author to proceed without delay—Responsible Performer: Mary J.—Activity: Communicating with the author to ensure they proceed promptly without further delays. In this BPMN process, Mary J. is assigned as the responsible performer for all the activities related to managing author delays. This indicates that she is responsible for coordinating and executing the necessary steps to address and mitigate any delays that may occur in the process</p>	<p>In the provided BPMN (Business Process Model and Notation) process, which seems to represent the editorial workflow for publishing academic articles, author delays are managed through a subprocess called “AuthorDelay.” Let’s break down the process and responsibilities associated with managing author delays: 1. Initial Notification of Delay: When a delay notification is received from the author (represented as the start event “delay request received”), the subprocess “AuthorDelay” is triggered. 2. Record and Confirm Delay: The delay is recorded by a manual task named “Record delay.” After recording the delay, a decision is made to confirm whether the delay is valid or not. This is depicted by the exclusive gateway “Confirm delay?” which diverges the flow into two possible paths based on the decision outcome. 3. Handling Confirmed Delay: If the delay is confirmed to be valid, the process proceeds to update the publication schedule by assigning a new deadline. This is managed by the user task “Update publication schedule.” After updating the schedule, the editorial team is notified about the delay through the service task “Notify editorial team.” Subsequently, approval confirmation is sent to the author along with the new deadline using the send task “Send approval confirmation to Author and Set new deadline.” Finally, an end event “approval confirmation sent” marks the completion of this path. 4. Handling Rejected Delay: If the delay is deemed invalid, the process moves to notify the author to proceed without delay. This is done through the send task “Notify Author to proceed without delay.” An end event “declined delay” indicates the completion of this path. 5. Responsibilities: The responsibilities for managing author delays primarily lie within the subprocess “AuthorDelay.” The tasks involved include recording the delay, confirming its validity, updating the publication schedule, notifying the editorial team, and communicating with the author. The editorial team and the publisher might be involved in decision-making regarding the validity of the delay and in communicating with the author accordingly. This BPMN process provides a structured approach for managing author delays in the publication workflow, ensuring timely communication and decision-making to minimize disruptions in the publishing schedule</p>

Although both CASE I and CASE II aim to correctly handle author delays within the provided BPMN process, CASE I stands out for explicitly mentioning the performers responsible for handling author delays, a detail that CASE II lacks. Both tools allow the assignment of responsibilities (e.g., RACI—responsible, accountable, consulted, informed); however, in the case of RDF, these are explicit semantic links to elements of an organigram, whereas in SAP Signavio, they are attributes that do not make it into the standard XML export, being used only by in-tool features or exported separately as Excel. This is a major difference between the open-ended semantics enabled by annotating BPMN elements with RDF resources (as available in Bee-Up) and the standard-grounded closed semantics of the XML export—probably not intended to be used in an open semantic space such as being explored by LLM products.

Scenario 3. Evaluation Prompt. “How effective do you think the communication between the Author and Publisher is in the BPMN process and what improvements could be made?”

Case I	Case II
answers generated by Ontotext GraphDB 10.5 using OpenAI gpt-3.5-turbo model on RDF formatted data	answers provided by ChatGPT 3.5 on XML formatted data
<p>The communication between the Author and Publisher in the BPMN process as depicted by the provided RDF triples seems to be structured through specific tasks with labels like “Check suitability”, “Finalize evaluations”, “Reject”, “Offer recommendations for refinement”, “Receive revised draft”, and “Format and Upload draft”. These tasks suggest a linear flow of activities from initial evaluation to final publishing stages. Improvements that could be made to enhance communication between the Author and Publisher in the BPMN process based on the provided triples include: 1. Clarity on Task Dependencies: There should be explicit relationships or dependencies defined between tasks to understand the sequence of activities better. This could be achieved by adding properties like “precedes” or “follows” to show the order of execution. 2. Feedback Mechanism: Introducing tasks related to feedback provision or review cycles can make the communication process more interactive. For example, after the Author uploads a draft, there could be a “Reviewer feedback” task that feeds back into the revision process. 3. Notification Mechanism: Implementing a task that notifies relevant parties (Author, Publisher) about the completion of specific tasks can streamline the communication process and avoid delays. 4. Collaborative Task Assignment: Enabling features that allow the assignment of tasks to specific individuals or roles can improve accountability and streamline the workflow. 5. Integration with Communication Tools: Integrating communication tools within the BPMN process, such as notifications via email or project management platforms, can facilitate quicker responses and updates between the Author and Publisher. Overall, incorporating these improvements could enhance the effectiveness of communication between the Author and Publisher in the BPMN process, ensuring smoother collaboration and efficient progression from manuscript submission to publication</p>	<p>Based on the BPMN process, the communication between the Author and Publisher seems fairly effective. The process includes clear steps for the Author to submit the manuscript and for the Publisher to review and accept or reject it. However, there are a few potential improvements that could be made to enhance communication and efficiency: Feedback Loop: Implement a feedback loop where the Publisher provides detailed feedback to the Author in case of rejection. This can help the Author understand reasons for rejection and improve the manuscript for future submissions. Real-time Notifications: Introduce real-time notifications to keep both parties informed about the status of the manuscript. This could include automated emails or notifications within the BPMN system to inform the Author when the manuscript is received, under review, or accepted/rejected. Collaborative Review: Allow for a collaborative review process where the Author and Publisher can discuss changes or improvements to the manuscript before a final decision is made. This can lead to better quality submissions and reduce the likelihood of misunderstandings. Transparent Workflow: Ensure transparency in the workflow by clearly defining each step of the process and providing access to status updates for both the Author and Publisher. This can help manage expectations and prevent delays or misunderstandings</p> <p>Overall, by implementing these improvements, the communication between the Author and Publisher can become more streamlined, transparent, and collaborative, leading to a more efficient manuscript submission and review process</p>

CASE I shows a more detailed comprehension of the graph elements, also pointing to the design detail that the graph treats sequence flow arrows as nodes, with no direct precedes/follows edge present in the graph. It appears to be more stimulated by the graph structure in generating extrapolations and recommendations. CASE II shows a good understanding of the importance of communication and suggests

improvements that can enhance the efficiency of the BPMN process, but it lacks the depth of specificity and granularity found in the first case.

Scenario 4. Prediction prompt. “Based on the BPMN process description, what do you anticipate would be the most common reasons for article rejection?”

Case I	Case II
<p>Based on the provided BPMN process description, the most common reasons for article rejection could include: 1. Lack of suitability: If the article does not meet the criteria for scientific suitability or general suitability, it may be rejected during the “Check suitability” and “Check scientific suitability” tasks. 2. Insufficient quality or content: Articles that do not meet certain quality standards or lack in-depth scientific content may be rejected during the evaluation process, leading to the “Reject” task. 3. Failure to address reviewer feedback: If authors do not adequately address feedback or make necessary revisions after receiving evaluations, their article may be rejected upon receiving the revised draft during the “Receive revised draft” task. 4. Formatting issues: Articles that are not properly formatted or uploaded as per the guidelines may also face rejection during the final stages of the process in the “Format and Upload draft” task. Overall, it seems that the rejection of articles in this BPMN process could be influenced by factors related to suitability, quality, responsiveness to feedback, and adherence to formatting requirements</p>	<p>Based on the BPMN process description provided, the most common reasons for article rejection could include: Incomplete Submission: If the article submission does not include all required components or information as specified in the submission guidelines, it may be rejected. Violation of Guidelines: Articles that fail to adhere to the journal’s formatting guidelines, citation style, or ethical standards may face rejection. Poor Quality: Articles with substandard writing, lack of clarity, or insufficient depth of research are likely to be rejected. Lack of Originality: Articles that do not present new or innovative findings, or that merely replicate existing research without significant contributions, may be rejected. Inappropriate Scope: Articles that fall outside the scope of the journal, or that do not address topics of relevance to its audience, may face rejection. Methodological Flaws: Articles with flawed research methodologies, inadequate data analysis, or unreliable results may be rejected due to concerns about the validity of findings. Language Issues: Articles with poor grammar, spelling errors, or language that is not appropriate for publication may be rejected or returned for revision. Lack of Significance: Articles that do not offer significant contributions to the field or fail to advance current understanding may be deemed unsuitable for publication. These reasons reflect common criteria used by journals to assess the quality and suitability of submitted articles for publication</p>

In CASE I, specific tasks within the BPMN process, such as “Check suitability”, “Check scientific suitability”, “Reject”, “Receive revised draft” and “Format and Upload draft” are indicated for potential reasons for article rejection. This focused approach provides clear and explicit answers tailored to the provided BPMN process description, making CASE I excellent concerning prediction prompts. On the contrary, CASE II provides a more generalized overview of potential reasons for article rejection, encompassing a broader range of criteria. While these reasons may apply to article rejections in general, they are not explicitly tied to the provided BPMN process description. As a result, CASE II offers a less targeted and more generalized analysis.

Scenario 5. Chain of thought prompt. “Starting from the submission of the article draft by the Author, trace the sequence of events leading to the preparation of the article for printing in the BPMN process.”

Case I	Case II
<p>1. The Author submits the article draft. 2. The submitted draft goes through the “Check suitability” task (Task_BPMN-15569). This task assesses if the draft is suitable for the intended purpose. 3. If the draft passes the suitability check, it moves on to the “Check scientific suitability” task (Task_BPMN-15573), which evaluates the scientific credibility and relevance of the content. 4. After passing the scientific suitability check, the draft proceeds to the “Finalise evaluations” task (Task_BPMN-15587), where the final evaluations and feedback are consolidated. 5. If the draft does not meet the criteria during the evaluation tasks, it is sent to the “Reject” task (Task_BPMN-15627). 6. If revisions are needed, the draft is sent to the “Offer recommendations for refinement” task (Task_BPMN-15635) for suggestions on how to improve it. 7. Once the Author incorporates the revisions and improvements, the updated draft is submitted again and goes to the “Receive revised draft” task (Task_BPMN-15651). 8. After receiving the revised draft, the final step in the process is the “Format and Upload draft” task (Task_BPMN-16315), where the document is prepared for printing and uploaded for publication</p>	<p>Based on the provided BPMN process, here is the sequence of events leading to the preparation of the article for printing: Submission of Article Draft by Author: The process starts with the submission of the article draft by the author. This triggers the process. Editorial Review: After the article draft is submitted, it undergoes an editorial review. During this phase, editors review the draft for quality, clarity, and adherence to guidelines. Review Decision: Following the editorial review, a decision is made regarding the draft. If the draft meets the publication criteria, it moves forward to the next stage. If it requires revisions, it is sent back to the author for revisions and resubmission. Revision by Author: If revisions are needed, the author revises the article accordingly and resubmits it. Final Review: Once the revised draft is submitted, it undergoes a final review. Editorial Approval: If the final review is satisfactory, the article receives editorial approval. This indicates that the article is ready for the next phase of the process. Preparation for Printing: After editorial approval, the article is prepared for printing. This involves formatting, layout design, and any other necessary preparations. Printing: Once the preparation is complete, the article is sent to the printing stage. This is the final stage before the article is published and distributed. Article Published: Finally, the printed article is published and made available to readers. This sequence outlines the steps involved in preparing the article for printing in the BPMN process, starting from the submission of the article draft by the author to its publication</p>

CASE I breaks down the process into distinct tasks, providing a thorough view of each step involved in the preparation of the article for printing. This level of detail allows for a more precise understanding of the workflow and potential points for optimization or improvement. CASE II presents a higher-level overview of the process, focusing on broader stages, such as review, revision, approval and others. Nonetheless, while this provides a clear outline of the overall flow, it may overlook specific tasks.

12.6 Conclusions and Future Work

Our examination reveals that RDF exports, characterized by their directed graph-based structure and labeling of edges and nodes, outperform standard XML exports in addressing natural language queries, even if the process description vocabulary is tool-specific and non-standard in contrast to the XML serialization. Our reliance on LLM services, notably GPT-3 by OpenAI, underscores the imperative of exploring alternative methodologies and platforms to ensure more comprehensive research coverage and diversity of prompts. A challenge arises with ChatGPT 3.5 during the examination of XML-based business processes: it tends to deviate from the focal BPMN process and furnish inaccurate responses after multiple prompts. To recalibrate the alignment of ChatGPT 3.5 with the provided BPMN process and subprocesses, it is required to reintroduce the XML contents as a prompt prior to advancing with the analysis. Our further work regards ongoing investigation on how results compare with the new generations of GPT models (GPT-4 and newer), particularly in multimodal setups where images are also used in combination with semantic links and narrative content.

Additionally, we plan to further employ the Retrieval Augmented Generation Assessment (RAGAs) framework [21] to systematically compare the accuracy and relevance of answers generated by LLM across RDF and XML formatted BPMN content. This framework will help us identify strengths and weaknesses in how each data format influences the performance of the model in handling queries over complex workflows.

References

1. OMILAB NPO, Bee-Up For Education. <https://bee-up.omilab.org/activities/bee-up/>. Accessed 30 Dec 2023
2. Uifălean, Ș., Ghiran, A.M., Buchmann, R.A.: Employing graph databases for business process management and representation. In: Silaghi, G.C., et al. *Advances in Information Systems Development. ISD 2022. Lecture Notes in Information Systems and Organisation*, vol 63. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-32418-5_5
3. Bachhofner, S., Kiesling, E., Revoredo, K., Waibel, P., Polleres, A.: Automated Process Knowledge Graph Construction from BPMN Models. In: Strauss, C., Cuzzocrea, A., Ko-tsis, G., Tjoa, A.M., Khalil, I. (eds) *Database and Expert Systems Applications. DEXA 2022. Lecture Notes in Computer Science*, vol 13426. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-12423-5_3
4. Cinpoeru, M., Ghiran, A.M., Harkai, A., Buchmann, R.A., Karagiannis, D.: Model-Driven Context Configuration in Business Process Management Systems: An Approach Based on Knowledge Graphs. In: Pańkowska, M., Sandkuhl, K. (eds.) *Perspectives in Business Informatics Research. BIR 2019. Lecture Notes in Business Information Processing*, vol 365. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-31143-8_14
5. Vidgof, M., Bachhofner, S., Mendling, J.: Large Language Models for Business Process Management: Opportunities and Challenges. In: Di Francescomarino, C., Burattin, A., Janiesch, C., Sadiq, S. (eds.) *Business Process Management Forum. BPM 2023. Lecture Notes in*

- Business Information Processing, vol 490. Springer, Cham (2023). <https://doi.org/10.48550/arXiv.2304.04309>
6. OpenAI, GPT-4. <https://openai.com/research/gpt-4>. Accessed 17 Oct 2023
 7. Polyvyanyy, A.: Process Querying: Methods, Techniques, and Applications. In: Poly-vyanyy, A. (eds.) Process Querying Methods. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92875-9_18
 8. Grohs, M., Abb, L., Elsayed, N., Rehse, JR.: Large Language Models Can Accomplish Business Process Management Tasks. In: De Weerd, J., Pufahl, L. (eds.) Business Process Management Workshops. BPM 2023. Lecture Notes in Business Information Processing, vol 492. Springer, Cham (2024). <https://doi.org/10.48550/arXiv.2307.09923>
 9. Fill, H., Fettke, P., Köpke, J.: Conceptual modeling and large language models: Impressions from first experiments with ChatGPT. EMISAJ. **18**(3), 1–15 (2023). <https://doi.org/10.18417/emisa.18.3>
 10. Busch, K., Rochlitz, A., Sola, D., Leopold, H.: Just Tell Me: Prompt Engineering in Business Process Management. In: Van der Aa, H., Bork, D., Proper, H.A., Schmidt, R. (eds.) Enterprise, Business-Process and Information Systems Modeling. BPMDS EMMSAD 2023. Lecture Notes in Business Information Processing, vol 479. Springer, Cham (2023). <https://doi.org/10.48550/arXiv.2304.07183>
 11. Shahin, M., Chen, F.F., Hosseinzadeh, A. et al.: A novel approach to voice of customer extraction using GPT-3.5 turbo: linking advanced NLP and lean six sigma 4.0. Int J Adv Manuf Technol. **131**, 3615–3630 (2024). <https://doi.org/10.21203/rs.3.rs-3246823/v1>
 12. Jasińska, K., Lewicz, M., Rostalski, M.: Digitization of the enterprise—prospects for process automation with using RPA and GPT integration. Procedia Computer Science **225**, 3243–3254 (2023). <https://doi.org/10.1016/j.procs.2023.10.318>
 13. Yang, L.F., Chen, H., Li, Z., Ding, X., Wu, X.: Give us the facts: enhancing large language models with knowledge graphs for fact-aware language modeling. In: IEEE transactions on knowledge and data engineering (2023). <https://doi.org/10.48550/arXiv.2306.11489>
 14. Karagiannis, D., Buchmann, R.A., Utz, W.: The OMiLAB digital innovation environment: agile conceptual models to bridge business value with digital and physical twins for product-service systems development. Comput. Ind. **138**, 103631 (2022). <https://doi.org/10.1016/j.compind.2022.103631>
 15. Buchmann, R.A., Karagiannis, D.: Pattern-based transformation of diagrammatic conceptual models for semantic enrichment in the web of data. Procedia Comput. Sci. **60**, 150–159 (2015). <https://doi.org/10.1016/j.procs.2015.08.114>
 16. Ontotext, What is GraphDB?—GraphDB 10.5 documentation. <https://graphdb.ontotext.com/documentation/10.5/>. Accessed 12 Feb 2024
 17. Ontotext, Querying OpenAI GPT Models, <https://graphdb.ontotext.com/documentation>. Accessed 12 Feb 2024
 18. SAP Signavio, SAP Signavio Process Transformation Suite, Academic Edition. <https://www.signavio.com/academic-and-research-alliances/>. Accessed 30 Dec 2023
 19. OpenAI, Security & privacy. <https://openai.com/security/>. Accessed 17 May 2024
 20. OpenAI, OpenAI Privacy Request Portal. <https://privacy.openai.com/policies>. Accessed 17 May 2024
 21. Es, S., James, J., Espinosa-Anke, L., Schockaert, S.: RAGAS: Automated Evaluation of Retrieval Augmented Generation (2023). arXiv:2309.15217. <https://arxiv.org/abs/2309.15217>

Chapter 13

A Hybrid Retrieval-Augmented Generation Approach for Heterogeneous Knowledge Bases



Andrei Chiş  and Ana-Maria Ghiran 

Abstract *Context and motivation.* Large language models have the potential to bring important benefits in the context of processing high volumes of data and documents. LLMs led to the appearance of retrieval-augmented generation in which chunks of relevant data are being injected into large language models prompts in order to ensure the generation of higher quality content, anchored in verified facts and data. *Question/problem.* Large language model generation capabilities can be improved by providing parts of data in a more structured format, expressing information with a lesser amount of characters and tokens, without losing its semantic richness, situation where knowledge graphs can play a vital role in storing, at least partially, facts and data. *Principal ideas/results.* The paper at hand present a hybrid RAG approach in which data is gathered from both natural language documents and RDF knowledge graphs in order to provide an interplay between multiple heterogeneous data sources. *Contribution.* The presented hybrid RAG aims to improve the quality of LLM-based retrieval and querying of data, subsequently offering more flexibility when storing information.

13.1 Introduction

Large Language Models or LLMs [1] became one of the most used technologies from the AI spectrum over the course of the last 2 years due to the advances made in this field by companies such as OpenAI or Google. It has demonstrated its usefulness in scenarios that include human language processing by offering capabilities

A. Chiş · A.-M. Ghiran (✉)

Faculty of Economics and Business Administration, OMILAB@FSEGA, Babeş-Bolyai University, Cluj-Napoca, Romania

e-mail: anamaria.ghiran@econ.ubbcluj.ro

A. Chiş

e-mail: andrei.chis@econ.ubbcluj.ro

of querying on rather unstructured documents through the use of vector stores [2], embeddings [3], and pre-trained models [4].

Having LLMs technology as a foundation, a new framework was developed, named Retrieval-Augmented Generation (RAG), which relies on chunking textual data and storing it in dedicated vector stores, such as ChromaDB [5] having the purpose to improve the quality of LLM-generated content by grounding LLMs on external sources of knowledge in order to supplement their capacity of representing information with access to trusted data, with a higher degree of domain specificity, tailored to different fields of work. The way RAG approaches [6] significantly increase the quality of generated content is by injecting retrieved information as context to the user prompts when they are sent to the large language models, ensuring better quality in response generation, reducing the risk of hallucinations.

Besides all the improvements and advantages provided by LLMs and retrieval-augmented generation, they also tend to have one prominent downside which is correlated to the size of the chunks, in tokens, when they are fed to large language models. Large chunks may lead to making the model less accurate in generating high-quality content as it uses more resources in processing the context. This shortcoming can be addressed by storing information from knowledge bases that are used in RAG processes, totally or partially, as RDF Knowledge Graphs [7] or in other structured data formats.

This paper presents a research in progress that follows the development of a hybrid RAG that aggregates information on one side from natural language documents and on the other side from an RDF graph stored in GraphDB [8], following a Design Science Research approach [9]. By doing so, parts of knowledge can be serialized in machine-readable format which is more efficient in terms of expressing knowledge with a reduced number of tokens. The rest of the paper is structured as follows: Sect. 13.2 presents the research objectives formulated after the Design Science problem template [10]. Section 13.3 provides a solution overview, while Sect. 13.4 discusses implementation details. Section 13.5 presents a literature review, while Sect. 13.6 concludes the paper.

13.2 Objectives Definition

By following the Design Science problem template [10], the following research objective was formulated:

Improving **retrieval-augmented generation capabilities of LLMs tailored for academic or organizational settings** (*problem context*).
... by treating it with a **hybrid layer that aggregates knowledge graphs and natural language documents** (*artifact*).

... to satisfy a need for **higher quality and more factually grounded LLM-generated content** (*requirements*).

... in order to **increase the quality of generated content and dependent organizational flows** (*goals*).

The DSR process is planned to undergo three engineering and evaluation cycles, the paper at hand presenting results obtained over the course of the first cycle.

13.3 Solution Overview

The artifact developed over the course of the first Design Science Research cycle consists of chatbot-like system that performs queries over local heterogeneous data sources, namely natural language text documents and knowledge graphs, by leveraging the models made available by OpenAI [11]. Concretely, the case of a fictional company was analyzed, for which a knowledge base was developed consisting of four elements:

- Textual documents consisting of data about the company's mission and vision;
- Textual documents consisting of data about the company's offered services;
- Textual documents consisting of data about the founder of the company, their previous ventures, and domain expertise;
- A GraphDB-hosted knowledge graph consisting of information about the company's budgets, contracts, clients, and suppliers. The decision to serialize this sort of information as a knowledge graph was taken in the aforementioned fictional company use-case because such data can be easily gathered from ERP systems, legacy databases, or accounting systems of each organization.

On top of the constructed knowledge base there is a web client developed with Streamlit [12] which sends requests to backend developed leveraging LangChain [13], ChromaDB [5], and the SPARQLWrapper Python library [14] for querying knowledge graph endpoints. The backend side also performs the embedding of textual document for being stored on ChromaDB.

Figure 13.1 displays the information flow in the approach presented in this paper. Firstly, a user query is inputted in the frontend component. Then, an execution chain is triggered, using LangChain-specific functions in order to send requests to the large language model. The middleware component also performs the chunking and vectorization of the textual data sources every time the web client is started. On the KG side, the procedure is different as no vectorization is performed but a query is sent in order to return graph fragments that include information relevant to the user query. After data from both textual documents and knowledge graph is obtained, it is used by the large language model as contextual information in order to provide a response for the user query, which is being sent to the frontend component.

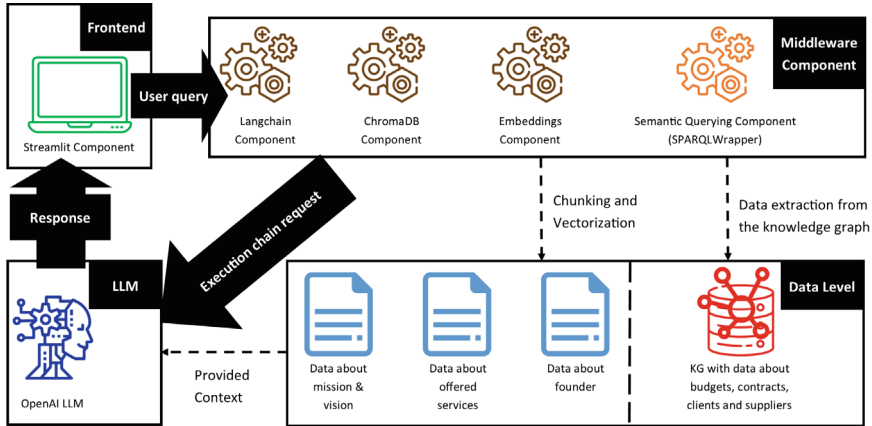


Fig. 13.1 Solution overview

13.4 Implementation Details

The technical implementation is split between two main components:

- The Natural Language Documents Component uses LangChain, ChromaDB, OpenAI LLM (GPT-4), and OpenAI Embeddings in order to perform the chunking, vectorization, and LLM querying over documents written in natural language;
- The Knowledge Graph Component relies on LangChain, OpenAI LLM (GPT-4), GraphDB, and SPARQLWrapper in order to extract query-relevant subgraphs from the main graph which are then used as context for the large language model.

13.4.1 Natural Language Documents Component

In order to process natural language documents a document loading then chunking strategy was needed. LangChain provides dedicated functions for both document loading through the “Directory Loader” function while chunking can be done by using the “Recursive Character Text Splitter” function. After the chunking part, data is being stored on the ChromaDB vector database, going through an embedding process done by accessing the OpenAI embeddings endpoint through LangChain. The algorithm that enables document querying can be split into two main parts: *document processing* and *document querying* using LLMs. These steps are described in Fig. 13.2, as a flowchart diagram.

A code fragment depicting document loading and document chunking functions is available in Fig. 13.3. One notable aspect is in the `split_docs` function used for chunking, namely using the Recursive Character Text Splitter function available in

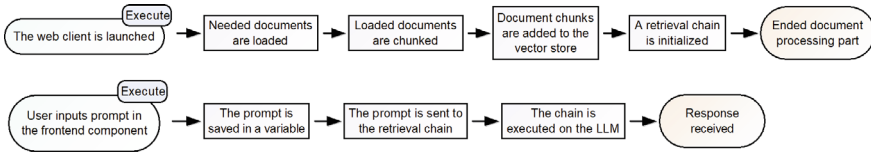


Fig. 13.2 Document processing (upper side) and document querying (lower side)

```

def load_documents(directory):
    loader = DirectoryLoader(directory)
    documents = loader.load()
    return documents

def split_docs(documents, chunk_size=1000, chunk_overlap=200):
    text_splitter = RecursiveCharacterTextSplitter(
        chunk_size=chunk_size,
        chunk_overlap=chunk_overlap,
    )
    docs = text_splitter.split_documents(documents)
    return docs
  
```

Fig. 13.3 Document loading and document splitting functions

LangChain which performs splitting by searching for special characters such as “\n” in order to preserve as much as possible of the context of each chunk unaltered. The chunk size is set to 1000 characters, which represents roughly around 220 tokens per chunk, while the overlapping is set to 200 characters. The overlapping parameter shows how many characters can be shared between two subsequent chunks, in order to help in preserving the context of chunks.

13.4.2 Knowledge Graph Component

In order to implement the knowledge graph component of the knowledge base, a series of concepts were defined starting from schema.org [15] terms. The ontology and instances obtained this way helped in populating the knowledge graph with data about the employees, clients, suppliers, contracts, and budgets of X Corp. The main classes from the obtained ontology can be seen in Table 13.1, while the main properties are available in Table 13.2.

On the properties side, there was a need to have a rich ontology tailored for the specified use-case, even including some properties that are inverse one to each other because in the graph retrieval-augmented generation process, there are selected

Table 13.1 Classes of the knowledge graph ontology

Class name	Superclass	Description
schema:Organization	schema:Thing	Used to instantiate the business entities available in the knowledge graph
schema:Person	schema:Thing	Used to instantiate persons available in the knowledge graph, such as employees or managers
ex:Budget	schema:Thing	Term defined as part of a locally created dictionary used to define URIs of yearly budgets for the company
ex:Business Contract	schema:Thing	Locally defined term to define business contracts of X Corp

only graph fragments, not the whole graph in order to reduce the number of tokens necessary for each call. An excerpt from the knowledge graph can be seen in Fig. 13.4.

The algorithm that is used for gathering graph fragments and adding them as a context for the large language model functions is described in Fig. 13.5. It is functioning in parallel with the natural language document querying algorithm as they both are called when a prompt is inputted by the user in the frontend.

One key aspect of the algorithm is represented by extracting graph fragments using a *DESCRIBE* SPARQL [16, 17] query to identify the subgraph containing information about the subject of the query inputted by the user. This way, instead of loading all the graph's content to the large language model, only the query-relevant part is sent as context. In order to obtain the subject of the user query a prompt is sent to OpenAI. The prompt content was obtained by running a few experiments and performing prompt engineering in order to reach the optimal form:

Identify the subject of the question {question}. When providing the response write only the subject without other comments or details.

The SPARQL query used for the graph fragment gathering part can be seen below:

```
PREFIX : <http://www.example.org#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
DESCRIBE ?x
WHERE{
  {?x schema:name ?name FILTER REGEX(?name , "" + ' ' + subject + ' ' + "" , "i")}
  UNION
  {?x rdfs:label ?label FILTER REGEX(?label , "" + ' ' + subject + ' ' + "" , "i")}
}
```

13.5 Related Works

The work presented in [18] explores the application of large language models over the course of the whole BPM lifecycle mostly relying on textual documents in order to identify processes from documentation or mapping processes through chatbots

Table 13.2 Properties of the knowledge graph ontology

Property name	Domain	Range	Description
schema:Name	schema:Thing	Text datatype	Adds labels to entities from the KG
schema:Url	schema:Thing	Url datatype	Adds URLs for different entities
schema:Employee	schema:Organization	schema:Person	Defines a relationship of employment between a person and an organization
ex:has Budget	schema:Organization	ex:Budget	Defines a relationship between an organization and its yearly budget
schema:Provider	schema:Organization	schema:Organization	Shows a business relationship between two organizations in which the domain organization has the range organization as a supplier
schema:Customer	schema:Organization	schema:Organization	Shows a business relationship in which the range organization is a client for the domain organization
ex:has Contract	schema:Organization	ex:Business Contract	Shows a relationship between an organization and the business contracts in which it is involved
schema:job Title	schema:Person	Text datatype	Used to add job titles for persons who are employees
schema:works For	schema:Person	schema:Organization	Shows the relationship depicting the fact that one person is the employee of an organization
ex:total Amount	ex:Budget	xsd:decimal	Shows the value of a yearly budget
ex:currency	ex:Budget	Text datatype	Shows the currency of a yearly budget
ex:for Year	ex:Budget	xsd:gYear	Shows the year for which a budget was defined
ex:supplies	schema:Organization	schema:Organization	The inverse of schema: provider
ex:client Of	schema:Organization	schema:Organization	The inverse of schema: customer
ex:contractor	ex:Business Contract	schema:Organization	Shows which company is the contractor of a specific business contract
ex:contractee	ex:Business Contract	schema:Organization	Shows which company is the contractee of a specific business contract

(continued)

Table 13.2 (continued)

Property name	Domain	Range	Description
ex:start Date	ex:Business Contract	xsd:date	Shows the start date of a specific business contract
ex:end Date	ex:Business Contract	xsd:date	Shows the end date of a specific business contract
ex:contract Value	ex:Business Contract	xsd:decimal	Shows the value of a specific business contract
ex:currency	ex:Business Contract	Text datatype	Shows the currency of a specific business contract

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix schema: <http://schema.org/> .
@prefix ex: <http://example.com/schema/> .

# Definition of X Corp
ex:XCorp a schema:Organization ;
  schema:name "X Corp" ;
  schema:url <http://www.x-corp.com> ;
  schema:employee ex:JohnDoe, ex:JaneSmith ;
  ex:hasBudget ex:Budget2024 ;
  schema:provider ex:TechSupplier1, ex:OfficeSupplier1 ;
  schema:customer ex:ClientA, ex:ClientB ;
  ex:hasContract ex:Contract1, ex:Contract2, ex:Contract3, ex:Contract4 .

# Budgets
ex:Budget2024 a ex:Budget ;
  ex:totalAmount "5000000"^^xsd:decimal ;
  ex:currency "USD" ;
  ex:forYear "2024"^^xsd:gYear .
    
```

Ontology defined classes and properties for storing use-case relevant knowledge

Fig. 13.4 Knowledge graph fragment

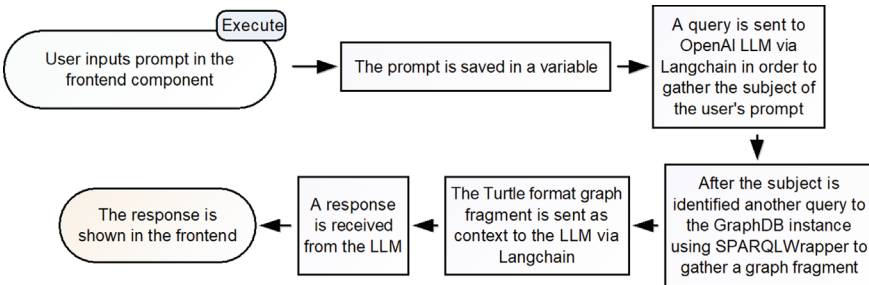


Fig. 13.5 The algorithm for graph querying

which interact with employees or stakeholders, also leveraging XML data formats. The hybrid RAG approach presented in our paper is oriented toward implementing a proof of concept for another organizational use-case, related to data processing from heterogeneous sources, putting an emphasis on RDF serializations of information rather than XML.

The article depicted in [19] describes a retrieval-augmented large language model for sentiment analysis in financial scenarios, by relying on information provided in financial news. The approach presented by Zhang et al. consists of an instruction-tuned LLM and a retrieval-augmented module that gathers additional context from external data sources. Our approach was tailored for a different scenario of providing information gathered from internal organizational databases, but through fine-tuning and different prompting operations, it could be used for sentiment analysis tasks as well.

Taju et al. [20] present an information service virtual assistant that was developed relying on OpenAI GPT-3 models. The mechanism showcased in the paper consists of an intent recognition part, which enhances a semantic search component which ensures better quality in offering responses to user queries. The information retrieval mechanism is based on selecting the closest chunk, obtained through embedding, from the external knowledge bases. Our approach discussed in the paper at hand presents several similarities, but is differentiated by obtaining information not only from textual document databases but also from knowledge graphs, implementing a query subject recognition mechanism supported by AI, which leads to obtaining subgraphs rather than chunks of textual data, as context for the large language model.

Zhao et al. [21] present a survey on different types of information that can be fed to multimodal large language models, such as the ones made available by OpenAI. The article discusses cases in which knowledge formats can range from text, images or audio to graphs, tables, or machine-readable codes. Wenhao [22] provides several supplementary approaches in RAG methods across heterogeneous knowledge by leveraging both structured and unstructured data approaches. Our work is another step in using multimodal knowledge in the context of large language models and retrieval-augmented generation as we aggregate information from textual and knowledge graph sources.

13.6 SWOT-Based Conclusions

This paper presents a Design Science motivated approach of combining data from natural language documents and knowledge graphs when building retrieval-augmented generation-based chatbots and virtual assistants. Unstructured and structured data can be combined, providing better results and more specificity in satisfying user prompts. The artifact is still in an early stage, being in its first Design Science cycle, therefore several evaluation mechanisms will be defined for future cycles from the foundations laid in [23]. Below, a SWOT analysis is summarized, which will be used as a starting point for future iterations of the presented artifact:

Strengths: The proposed hybrid RAG supports integrating data from heterogeneous data sources, namely natural language documents and semantic graph databases. This approach improves generation capabilities as at least parts of the knowledge bases that are fed to large language models are serialized in a machine-readable format, expressing more information as structured data, to help in generating better quality responses.

Weaknesses: Currently, this approach is built and limited to using OpenAI endpoints, but this can be mitigated thanks to the integrations to other large language models made available by LangChain. Another weak point resides in the need for defining a reliable evaluation framework for this approach, in order to furtherly investigate its practical applicability across different settings and scenarios.

Opportunities: Future work can explore the addition of LPG [24] graphs alongside RDF ones as sources of knowledge for the RAG artifact. Moreover, the presented ontology can be refined in order to support multiple use-cases, also exploring the possibilities of automated text-to-knowledge graph generation by implementing large language model integrations.

Threats: The current approach is dependent on dedicated graph triple stores which gained more popularity over the course of the last years, but their licensing costs and implementation complexity can still pose blockades for small or medium organizations.

References

1. Shanahan, M.: Talking about large language models. *Commun. ACM* **67**(2), 68–79 (2024)
2. Chaki, E., Microsoft, What is a vector database?. <https://learn.microsoft.com/en-us/semantic-kernel/memories/vector-db>. Accessed 26 Feb 2024
3. OpenAI Foundation, OpenAI Embeddings. <https://platform.openai.com/docs/guides/embeddings>. Accessed 01 Mar 2024
4. Han, X., Zhang, Z., Ding, N., et al.: Pre-trained models: past, present and future. *AI Open* **2**, 225–250 (2021)
5. Chroma, ChromaDB. <https://docs.trychroma.com/>. Accessed 28 Feb 2024
6. Gao, Y., Xiong, Y., Gao, X. et al.: Retrieval-augmented generation for large language models: a survey. *ArXiv* (2023)
7. W3C Recommendation, RDF resource description framework 1.1 concepts and abstract syntax. <https://www.w3.org/TR/rdf11-concepts/>. Accessed 25 Feb 2024
8. Ontotext, GraphDB Semantic Graph Database. <https://graphdb.ontotext.com/>. Accessed 29 Feb 2024
9. Wieringa R. J.: *Design science methodology for information systems and software engineering*. Springer (2014)
10. Wieringa R. J.: *Design Science Research Methods*. <https://wwwhome.ewi.utwente.nl/~roelw/DSM90minutes.pdf>. Accessed 10 Feb 2024
11. OpenAI Foundation, OpenAI Models. <https://platform.openai.com/docs/models>. Accessed 15 Feb 2024
12. Snowflake Inc., Streamlit. <https://docs.streamlit.io/>. Accessed 20 Feb 2024
13. LangChain Inc., LangChain Framework. https://python.langchain.com/docs/get_started/introduction. Accessed 20 Feb 2024

14. Herman, I., Fernández, S. et al., SPARQLWrapper, <https://sparqlwrapper.readthedocs.io/en/latest/>. Accessed 22 Feb 2024
15. W3C Schema.org Community Group, Schema.org. <https://schema.org/>. Accessed 22 Feb 2024
16. W3C, SPARQL 1.1 Query Language. <https://www.w3.org/TR/sparql11-query/>. Accessed 22 Feb 2024
17. Pérez, J., Arenas, M., Gutierrez, C.: Semantics and complexity of SPARQL. *ACM Trans. Database Syst.* **34**(3), 1–45 (2009)
18. Vidgof, M., Bachhofner, S., Mendling, J.: Large Language Models for Business Process Management: Opportunities and Challenges. In: *Business Process Management Forum. BPM 2023, Lecture Notes in Business Information Processing*, vol. 490, pp. 107–123, Springer, Cham (2023)
19. Zhang, B., Yang, H. et al.: Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models. In: *Proceedings of the Fourth ACM International Conference on AI in Finance*, pp. 349–356, Association for Computing Machinery (2023)
20. Taju, S. W., Ferry Mandias, G., et al.: AI-powered Chatbot for Information Service at Klabat University by Integrating OpenAI GPT-3 with Intent Recognition and Semantic Search. In: *2023 5th International Conference on Cybernetics and Intelligent System*, pp. 1–6, IEEE Xplore (2023)
21. Zhao, R., Chen, H. et al: Retrieving Multimodal Information for Augmented Generation: A Survey. *Findings of the Association for Computational Linguistics: EMNLP 2023*, Association for Computational Linguistics, 4736–4756 (2023)
22. Wenhao, Y.: Retrieval-augmented Generation across Heterogeneous Knowledge. In: *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop*, Association for Computational Linguistics, pp. 52–58 (2022)
23. Prat, N., Comyn-Wattiau, I., Akoka, J.: Artifact evaluation in information systems design-science research – a holistic view. In: *Proceedings of Pacific Asia Conference on Information Systems*, **23**, AIS eLibrary (2014)
24. Sharma, C., Sinha, R.: A Schema-First Formalism for Labeled Property Graph Databases: Enabling Structured Data Loading and Analytics. In: *Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, pp. 71–80, Association for Computing Machinery (2019)

Chapter 14

Digitalization in Higher Education—Designing and Implementing an Application to Facilitate the University Final Thesis Collaboration Process



Cristina Iancu , Simona Vasilica Oprea , and Cristian-Eugen Ciurea 

Abstract In the academic universe, the process of registering for the bachelor’s degree final examination often poses major logistical challenges for both parties: students and faculty. Because traditional procedures are marked by bureaucracy and inefficient communication, this important part of scholarly study needs to be streamlined using new technological solutions. In order to streamline administrative procedures and promote smooth collaboration between students and instructors, this paper presents a novel web application that aims to transform the registration process. With the aid of contemporary technology and innovative design principles, this application aims to lessen current challenges and promote greater collaboration between students and professors, by seamlessly bridging the gap between both categories. Additionally, this solution offers a unique point of access for the head of department, secretaries, students, and professors, where the flow of documents can be verified at any time and from any location.

14.1 Introduction

With the onset of the “Digitalization in A.S.E.” competition, an interest was piqued in developing the application that will be described in the following pages. This application seeks to empower students to seamlessly register for their bachelor’s

C. Iancu (✉) · S. V. Oprea · C.-E. Ciurea
Department of Economic Informatics and Cybernetics, The Bucharest University of Economic Studies, Bucharest, Romania
e-mail: iancu2cristina19@stud.ase.ro

S. V. Oprea
e-mail: simona.oprea@csie.ase.ro

C.-E. Ciurea
e-mail: cristian.ciurea@ie.ase.ro

degree thesis defenses while also facilitating the process of connecting with a suitable collaborating professor. Inspired by the actual challenges faced during being a senior year student, we embarked on a mission to craft an innovative solution, thus choosing this field. Students' experiences, in which they struggle with the difficulties of communicating with professors during the final stage of their undergraduate education, served as inspiration for our choice [7].

Before the creation of this application, students would often reach out to several professors via email seeking supervision for their thesis, with uncertainty concerning the availability of spots, given the absence of a digital platform to streamline the allocation process. Furthermore, finalizing the thesis title and formalizing collaboration agreements entailed navigating through a deluge of emails (Chelsea [2]).

The need of a platform to facilitate interaction and communication between students wishing to enroll in the thesis examination and available supervising professors was evident. Through a client-server web application, any disadvantages from the scenario are eliminated, such as the uncertainty of finding a professor with available spots. Moreover, a concrete and official record is maintained throughout the entire search, request, confirmation, and signing process. With that being said, efficiency, speed, and security are combined in a platform that can streamline the allocation process through which hundreds of students and professors pass annually (Noble 2002).

14.2 Similar Applications in the Academic Field

Romanian universities have developed, to varying degrees, platforms, websites, or applications aimed at facilitating the enrollment process for thesis defense and graduation process. But the majority of these systems fall short in covering a good deal of the functionality needed for a completely automated procedure. Generally, they only provide information, necessary steps, a few documents, and maybe some forms for uploading files, with the remaining steps needing to be completed through other means.

Faculty of Mathematics and Computer Science at the University of Bucharest

By far, the most comprehensive and intriguing platform we will discuss is the one utilized within this faculty. Graduating students can easily and efficiently find a supervising professor for their thesis through the creation of an account (using their institutional email), providing personal details, and selecting an area of interest to receive suggestions. The entire collaboration process is covered by this application, serving as a true inspiration model for our application. Students can see which professors are available at the time of authentication and what topics they can address and then request collaboration. One aspect that stands out, and which the application we have developed possesses, is the ability to automatically generate a thesis request based on the details provided in the personal account, as well as the existence of

these data. But, within the platform of this other faculty, students can add very few personal details.

14.3 Method Used for Design and Implementation

14.3.1 Functionalities

The functionalities encapsulated within the application are meticulously defined and seamlessly align with the specified requirements. This owes much to the recurrent consultations convened with the guiding professors, alongside the invaluable support extended by other esteemed faculty members throughout this journey.

The application caters to both categories of users: professors and students, clearly defining them as two distinct user types, with special pages. For authentication, the application utilizes the Google account associated with the institutional email provided by the faculty. This process relies on an external service integrated into the application, known as the Google API. Furthermore, the system incorporates an Administrator account type, overseen by the department director within the faculty. This account has specialized functionalities, such as adjusting the maximum student limit allocated to coordinating professors and generating comprehensive reports. As a result, the administrator gains access to view and modify any information within the application. Authentication for this account type utilizes a distinct no-reply mechanism, featuring a unique interface accessed via a separate URL.

The application shows a user-friendly interface catering to the specific user's needs as soon as they log in. They can easily enter, view, and modify essential personal details required for the registration process by creating a detailed profile that captures everything the program needs. This feature ensures that all information provided is accurate and up-to-date and that both parties are guided through the maze of bachelor's degree registration with just a few clicks because, at its core, this web app is all about efficiency.

Professors gain a comprehensive insight into their academic cohorts, empowering them to efficiently manage their engagements. This includes access to a detailed page of accepted students, real-time updates on available slots and current occupancy, as well as visibility into pending collaboration requests and documents awaiting signature. Such transparency empowers professors to make informed decisions and optimize their academic collaborations.

14.3.2 Projecting the Process

In this subchapter, the structure and design of the application will be depicted with the aid of visual representations like UML diagrams [4].

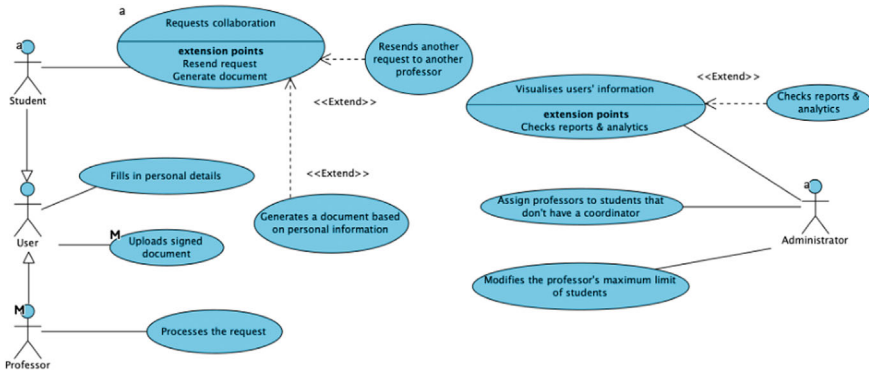


Fig. 14.1 General use case diagram

The primary feature of the app is its ability to connect students with professors. Through a filtered list, students can easily find professors who match their academic profile and who are available to collaborate with for their final paper project. With just a click of a button, they can express their interest in working with a specific professor.

Figure 14.1 below illustrates the main use case diagram corresponding to the computer system presented in this paper. It outlines, overall, how the flow of cases will unfold, but also the general functionalities of the web application. The main actors who will interact with the system are highlighted, each capable of performing various operations, which will be detailed further.

Following a student’s initial request, a 3-day grace period is established for the professor to provide a response. If the response is not received within this timeframe, the student can explore alternative avenues of collaboration by submitting a request to another professor. Importantly, to maintain integrity and prevent redundancy, the app prohibits the submission of duplicate requests to the same professor, thus ensuring a balanced and efficient process for both students and professors alike.

On the other side, professors are provided with a user-friendly interface to review and respond to these requests within a designated timeframe, thereby promoting transparency and accountability in the selection process. They are afforded with the flexibility to select suitable candidates for academic collaboration based on their expertise and availability. This not only saves time but also opens possibilities beyond the traditional classroom.

Upon acceptance of a collaboration request, students can generate a standardized PDF document incorporating their profile information and project details. This document serves as a formal agreement between the student and professor, outlining their respective roles and responsibilities throughout this academic process. Feedback mechanisms are integrated to facilitate ongoing communication and revisions as necessary, ensuring a collaborative and iterative approach to academic inquiry.

Administrative personnel play a pivotal role in overseeing the system, possessing the capability to adjust critical parameters such as the maximum student quota allotted

to individual professors. Moreover, they hold jurisdiction over approving requests for altering bachelor’s degree paper titles and allocating unassigned students, ensuring optimal resource allocation and utilization.

Thus, focusing on administrative efficiency, this type of user is vested with comprehensive administrative privileges, enabling him to oversee, append, amend, or remove data as necessary. Furthermore, they have the authority to generate detailed reports, providing valuable insights for informed decision-making and strategic planning.

14.3.3 The Application Structure

The Class Diagram serves the purpose of describing the static component used to efficiently implement all the functionalities outlined in the diagrams above, namely the entities within the system. Figure 14.2 shows the existing eight classes, along with their attributes, operations, and relations [5].

User—this class encapsulates the concept of the application’s user, containing the necessary authentication details, along with some common methods.

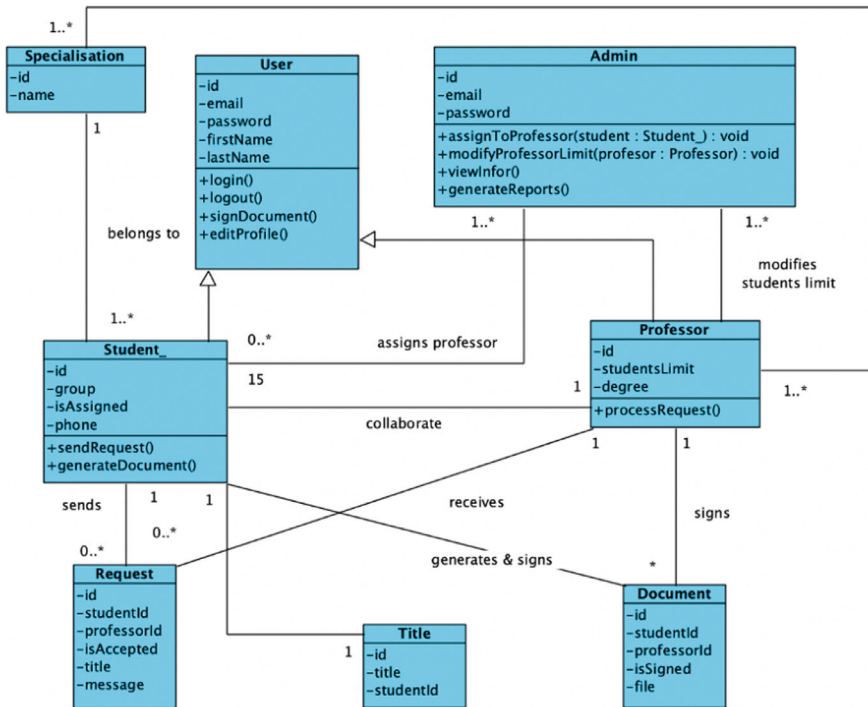


Fig. 14.2 Class diagram

Student—derived from **User**, this class represents any student within our faculty who is in their final or additional year and wishes to defend their bachelor’s final paper in the current academic year.

Professor—inheriting the **User** class, the professor is also a member of this faculty, capable of supervising students in the development of their project.

Administrator—this class is the last of the concrete representatives of the application, also acting as a stakeholder within the system. They have special duties, being able to view and modify information within the application.

Specialization—represents an attribute of the user, with each of the two derived classes belonging to a specialization of the faculty.

Request—within the collaboration request flow, this class is essential, serving as the first step in the communication between student and professor. A student can submit multiple requests to several professors, but never multiple requests to the same professor.

Title—this entity encapsulates the potential title of the paper that a student can save in the database, before deciding its final version.

Document—like a “contract” between the student and professor, the request seals the collaboration between the two actors of the system. Once a collaboration request is accepted, a document is generated, signed, and uploaded containing the necessary details for this targeted collaboration within the application.

14.4 Results Obtained in a Real Scenario

The main page, featuring the authentication form, is the first encounter users have with the application, accessible at <https://als.csie.ase.ro>. The login button triggers a modal redirect to Google since a Gmail account, specifically the institutional one provided by the university, is mandatory. Additionally, users can review the terms and conditions required for the application to connect to their Google account and process the requested information.

Both users are initially prompted with an on-screen message indicating the necessity of completing their personal details to initiate the application process. Thus, the student provides information such as group, phone number, year of enrollment, study year, form of education, and specialization to view corresponding professors. Similarly, the professor must specify their specialization or specializations, with the possibility of multiple selections. Following this step, professors from the same specialization become visible, but only after the administrator validates the professor’s account by setting the maximum student limit to 15 or another nonzero natural number. By default, the limit is set to 0 to prevent any university employee with an institutional account from logging in, thereby misleading students and disrupting the process.

Upon completion of the data by both parties, the student encounters an interactive table facilitating filtering, pagination, and sorting to streamline the search process. The displayed professors are those within the same specialization, with positive and

nonzero available spots. If a student has been previously rejected by a professor, they will not appear in the table.

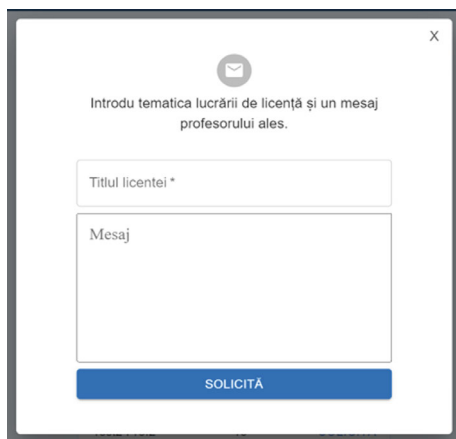
In this stage, any professor can be selected by clicking the “Request” button. This button opens a personalized modal window with two fields (Fig. 14.3): a textbox for entering the initial thesis topic and a text area for conveying a message to the professor to facilitate the acceptance process. A brief overview of academic progress, chosen topics, and technologies can be provided, along with an expression of collaboration interest. By clicking “Request,” the entered data is transmitted to the professor within the application, accompanied by an email alert.

Following the submission of the request, the student cannot submit a new one to another professor, nor can they withdraw the pending one earlier than three days. Subsequently, the “scene” shifts to the professor’s interface, who is notified of the pending response while the student awaits. The professor reviews the details provided by the student, with the option to reject or accept, as seen in Fig. 14.4.

Upon clicking either button, the professor confirms their choice, and the response is transmitted to the student within the application and via email. In the case of an affirmative response, a spot is reserved from the available ones, and the list of accepted students in the professor’s interface’s second tab is updated. Otherwise, the student cannot resubmit the request to the same professor and must repeat the process until they find a coordinator.

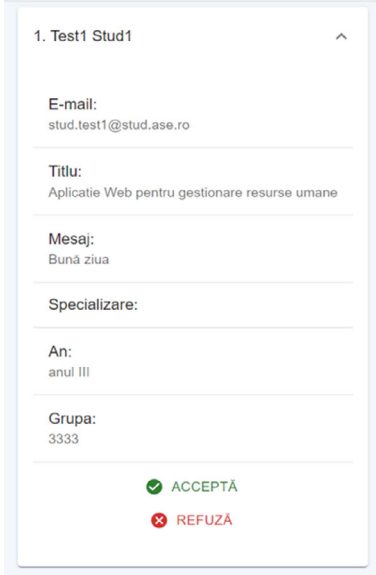
The subsequent major step in this flow is the request cycle. It begins with the automatic document generation within the platform by the student. The accepted title by both parties is all that is needed to generate and upload the request. An important aspect is that the document must be initially signed by the prospective graduate before being transmitted to the coordinator. Similarly to the request submission, once a PDF is submitted, the student must await a response from the professor, whether positive or negative. Consequently, they cannot resubmit or overwrite what has already been sent, ensuring data persistence.

Fig. 14.3 Request form



The image shows a modal window titled "Request form" with a close button (X) in the top right corner. At the top center is an envelope icon. Below it, the text reads: "Introdu tematica lucrării de licență și un mesaj profesorului ales." (Enter the thesis topic and a message to the chosen professor). There are two input fields: a text box labeled "Titlul licenței *" (Thesis title *) and a larger text area labeled "Mesaj" (Message). At the bottom center is a blue button labeled "SOLICITĂ" (Request).

Fig. 14.4 Awaiting student details



1. Test1 Stud1

E-mail:
stud.test1@stud.ase.ro

Titlu:
Aplicatie Web pentru gestionare resurse umane

Mesaj:
Bună ziua

Specializare:

An:
anul III

Grupa:
3333

✓ ACCEPTĂ

✗ REFUZĂ

Upon receiving an email notification, as well as within the website, the professor can view the title and student details of who has just uploaded the document and can either directly refuse or download, sign, and reupload the document.

Regardless of the professor's response, their coordinated student will be alerted via email. In the case of rejection, a reason can also be provided to offer guidance and suggestions for the future request to be accepted. If the request is ultimately signed by both parties, the student can download the final version from the third tab—otherwise, they will resubmit the request until accepted.

Once this step is completed, the collaboration process concludes successfully. All that remains is for the student to stay in touch with the coordinating professor for developing the thesis.

However, if the student wishes to change their thesis title after a request has already been accepted during the registration period, they can do so by simply generating and uploading a new request. Similarly to before, the modification can be either accepted or refused. The final request will always be the most recently accepted, regardless of the number of submissions.

If the coordinating professor runs out of available spots, the administrator can increase the maximum limit allowed. This update can be made in exceptional cases or if there are students unassigned to a coordinator. Additionally, the administrator can manually allocate those less fortunate (without any coordinator) via their dedicated page.

14.4.1 Importance of the Application

The significance of this web application extends beyond simple convenience; it represents a paradigm shift in the way academic collaboration is facilitated and administered. The bachelor's degree registration procedure has been streamlined to enable students to participate more actively in their research projects and to give faculty members better insight into the interests and skills of their students [8].

Furthermore, the application serves as an inspiration for fostering mentorship and professional development opportunities within the academic community. By facilitating direct communication and collaboration between students and professors, it creates an appropriate environment for knowledge exchange and communication.

Compared to traditional approaches, web applications offer unparalleled accessibility and flexibility. Users may manage their profiles from anywhere at any time by using this platform, which is accessible from any device with an internet connection. This is essential in a dynamic environment in which we live nowadays, where both professors and students have different, sometimes chaotic schedules.

Furthermore, the openness and visibility of this whole coordinating process are significantly increased by centralizing information into a single database. Clearer and more effective communication between the parties is made possible by the ability for students to rapidly check the status of their applications and for professors to manage lists of accepted students and incoming applications.

Lastly, a fluid transfer from conventional techniques to a contemporary digital working environment is ensured by this application's user-friendly interface, which is developed to different levels of digital competence and aids all types of users.

14.4.2 Impact of Administrative Efficiency

The adoption of this program has greatly increased efficiency from the standpoint of faculty management. The head of the department, known as administrator, can simply oversee and control every step of the procedure, from assigning professors to approving and signing applications, thanks to the admin interface. Wisely taking decisions when necessary is made easier by the capacity to create comprehensive reports and monitor the current status of requests in real time. By doing this, people from management can guarantee the fair and effective distribution of resources, ensuring a more structured and functioning learning environment. In addition to this, the program has raised student happiness and made faculty management's administrative tasks easier. The significance of using technology in education and academic administration is shown by these combined advantages. Also, the data speaks for itself. As shown in Table 14.1 below, a great number of students have registered the platform, generating one or more requests, judging by the numbers. The reason behind the discrepancy between the smaller number of students (1155) compared to the one for requests (1224) is that a request can be denied by the professor, thus

Table 14.1 Real number of entries in the database

Entity name	Number of entries
Student	1155
Professor	78
Request	1224
Document	1083

leading to the generation of a new one. Unfortunately, not every student manages to find a collaborator until the end, so the data shows us that last year 72 students did not upload a document that would seal an agreement with a professor.

14.5 Conclusion

A technologized university represents an academic environment fully adapted and integrated into the comprehensive informational context of society, a conducive and perfectly suitable environment for innovation, through which both performance and academic excellence can be achieved. It is also important to consider that technology is not a substitute for existing activities but rather an adjunct. Although the physical mode of teaching, face-to-face, will continue to dominate among teaching methods, technology adds value and enhances the efficiency of the education system. The complete automation of routine activities, such as registration for final exams, allows all involved parties to focus their attention and interest on aspects that are truly important, such as imparting knowledge to future generations, innovation, or even research [1].

The idea of transforming a repetitive and manual activity into an automated one is the foundation of modern technology, as people seek to optimize their time and resources to redirect them to areas that truly require human intervention. Thus, bureaucratic processes, which involve soliciting and completing information and requests, can easily be replaced by a specially designed web system.

Recognizing that this remains a work in progress, the commitment is to continually refine its functionality to ensure accuracy, coherence, and efficiency. By addressing any emerging issues promptly, the aim is to create a seamless experience for all users. Moreover, as the scope expands to encompass dissertation coordination, the application will serve as an invaluable guide for every student in our faculty. Looking ahead, there is potential for further expansion to encompass all faculties within our university, making it an indispensable tool for our academic community.

In conclusion, the application described in this paper represents a pioneering step toward enhancing academic collaboration and administrative efficiency within the realm of higher education. By using the power of technology and innovation, it offers a scalable solution to the various challenges associated with the enrolling for the bachelor's degree examination, thereby empowering students to navigate their academic journey with confidence and clarity. As the educational landscape continues

to evolve, the role of digital tools in facilitating scholarly inquiry and collaboration will undoubtedly become increasingly indispensable.

In the future the application will be extended with features to measure the users' behavior, in order to improve the correspondence between students and professors in terms of bachelor collaboration. Indicators such as the number of authentications, the number of requests accepted or rejected, and the number of titles changed during the collaboration will be introduced in the application. These characteristics will improve the overall process and it will facilitate a better collaboration between all the actors involved (students, professors, secretaries, and department director) [3].

Acknowledgements This work was supported by a grant from the Ministry of Research, Innovation, and Digitization, CNCS/CCCDI—UEFISCDI, project number COFUND-DUT-OPEN4CEC-1, within PNCDI IV.

References

1. Angie.: Why Automation for Universities Is a Must-Have. Retrieved 20 Mar 2024 (2024). <https://www.cazoomi.com/blog/why-automation-for-universities-is-a-must-have>
2. Chelsea Maher.: How Automation Can Make Educational Institutions More Efficient | Culture Shift. Retrieved 20 Mar 2024 (2021). <https://culture-shift.co.uk/resources/higher-education/how-automation-can-make-educational-institutions-more-efficient/>
3. Ciurea, C., Oprea, S.V.: Digital Humanism and a quantitative analysis of users behavior in virtual exhibitions during the pandemic context. *Procedia Comput. Sci.* **214**, 384–390 (2022)
4. GeeksforGeeks.: Class Diagram | Unified Modeling Language (UML) - GeeksforGeeks. Retrieved 20 Mar 2024 (2023) <https://www.geeksforgeeks.org/unified-modeling-language-uml-class-diagrams>
5. GeeksforGeeks.: Use Case Diagrams | Unified Modeling Language (UML) - GeeksforGeeks. Retrieved 20 Mar 2024 (2023). <https://www.geeksforgeeks.org/use-case-diagram>
6. Noble, David F.: Digital Diploma Mills : The automation of higher education, Monthly Review Press, pp. 116 (2003)
7. Selwyn, N., Hillman, T., Rensfeldt, A.B., Perrotta, C.: Digital technologies and the automation of education — key questions and concerns. *Postdigital Sci. Educ.* **5**(1), 15–24 (2023)
8. Williamson, Ben .: The Automatic University: A Review of Datafication and Automation in Higher Education, University and College Union, Retrieved 20 Mar 2024 (2020) https://www.ucu.org.uk/media/10947/The-automatic-university/pdf/ucus_the-automatic-university_un20.pdf

Part III
IoT, Mobile and Multimedia Solutions;
Cybersecurity and Critical Infrastructures

Chapter 15

An Overlook of IoT Finance and Pre-existing Security Challenges in Financial Transactions via IoT Devices on Blockchain Networks



Alisa Harkai

Abstract The nearness of IoT (Internet of Things) gadgets in everyday life these days is expanding, each of us owning at slightest one IoT gadget. Therefore, the adoption of this technology is as global as possible and spread on a very large scale throughout the civilized world. Utilizing IoT gadgets most of the time on a day-by-day premise, we end up doing nearly any action with their offer assistance, as we also make online payments through different web or mobile applications, or specifically with the phone/smartwatch at physical portable POS (Point of Sale). For people, this possibility of carrying out financial transactions online is no longer a novelty, but we must indicate the fact that these exchanges are attempted through a mediator, specifically the issuing Bank of each client. Regarding financial transactions, the novelty exists when we talk about making them with cryptocurrencies, so using the same smart or IoT devices to make payments, but this time we no longer do them through an intermediary, i.e., the Bank, we do them directly on the so-called blockchain. The main purpose of this paper is to cover a brief overlook of how these financial transactions made on blockchain through IoT devices are vulnerable from a cybersecurity point of view.

15.1 Introduction

The Internet of Things (IoT) has become a transformative paradigm, revolutionizing the way devices interact and communicate across a variety of sectors, including finance [1].

IoT devices, from smart cards to wearables, have become ubiquitous, facilitating seamless data exchange and automating tasks. The widespread popularity of these

A. Harkai (✉)

Faculty of Cybernetics, Statistics and Economic Informatics, Bucharest University of Economic Studies, Dorobanti 15-17, 010552 Bucharest, Romania
e-mail: harkaialisa22@stud.ase.ro

IoT devices has expanded their utility to financial transactions, providing unparalleled convenience and accessibility.

However, integrating IoT devices into financial systems poses new security challenges that require innovative solutions.

Financial transactions conducted through IoT devices present unique opportunities and risks. On one hand, IoT-enabled payment methods, such as contactless payments and mobile wallets, streamline transactions, improve user experience, and promote financial inclusion. On the other hand, the connected nature of IoT devices exposes them to countless security threats, including data breaches, identity theft, and unauthorized access. As financial institutions increasingly adopt IoT technology to meet growing consumer demands, protecting the integrity and security of transactions becomes paramount.

In response to these challenges, blockchain technology has emerged as a promising solution to improve the security and transparency of financial transactions conducted through IoT devices [2].

Blockchain, a decentralized and immutable ledger, provides a tamper-proof platform for recording and verifying transactions in a trustless environment. By leveraging cryptographic techniques and consensus algorithms, blockchain ensures the integrity and authenticity of financial data exchanged between IoT devices.

Additionally, the decentralized nature of the blockchain network eliminates single points of failure, reducing the risk of cyberattacks and unauthorized modifications.

This article aims to explore the convergence of IoT devices and blockchain technology to facilitate financial transactions. Drawing on insights from academic literature and scientific research, we examine the security implications of using IoT devices for financial transactions and the role of blockchain in mitigating risk-related risks.

15.1.1 General Aspects of Financial Transactions

We can say that a financial transaction represents any act of exchanging or substituting financial capital for another with a different maturity [3, 4]. Financial transactions are an integral part of the economic system, facilitating the exchange of goods, services and assets between different stakeholders.

They cover a wide range of activities, from everyday purchases to complex investment transactions. Because of their importance, researchers have extensively studied the role of financial transactions in economic development and the functioning of financial markets.

You don't have to be a specialist to carry out financial transactions. In fact, almost everyone can do it. Most people interact with different banks to accomplish specific financial objectives. These interactions can take place both in person and digitally, and we'll look at what are the main types of banking activities below. There are two important types of banking, namely: traditional banking and modern banking, the second one being classified in online and digital banking.

1. Traditional banking:

- Known by the fact that it refers to financial institutions that operate in physical branches and also hold national banking licenses [5];
- Financial transactions are conducted face-to-face or in one-to-one customer service [5];
- Considered to be safe in general and easier for some of users in making withdrawals or in handling cash payments [5];

2. Online banking:

- Any type of personal banking that we perform online is referred to as online banking: from monitoring our balance and past statements to establishing direct debits, making payments, or transferring funds to a different person's account [6];
- Simple banking transactions are available seven days a week, at any time of day or night [7];
- Much faster than traditional banking, but comes with other challenges such as cyber vulnerabilities (e.g., unauthorized access) [7];

3. Digital banking

- All financial transactions including technology are referred to as “digital banking”. Considering this fact, we can say that online banking is only a part of digital banking. [6];
- Digital banking refers to the various digital solutions offered to businesses to help them run more efficiently, such as accounting and payroll software and payment technologies. This could be a physical store or an internet merchant [6];

In this article we will refer exclusively to online banking and financial activities undertaken online and we will see in the next subsection some general information about how IoT devices are used in their implementation.

15.1.2 IoT's Impact on Banking and Finance

In recent years, the integration of Internet of Things (IoT) technology has emerged as a promising avenue for revolutionizing banking and financial transactions. IoT devices, ranging from smart sensors to wearables, are increasingly being utilized to streamline processes, enhance security measures, and provide real-time insights into customer behavior and preferences.

One significant application of IoT in finance is the optimization of ATM (Automated Teller Machine) networks. Internet-connected Automated Teller Machines (ATMs) are a critical development in IoT for financial transactions. Unlike traditional ATMs, these devices leverage IoT technology to offer enhanced services and connectivity. They can connect to banking networks in real time, providing customers

with up-to-date account information and enabling a range of transactions beyond cash withdrawals, such as bill payments and fund transfers. Advanced ATMs also feature biometric authentication methods, including fingerprint and facial recognition, to improve security and user experience. The real-time data capabilities of IoT-enabled ATMs facilitate better service delivery, immediate fraud detection, and reduced downtime through predictive maintenance. This evolution of ATMs underscores the transformative impact of IoT on traditional banking infrastructure. We can sure make financial transactions through other IoT digital devices like:

- **Smart Payment Wearables**

Smart payment wearables, such as smartwatches and fitness trackers, are increasingly popular IoT devices facilitating financial transactions. These devices often incorporate Near Field Communication (NFC) technology, allowing users to make contactless payments. For instance, Apple Watch and Fitbit Pay enable users to link their bank accounts or credit cards to the device and pay for goods and services by simply tapping their wearable on a compatible point-of-sale terminal. This technology not only enhances convenience but also adds a layer of security as it typically requires biometric authentication (e.g., a fingerprint or facial recognition) to complete a transaction. The integration of IoT in wearables exemplifies how technology can simplify everyday financial interactions, making transactions seamless and secure.

- **Connected Point-of-Sale (POS) Systems**

Modern connected POS systems are another prime example of IoT devices revolutionizing financial transactions. These systems, such as Square and Clover, are equipped with advanced sensors and connectivity options, including Wi-Fi and Bluetooth, enabling them to process payments quickly and efficiently. Connected POS systems can accept various payment methods, including traditional magnetic stripe cards, chip cards, and contactless payments from mobile devices or wearables. They also offer additional functionalities like inventory management, sales analytics, and customer relationship management. The adoption of IoT in POS systems enhances the transaction process by providing real-time data and improving overall business operations through automation and connectivity.

- **Smart Vending Machines**

Smart vending machines represent a significant leap forward in the application of IoT in financial transactions. These machines are equipped with various sensors and internet connectivity, allowing them to accept multiple forms of payment, such as credit/debit cards, mobile payments, and even cryptocurrencies. For example, Coca-Cola has implemented smart vending machines that accept payments through mobile wallets like Apple Pay and Google Wallet. These machines also provide real-time inventory data to operators, enabling more efficient restocking and maintenance. Additionally, smart vending machines can offer personalized promotions and discounts based on customer data, enhancing the purchasing experience and driving sales. The integration of IoT in vending machines illustrates the potential for automated and intelligent transaction processes in everyday scenarios.

By equipping ATMs and the rest of devices with IoT sensors and monitoring devices, banks can remotely track the operational status of machines, detect potential malfunctions in advance, and ensure timely maintenance. According to a study conducted by McKinsey & Company [8], this proactive approach not only reduces downtime but also minimizes the risk of fraudulent activities, thereby safeguarding customers' funds.

Furthermore, IoT-enabled devices play an important role in increasing the security of financial transactions. For example, biometric authentication technologies included in IoT-enabled payment terminals provide a more secure and convenient alternative to traditional methods like PINs or signatures. This strategy reduces the danger of identity theft and unauthorized access, according to research published in the *Journal of Information Security and Applications*.

In the next paragraphs, we will talk a little about the trend of using IoT devices when making financial transactions. We can see in different research studies by that a significant surge was indicated in the usage of IoT devices among users for banking and financial purposes. As consumers increasingly seek seamless and personalized experiences, the convenience and efficiency offered by IoT-enabled solutions are expected to drive adoption rates. One prominent trend anticipated in the coming years is the proliferation of IoT-powered wearables for financial transactions. Many studies suggest that wearables, including smartwatches and payment-enabled accessories, will become ubiquitous tools for conducting secure and contactless payments. With built-in biometric authentication and tokenization capabilities, these wearables offer users a convenient and secure means of accessing their financial accounts and making transactions on the go.

Moreover, the utilization of IoT devices for personalized financial management is forecasted to gain traction. Through IoT-powered analytics and recommendation systems, users can expect tailored financial advice and insights to optimize their financial health and achieve their long-term goals.

So, we can see that there is indeed a growing trend in terms of the use of IoT devices to do banking, trend which is also shown in Table 15.1 which can be seen below, as well as in Fig. 15.1.

Table 15.1 illustrates actual data for the year 2020 concerning mobile payment transactions, along with some forecasted data for the year 2025, indicating their growth in several European countries, with CAGR representing compound annual growth. We can surely see that the trends will grow and the growing rate will in some cases increase exponentially.

Figure 15.1 also emphasizes the same forecasted growth of mobile payment transactions conducted by the study from [9] but in a more graphical way.

For a brief discussion on Romania, Table 15.2 provides specific data from 2023, as per a study conducted in January 2024 [10], detailing the manner in which Romanians utilize digital devices for diverse online operations, payments, and financial transactions, specifically highlighting the percentage of total Romanian users (users over fifteen years) who employed online apps for these purposes. Additionally, the table also provides the same activities represented as percentages for the year 2021 [11], clearly indicating the increasing use of IoT for transactions in Romania.

Table 15.1 Growth of mobile payment transactions (Adapted from [9])

Europe country	2020	2025	CAGR (Compound annual growth) (%)
Germany	17,120.40	105,219.80	43.8
France	16,822.70	66,892.60	31.8
Spain	13,319.60	40,110.70	24.7
Netherlands	7,317.60	20,371	22.7
Sweden	6,818.90	29,978.40	34.5
Denmark	4,273.40	12,483	23.9
Norway	2,876.70	10,825.60	30.4
Portugal	1,888.50	9,823.40	39.1

Forecasted Growth of Mobile Payment Transactions in Chosen European Nations: 2020 Baseline and 2025 Projections

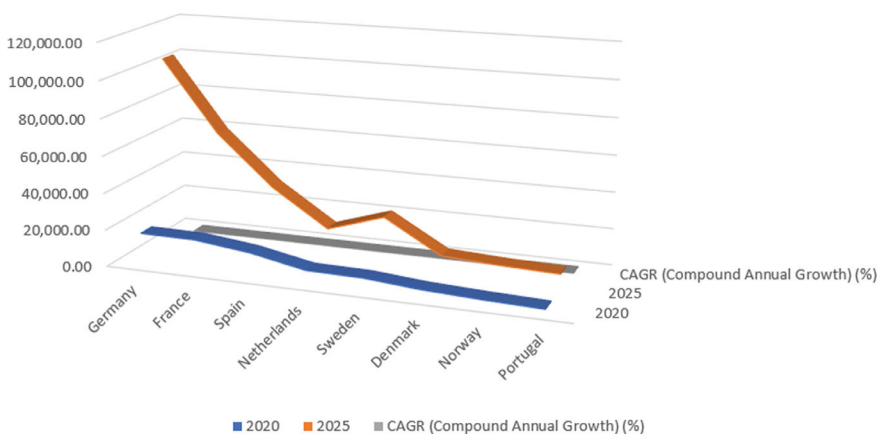


Fig. 15.1 Graph of forecasted growth of mobile payments transactions in Europe

Table 15.2 Romanian trends in mobile online financial transactions behavior: 2023 insights

Romania	2021 (%)	2023 (%)
Made a digital payment	47.2	56.1
Made a purchase using a mobile phone or the Internet	15.6	37.3
Used a mobile phone or the Internet for mobile banking	12.3	16.8
Used a mobile phone or the Internet to pay bills	12.0	34.1

As we see this growth in the use of IoT devices for banking in Europe, it is certainly about as high, or higher, globally. However, we come with a novelty, namely the use of these devices but in a blockchain network.

15.2 Types of Blockchain and the Impact of Blockchain in Finance

Further on, we will briefly see a definition of blockchain, the classification of blockchain networks, some information about the adoption of blockchain in making payments or other types of financial transactions directly with cryptocurrencies, as well as some of the security challenges at this level.

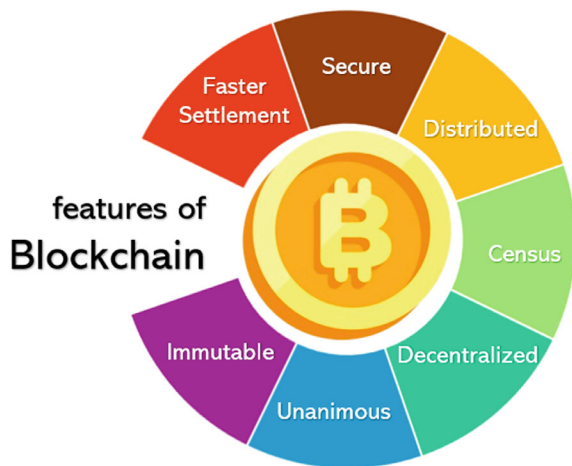
15.2.1 General Blockchain Definitions and Types of Blockchain

As a short general definition, we can say that the term “blockchain” refers to a paradigm to maintain information in a distributed system. However, we must make a distinction between the definition of blockchain and its characteristics, even if the definition of blockchain is not universally agreed upon, the basic properties of this technology are increasingly being explored [12].

We will next mention above some key features of blockchain (see Fig. 15.2).

As we can see in Fig. 15.2, some key features of blockchain are:

Fig. 15.2 Features of blockchain (Adapted from [12])



- **Immutable:** Immutability refers to the inherent quality of the blockchain as a network that is permanent and unchangeable [13].
- **Distributed.**
- **Decentralized:** Blockchain technology operates as a decentralized system, indicating the absence of a central authority governing the network. Rather, the network comprises numerous nodes collaborating to authenticate and validate transactions [13].
- **Secure:** Each record within the blockchain undergoes individual encryption, enhancing the overall security of the network. Encryption serves to fortify the security of the entire process within the blockchain network [13].
- **Consensus.**
- **Unanimous.**
- **Faster settlement.**

Now that we have seen some fundamental characteristics of blockchain, the following figure highlights the types of blockchain networks and some characteristics of each type (Public, Private, Consortium, and Hybrid):

- **Public Blockchain:** it is an open access network where anyone can participate in the validation of transactions and operates in a decentralized system. Also, transactions are transparent and accessible to all network participants. It offers a very high level of security thanks to the so-called Proof of Work (PoW [14]) mechanism;
- **Private Blockchain:** a restricted-access network, often utilized within organizations or consortia, operates on a centralized or semi-decentralized model with limited permissions, offering enhanced control over access, governance, privacy, and scalability compared to public blockchains;
- **Consortium Blockchain:** a collaborative network overseen by organizations with aligned interests, blending features of both public and private blockchains. Participants collectively manage the network, fostering consensus and governance, leading to enhanced efficiency and trust compared to conventional centralized systems.
- **Hybrid Blockchain:** integrates features from both public and private blockchains, providing flexibility in access control and transparency, enabling the integration of private and public networks to harness their individual advantages.

15.2.2 Blockchain Integration in IoT Financial Operations

The integration of blockchain technology in IoT finance represents a significant advancement in financial transactions, offering enhanced security and efficiency. Blockchain's decentralized ledger system ensures that financial transactions conducted through IoT devices are secure, transparent, and tamper-proof. Unlike traditional bank intermediaries, blockchain enables direct peer-to-peer transactions, eliminating the need for third-party involvement. This direct transfer of value between

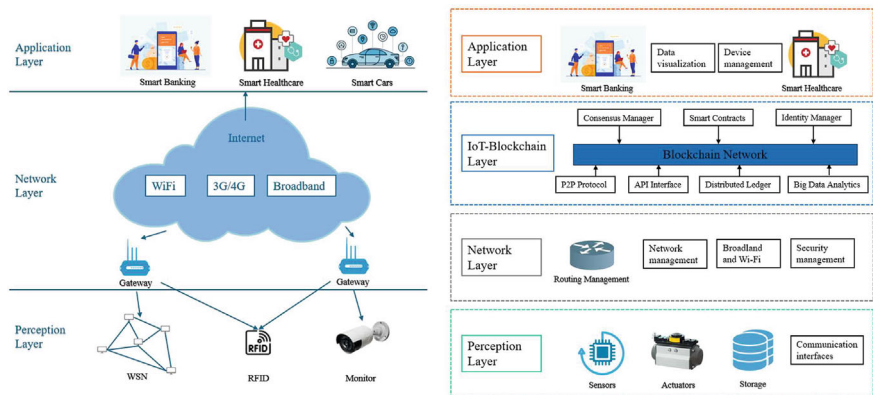


Fig. 15.3 Three-layer architecture of IoT comparison with IoT systems integrated with blockchain (Adapted from [15])

IoT devices, such as transferring money to a colleague, is inherently safer as it eliminates the risk of potential data breaches or unauthorized access associated with centralized banking systems.

With blockchain, transactions are cryptographically secured and recorded on a distributed ledger, making them immutable and resistant to manipulation (could be seen in Fig. 15.3). As a result, payments made through blockchain technology offer greater security and trust, making them an increasingly attractive option for financial transactions in the IoT ecosystem.

Through this comparison depicted in Fig. 15.3, between the simple IoT architecture and the IoT architecture integrating an additional blockchain layer, it can be observed that the blockchain layer is positioned at the transmission protocol level, ensuring information security. Here, the connected objects situated at the application level take on the role of generating transactions on the blockchain. Employing cryptographic mechanisms, each object possesses a private key used to authenticate transactions. The blockchain aims to validate these transactions, incorporating them into the chain associated with the object, thereby ensuring the ownership of the information tied to the connected object [15].

Nowadays, more and more users have already adopted the crypto trend and own cryptocurrencies with which they of course carry out financial operations. We saw briefly why we prefer to carry out financial transactions via blockchain, for example a transfer, but in this article, we will no longer insist on the details of the benefits and how exactly they are carried out, but we will focus further on the security issues that still exist and in this system.

15.3 Navigating Security Challenges in Blockchain Financial Operations

As we have seen in the previous descriptions, even if the blockchain networks are secure, there are still some security problems at the level of these networks and we will see some of them below. These security vulnerabilities.

- **Double spending:** In some blockchain systems, particularly those with weaker consensus processes, attackers can use flaws to spend the same bitcoin units' numerous times;
- **Smart Contract Security Vulnerabilities:** Vulnerabilities in smart contracts, which are self-contained code segments kept on the blockchain, may result in financial losses and exploits. For example, flaws or logical mistakes in smart contracts could allow attackers to take advantage of them to withdraw money or carry out illegal activities;
- **Sybil Attacks:** Sybil attacks [16] in decentralized networks entail the creation of numerous fictitious identities (Sybil nodes) in an effort to take over or exert influence over the network. These assaults have the ability to alter transactions, jeopardize security, and interfere with network functionality.
- **Insider Threats:** Those having special access to blockchain systems, like developers or administrators, run the risk of endangering the integrity of the system by deliberate or inadvertent means. Malicious acts, data breaches, or unapproved blockchain updates are examples of insider dangers;
- **51% Attacks:** It means that if one person or organization has more than 50% of the mining power on a proof-of-work blockchain network, they may be able to alter transactions, reverse transactions, or stop new transactions from being completed;
- **Routing Attacks:** By now, it should be clear to us that the operation of blockchain technology depends on a strong network. Through the BGP (Border Gateway Protocol), the ISPs are able to communicate with one another and exchange route data. Due to its age, this protocol contains certain vulnerabilities that could be used by an attacker [17]. An attacker in charge of an ISP, for example, might split the blockchain network in half or broadcast a fake route, which would prevent transactions for certain nodes.
- **Private Key Security Attacks:** Blockchain technology is based on public-key cryptography. As a result, poor public-key cryptography handling or implementation can result in some very major blockchain security problems. If your blockchain's key signing isn't done right—for example, by utilizing the same key for repeated signings rather than a Merkle tree—an attacker may be able to deduce your private key from the public key. Being in charge of your private key essentially entails being the owner of every piece of information linked to you on a blockchain. Holding all of your coins in the case of cryptocurrencies. That being said, the likelihood of this occurring is quite low unless your blockchain is built with incredibly flawed code. The misuse of your private key is the main problem [17].

- **Flash Loan Attacks:** These are attacks to steal assets from smart contracts intended to enable flash loans, which allow borrowers to borrow money without requiring collateral. These exploits use smart contract inputs to take advantage of uncollateralized loans [18].
- **Phishing Attacks:** These entail the use of social engineering tactics by hostile actors to get users' credentials, infect users' devices with malware, and obtain users' private keys and seed phrases—recovery phrases produced by cryptocurrency wallets during setup that allow users to access their wallets in the event that they misplace their device or forget their password [18].

We have listed as it can be seen, in general, some of the types of vulnerabilities when it comes to blockchain networks, following that in the next conducted studies we will also see how they are in detail and also how they can be prevented at the user level as well as on a larger scale wide.

15.4 Conclusion

Finally, integrating Internet of Things (IoT) devices with blockchain technology in financial transactions creates both intriguing prospects and serious obstacles. Throughout this article, we have depicted how IoT influenced financial sector and also the possible benefits of using IoT devices for financial transactions on blockchain networks, such as increased efficiency, transparency, and accessibility. However, it is critical to recognize the pre-existing security risks in this evolving ecosystem.

Moving forward, resolving these security concerns will necessitate a multifaceted approach combining collaboration among industry players, legislators, and cybersecurity specialists. Implementing strong encryption standards, improving authentication procedures, conducting frequent security audits, and increasing user knowledge and education are all critical steps toward mitigating risks and strengthening the resilience of IoT finance on blockchain networks.

Furthermore, continued research and development activities aimed at improving the security, scalability, and interoperability of blockchain technology will be critical in instilling trust and confidence in IoT-enabled financial transactions. By emphasizing security and implementing proactive tactics, the potential of IoT finance on blockchain networks to transform financial services while protecting user data and assets can be fully realized in the coming years.

References

1. Zhang, Z., Chen, M., Zhou, J.: A lightweight cache-assisted data recovery mechanism for latency-critical data in IoT. In: International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, pp. 443–448, IEEE. (2023). <https://doi.org/10.1109/WCSP58612.2023.10404447>
2. Onyeashie, B.I., Leimich, P., McKeown, S., Russell, G.: A bibliometric analysis and systematic review of a Blockchain-based chain of custody for digital evidence. In: Tan, Z., Wu, Y., Xu, M. (eds.) Big Data Technologies and Applications. BDTA. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 555. Springer, Cham. (2023). https://doi.org/10.1007/978-3-031-52265-9_8
3. Abad-Segura E, González-Zamar M-D.: Global research trends in financial transactions. *Mathematics*. **8**(4), 614 (2020). <https://doi.org/10.3390/math8040614>
4. Garis, R.L., Mead, E.S.: The business corporation: its financial organization and operation. *South. Econ. J.* **8**(1), 117–119 (1941)
5. Doing Business International, Traditional banking. <https://www.ng-business-international.com/traditional-banking>. Accessed (2024)
6. Suitsmecard, What are the main differences between digital banking & online banking. <https://suitsmecard.com/blog/what-are-the-main-differences-between-digital-banking-and-online-banking>. Accessed (2024)
7. Investopedia, What Is Online Banking? Definition and How It works. <https://www.investopedia.com/terms/o/onlinebanking.asp>. Accessed (2024)
8. Mckinsey, Unlocking the potential of the internet of things. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world>. Accessed (2024)
9. Statista, Transaction volume of mobile payments in selected countries in Europe in 2020 with a forecast for 2025. <https://www.statista.com/statistics/1270574/mobile-payment-transaction-forecast-europe/>. Accessed (2024)
10. Datareportal, Digital 2024: Romania. <https://datareportal.com/reports/digital-2024-romania>. Accessed (2024)
11. Datareportal, Digital 2022: Romania. <https://datareportal.com/reports/digital-2022-romania>. Accessed (2024)
12. Nicoletti, S.M., Peppelman, M., Kolb, C., Stoelinga, M.: Model-based joint analysis of safety and security: Survey and identification of gaps. *Computer Science Review* **50**. (2023). <https://doi.org/10.1016/j.cosrev.2023.100597>
13. Geeksforgeeks, Features of Blockchain. <https://www.geeksforgeeks.org/features-of-blockchain/>. Accessed (2024)
14. Investopedia, What Is Proof of Work (PoW) in Blockchain? <https://www.investopedia.com/terms/p/proof-work.asp>. Accessed (2024)
15. Abbasi, Y., Benlahmer, H.: IoT and Blockchain combined: for decentralized security. *Procedia Computer Science*, (2021), **191**, 337–342. In: the 2nd International Workshop on Artificial Intelligence & Internet of Things (A2IOT), Leuven, Belgium (2021). <https://doi.org/10.1016/j.procs>
16. Imperva, Sybil attack. <https://www.imperva.com/learn/application-security/sybil-attack/>. Accessed (2024)
17. Getastra, Blockchain Security issues – a complete guide. <https://www.getastra.com/blog/knowledge-base/blockchain-security-issues/>. Accessed (2024)
18. Techtarjet, Top blockchain attacks, hacks and security issues explained. <https://www.techtarjet.com/searchsecurity/tip/Top-blockchain-security-attacks-hacks-and-issues>. Accessed (2024)

Chapter 16

Models for Network Traffic Behavioral Analysis in IoT Systems



Alin Zamfiroiu , Justin Floroiu , Daniel Savu , Lorena Bătăgan ,
and Anupriya Sharma Ghai 

Abstract In today's interconnected world, the proliferation of Internet of Things (IoT) devices has revolutionized the way we interact with technology. From smart homes and wearable devices to industrial sensors and autonomous vehicles, IoT systems have permeated nearly every aspect of our daily lives. However, with this exponential growth in IoT adoption comes a host of challenges, particularly concerning the security and performance of these interconnected networks. One of the key methodologies employed to address these challenges is network traffic behavioral analysis. Moreover, network traffic behavioral analysis plays a crucial role in threat intelligence and incident response. By aggregating and analyzing network data over time, organizations can identify recurring attack patterns and indicators of compromise. This rich contextual information enhances their ability to anticipate and preempt future cyber threats. Additionally, in the event of a security incident, network traffic analysis provides invaluable forensic evidence, allowing investigators to reconstruct the sequence of events, attribute responsibility, and mitigate the impact of the breach. In this paper, we propose some models for the network traffic generated by sensors and actuators in IoT systems.

A. Zamfiroiu (✉) · I. Floroiu · D. Savu
National Institute for Research & Development in Informatics, Bucharest, România
e-mail: alin.zamfiroiu@csie.ase.ro

I. Floroiu
e-mail: iustin.floroiu@ici.ro

D. Savu
e-mail: daniel.savu@ici.ro

A. Zamfiroiu · L. Bătăgan
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: lorena.pocatilu@ie.ase.ro

A. S. Ghai
School of Computing, Graphic Era Hill University, Dehradun, India
e-mail: anupriya@gehu.ac.in

16.1 Introduction

At its core, network traffic behavioral analysis involves the systematic monitoring, analysis, and interpretation of data flows within an IoT network. By scrutinizing the patterns, trends, and anomalies present in network traffic, organizations can gain valuable insights into the behavior of IoT devices, users, and applications [1]. This proactive approach enables them to detect and mitigate security threats, optimize network performance, and ensure the reliability of IoT systems [2].

Security is undeniably one of the foremost concerns surrounding IoT deployments. With an ever-expanding attack surface and increasingly sophisticated cyber threats, safeguarding IoT networks against malicious activities is paramount. Network traffic behavioral analysis serves as a linchpin in the defense against such threats by providing continuous visibility into network activities. By monitoring traffic patterns in real time, organizations can detect and respond to anomalous behavior indicative of potential security breaches. Whether it's unusual data exfiltration attempts, unauthorized access attempts, or suspicious communication patterns, network traffic analysis enables security teams to swiftly identify and neutralize threats before they escalate into full-blown incidents [3].

Beyond security considerations, network traffic behavioral analysis also holds significant implications for optimizing the performance and efficiency of IoT systems. In complex and dynamic IoT environments, understanding the flow of data between devices, gateways, and cloud services is essential for ensuring smooth operation and timely delivery of services. By analyzing traffic patterns and identifying bottlenecks or latency issues, organizations can fine-tune their network infrastructure, allocate resources more effectively, and enhance the overall performance of IoT applications. Furthermore, insights gleaned from traffic analysis can inform capacity planning, resource allocation, and network optimization strategies, enabling organizations to scale their IoT deployments to meet growing demand and evolving business requirements.

One of the distinctive challenges of network traffic analysis in IoT systems lies in the sheer diversity and heterogeneity of devices, protocols, and communication patterns involved. Unlike traditional IT networks characterized by standardized architectures and protocols, IoT ecosystems encompass a myriad of interconnected devices with varying capabilities, interfaces, and communication protocols. This complexity necessitates sophisticated analytical techniques capable of handling diverse data formats, extracting meaningful insights, and adapting to dynamic network environments [4].

Machine learning and artificial intelligence (AI) have emerged as powerful tools in the arsenal of network traffic analysis, enabling organizations to tackle the intricacies of IoT data at scale. By leveraging advanced algorithms and statistical models, machine learning algorithms can discern subtle patterns and anomalies within vast volumes of network traffic data. Supervised learning techniques can be trained on labeled datasets to classify normal and malicious behavior, while unsupervised learning algorithms can detect outliers and deviations from established norms.

Furthermore, reinforcement learning approaches enable adaptive decision-making in response to evolving threats and changing network conditions, enhancing the resilience and effectiveness of network security measures.

16.2 IoT Systems in Smart Homes

In the last decades, technology has rapidly evolved and brought with it a series of innovations that have completely changed the way people carry out their activities and interact with the environment. More than ever the present is marked by the impressive technological advances of recent years, with every aspect of everyday life changing at an amazing pace. Even our homes are changing more and more with the help of today's extraordinary technology.

Transforming home into a smart ecosystem (smart home) has become a necessity for those who want a comfortable and energy-efficient lifestyle [5]. The "smart home" concept involves the integration of cutting-edge technologies that allow residents to remotely control various systems and devices, such as lighting, heating, air conditioning, security systems, and electronics at home [6]. The Smart Homes Association defines a smart home as "the integration of technology and services through home networking for a better quality of living" [7].

Driven by technological advancements, the smart home market has been rapidly evolving across the globe. In 2023, the global smart home market was valued at USD 102.3 billion and is expected to grow to USD 263.6 billion by 2032, at a compound annual growth rate (CAGR) of 10.0% during the projected period [8].

Smart home technology is based on the Internet of Things (IoT) concept and allows users to manage and control systems and devices through a single device, such as a smartphone, tablet, or laptop, or even through voice commands and personal assistants (such as Alexa, Google Assistant, and Siri). IoT creates a connected ecosystem where devices work together to create an intelligent living environment. The IoT smart home is a concept that, in addition to automation and remote control, benefits from two extremely important attributes: interconnection and contextual adaptation [9].

The concept of the (IoT) refers to all devices, phones, computers, gadgets, and household appliances in the home, that are interconnected in a single network, gather information from around, and exchange data. IoT applications allow devices to connect to each other and communicate without any human action.

The most commonly used IoT technologies that enable connectivity and automation in smart homes are smart sensors, voice assistants, smart appliances, security systems, and energy management.

There is a variety of IoT devices that can be used to create smart home systems. Smart home devices are designed to automate and simplify various aspects of everyday life. Among the most commonly used devices of this type are the following [6, 10]:

- **Smart thermostats.** Devices such as Nest and Ecobee allow for remote control of ambient temperature and have the ability to learn the user's temperature preferences and adjust heating and cooling systems to save energy, thereby reducing energy consumption and lowering costs.
- **Smart security cameras.** Allow remote home monitoring and send alerts to your smartphone if they detect any unusual activity.
- **Smart doors lock.** Allow doors to be locked or unlocked remotely, eliminating the need to use a physical key.
- **Smart lighting.** Bulbs, LEDs, and smart lamps can be controlled remotely to adjust their intensity, thus creating a relaxing environment. Also, they contribute to saving energy and reducing costs.
- **Smart sockets.** Can be programmed to turn on or off the power supply to devices. The smart sockets are equipped with sensors that automatically turn off when they detect high temperatures. They also include safety systems to protect children and pets.
- **Smart kitchen appliances.** IoT technology for the kitchen includes smart refrigerators, ovens, and coffee machines. These devices can be remotely controlled and provide information on food supplies, recipes, and cooking times.
- **Health monitoring devices.** Provide real-time health data, encouraging a healthier lifestyle.
- **Entertainment and streaming devices.** Smart TVs, game consoles, and streaming devices include built-in IoT capabilities, allowing users to access content and control their entertainment systems through voice commands or mobile apps.
- **Smart window blinds.** Light sensors on the outside of the blinds will automatically open them in the morning to provide natural sunlight. These blinds can be opened or closed remotely. Thermal blinds insulate the windows, thus contributing to ensuring an optimal ambient temperature.
- **Smoke and flood detection systems.** In cases of smoke detection, the sensors stop the consumers and trigger the fire protection system. The flood detection sensors can determine the unauthorized consumption of water if the taps are closed and can interpret it as a pipe break and flood hazard.
- **Lawn and garden irrigation systems.** The sprinkler system for watering the lawn can be programmed according to the temperature over the day and the existing evaporation. For the garden, there are programmable drip systems.

Some of these devices are presented in Fig. 16.1. IoT smart home technologies offer a number of important benefits, such as [11, 12]:

- **Control and monitoring.** IoT significantly improves the way events and processes that take place at home can be controlled and monitored. Thus, the IoT application collects day-to-day data about how a household works, processes it, and presents the most relevant insights.
- **Cost and energy savings.** Benefiting from detailed statistics on the energy consumption of each smart device, the user can optimize and adjust the settings

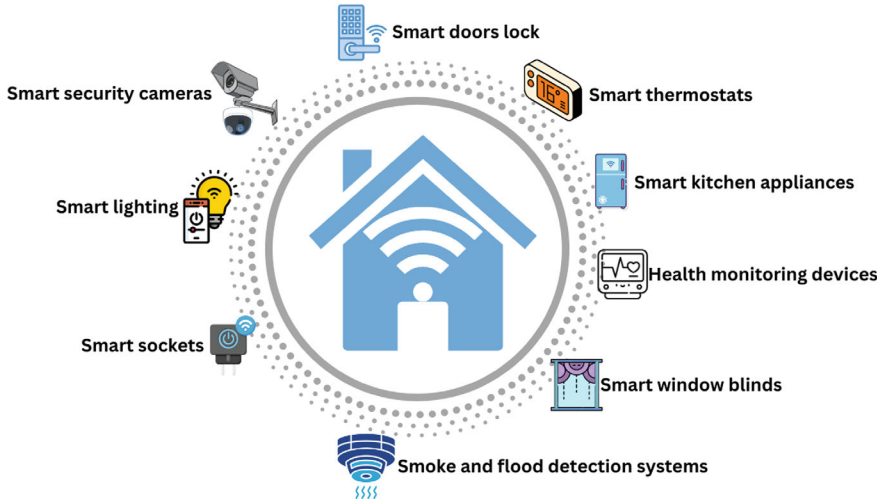


Fig. 16.1 IoT devices in a smart home

of the smart home in a more economical way. The smart home system can be allowed to do everything by itself. It will turn off devices that are not in use and reduce power consumption if no one is at home.

- Impact on the environment. Energy saving also ensures, in addition to reducing costs, reducing the negative impact on the environment.
- **Improved security.** The smart home security system provides total control over the entire home, both inside and outside. The smart devices that make up this system allow the user to check what the situation is, regardless of where they are.
- **Comfort.** IoT for smart homes simplifies and improves everyday life. The user is always in control, even though the processes in the home run autonomously.

In order for smart home devices to integrate into a smart home ecosystem, it is necessary to use an appropriate protocol. A smart home protocol is a set of standards that determine how devices within a smart home ecosystem communicate. Among the most used smart home protocols are Z-Wave, ZigBee, Wi-Fi, Bluetooth Low Energy (BLE), Ethernet, Thread, and Matter [13, 14].

Creating a smart home is a complex process that requires knowledge of available technologies and how they can be integrated to provide comfort, efficiency, and security. With proper arrangement and the use of suitable smart devices, the house can be transformed into a modern and connected space that offers total control over the living environment.

16.3 Behavioral Analysis of Network Traffic

Amar et al. 2018 [15] address the increasing presence of Internet-connected devices in homes and the associated rise in privacy breaches, data thefts, and security threats. The authors analyze network traces from a testbed of common IoT devices and describe methods for fingerprinting their behavior. They assess privacy and security risks and how device behavior affects bandwidth. The paper discusses simple measures that can be implemented to circumvent attempts at securing devices and protecting privacy. The study suggests that understanding the in-the-wild behavior of IoT devices is crucial for containing them and mitigating associated risks.

In intelligent living and buildings, the network sizes are typically small to medium. The number of sensor devices in an ordinary living area usually does not exceed 1,000 [16]. Data transmission rates for these sensor devices are infrequent and irregular. Most of these devices (such as the Indoor Ambiance Monitoring Sensor) send data only when an abnormality occurs. Emergency data from these sensors require real-time processing. Sensor devices in intelligent living/buildings may use various energy sources like batteries, fixed power supplies, and self-charging solar equipment.

Energy efficiency is crucial for network communications in these devices. The integration of advanced IoT technologies into healthcare systems, as outlined in the discourse on smart health, holds promise for revolutionizing medical services. This synthesis, coupled with the efficient utilization of hospital applications, not only enhances the quality of healthcare but also streamlines healthcare systems. The prospect of a unified barcode granting access to a spectrum of services, from identity management to mobile drug administration, exemplifies the potential efficiency gains [16].

At the forefront of energy management, smart grids aim to revolutionize utility usage by automating data collection and enabling dynamic adjustments. These systems, spanning from small-scale setups to extensive networks, rely on fixed power supplies to transmit data at relatively low rates. While regular traffic loads dominate, interruptions like power outages introduce variability, demanding nearly real-time data delivery for anomaly detection. Security and data integrity are paramount, outweighing energy efficiency concerns given the fixed power supplies [16].

The exponential growth of smart home IoT devices has ushered in a new era of convenience and connectivity, but it also presents a host of challenges, particularly in the realms of management, security, and privacy. According to [17], the number of IoT users of smart homes will increase by 117.69% by 2028, compared with 2023.

As IoT devices become ubiquitous in households worldwide, understanding their behavior and addressing associated concerns are imperative for researchers and operators alike. The heterogeneous nature of smart home IoT devices, coupled with their lack of standardized interfaces, poses a significant challenge in comprehensively studying their behavior. Studying IoT devices in controlled environments may not accurately capture their behavior in real-world settings, where environmental factors influence their performance and interactions. The sheer scale and diversity of the

smart home IoT market make it challenging for researchers to gain insights applicable to the broader ecosystem [18].

The functionality of IoT devices dictates their communication patterns, with media-related functionalities generating substantial traffic. Understanding these patterns enables better network management and provisioning. Despite the fragmentation on the front-end, smart home IoT back-ends are increasingly centralized, predominantly hosted on major cloud providers. This centralized infrastructure influences DNS resolution patterns, with a significant proportion of devices relying on hard-coded DNS servers. The lack of traffic encryption and reliance on HTTP communication pose serious privacy risks, leaving smart home IoT devices vulnerable to eavesdropping and manipulation. Additionally, communication with third-party advertising and tracking services raises concerns about user data privacy and security [18].

Smart home IoT devices communicate using a diverse array of protocols tailored to specific functionalities and requirements. These include HTTP/HTTPS for web-based communication, MQTT for lightweight messaging, DNS for domain resolution, and UPnP for device discovery and control. Each protocol serves a distinct purpose, facilitating efficient communication between devices and the broader network infrastructure. At the physical layer, smart home IoT devices utilize a variety of network technologies to establish local and internet connectivity. Zigbee, Z-Wave, and Bluetooth Low Energy (BLE) are commonly employed for short-range communication within the home environment, offering low power consumption and robust connectivity. Wi-Fi and Ethernet serve as the primary means for internet connectivity, enabling seamless access to cloud-based services and remote management capabilities.

16.4 Proposed Smart Home System Architecture for Behavioral Analysis of IoT Systems

In a Smart home system, all devices are connected to the same router and for this router the network traffic is checked by means of the software tool: Wireshark. This tool allows the capture and interactive analysis of network traffic in real time. Normally it is used for troubleshooting the network, we used it for gathering information about the packets transmitted by the devices installed in a Smart Home system. The systems are installed separately, so they are not from the same manufacturer and thus will have different behavior.

The analysis was carried out for 20 min. This analysis is demonstrative and to create faithful profiles for the installed sensors and systems, more analyses are needed. But for this material, we want to present how to create behavioral analysis models for them.

In this analysis, a total number of 17,580 packets were identified for the entire network. The following devices are installed in the network:

- three smart TVs;
- four smart window blinds;
- a cleaning robot;
- an intelligent visor;
- two computers;
- a washing machine;
- an intelligent thermostat;
- smart air conditioning.

These systems are shown in Fig. 16.2. The two computers are connected to the same network and for this reason in the analysis carried out certain packets were generated by these computers, even if they are not part of the Smart home system. The smart home environment is depicted in Fig. 16.2.

Based on traffic analysis, the characteristics that must be included in the recognition model of devices connected to the network were identified. Table 16.1 presents the used characteristics.

These characteristics are considered for each sensor or device that is connected to the network. In order to create the profile, the information for each device or sensor is saved, so that on future connections to the Internet or occasionally, the new measurements are compared with the saved ones. If the new measurements are different, it can be deduced that the respective sensor or system has been attacked or contains malicious code that changes its behavior.

For the created analysis that led to the identification of the characteristics, a comparative analysis was made for a set of blinds (Smart Windows Blinds) and

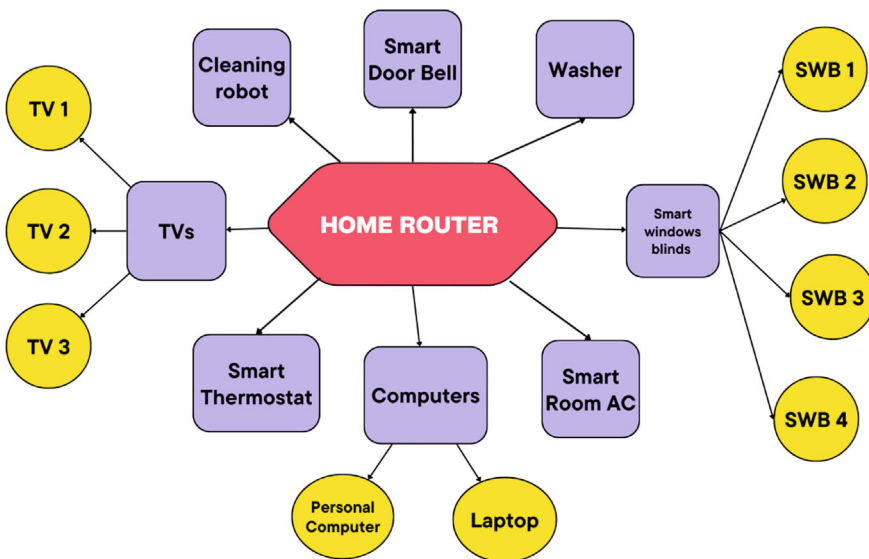


Fig. 16.2 Smart home environment

Table 16.1 Characteristics for traffic analysis

No	Characteristic	Description
1	Data length	Total length of useful data (without overheads) (bytes)
2	Frame length	Total length of the frame (bytes)
3	Capture length	Total length of the captured data (bytes)
4	Arrival time interval	The difference between the arrival time of a packet and the arrival time of the previous packet (seconds)
5	Used protocol	The analyzed network protocol (Ethernet, TCP, UDP, HTTP)
6	Source IP	The IP address of the source device
7	Destination IP	The IP address of the destination device
8	Source port	The port number used by the source device
9	Destination port	The port number used by the destination device

for the cleaning robot. The brief analysis can be seen in Table 16.2. For each of the two devices, 257 packets were analyzed in the analysis interval. The values for each individual device were similar for all 257 packages for all the characteristics analyzed, Table 16.1.

In future analyses, the measurements made will be compared with these and thus it can be decided whether any system has been attacked or not. It is also possible to determine the moment when a certain device changes its behavior.

Table 16.2 Comparative analysis between two systems

No	Characteristic	Smart windows blinds	Cleaning robot
1	Data length (bytes)	172	79
2	Used Protocol	UDP	UDP
3	Source IP	((hidden))	((hidden))
4	Destination IP	((hidden))	((hidden))
5	Frame Length (bytes)	214	121
6	Capture length (bytes)	214	121
7	Arrival time interval (seconds)	5	5
8	Source port	49,153	44,080
9	Destination port	6667	58,866

16.5 Conclusions and Future Work

Network traffic behavioral analysis represents a cornerstone in the quest to secure and optimize IoT systems. By scrutinizing the digital footprints left by IoT devices, applications, and users, organizations can fortify their defenses against cyber threats, enhance the performance and reliability of their networks, and unlock new opportunities for innovation and growth in the IoT landscape. As IoT adoption continues to soar and the digital ecosystem becomes increasingly interconnected, the role of network traffic analysis will only grow in importance, serving as a beacon of insight and foresight in the ever-expanding realm of connected devices and smart technologies. The next steps will focus on model development, data gathering, and model validation.

Acknowledgement This work was carried out through the Core Program within the National Research Development and Innovation Plan 2022-2027, carried out with the support of MCID, project no. 23 38 01 01, “Contributions to the consolidation of emerging technologies specific to the Internet of Things and complex systems.”

References

1. Boja, C.: Methodology for evaluating IoT security based on devices behavior. In: Proceedings of the IE 2020 International Conference (2020)
2. Aphorpe, N., Huang, D.Y., Reisman, D., Narayanan, A., Feamster, N.: Keeping the smart home private with smart (er) IoT traffic shaping (2018). arXiv preprint [arXiv:1812.00955](https://arxiv.org/abs/1812.00955)
3. Wan, Y., Xu, K., Wang, F., Xue, G.: Characterizing and mining traffic patterns of IoT devices in edge networks. *IEEE Trans. Netw. Sci. Eng.* **8**(1), 89–101 (2020)
4. Wan, Y., Xu, K., Wang, F., Xue, G.: IoT Athena: unveiling IoT device activities from network traffic. *IEEE Trans. Wireless Commun.* **21**(1), 651–664 (2021)
5. Yasar, K., Shea, S.: What is a Smart Home? Everything You Need to Know (2023). <https://www.techtarget.com/iotagenda/definition/smart-home-or-building>
6. Chakraborty, A., Islam, M., Shahriyar, F., Islam, S., Zaman, H.U., Hasan, M.: Smart home system: a comprehensive review. *J. Electr. Comput. Eng.* **2023**, 1–30 (2023)
7. El-Azab, R.: Smart homes: potentials and challenges. *Clean Energy* **5**(2), 302–315 (2021)
8. Smart Home Market (By Application: New Construction, Retrofit; By Product: Light Control, Security And Access Control, HVAC Control, Entertainment Control, Smart Speaker, Home Healthcare, Smart Kitchen, Home Appliances, Home Furniture; By Sales Channel: Direct, Indirect; By Protocol, Wired, Wireless, Hybrid) - Global Industry Analysis, Size, Share, Growth, Trends, Regional Outlook, and Forecast 2023–2032. <https://www.precedenceresearch.com/smart-home-market>
9. Yang, H., Lee, W., Lee, H.: IoT smart home adoption: the importance of proper level automation. *J. Sens.* (2018)
10. Ghafurian, M., Ellard, C., Dautenhahn, K.: An investigation into the use of smart home devices, user preferences, and impact during COVID-19. *Comput. Hum. Behav. Rep.* **11**, 100300 (2023)
11. Nadeem, M.: Transforming your home into a smart home heaven: modern smart home (2023). <https://www.linkedin.com/pulse/transforming-your-home-smart-heaven-modern-written-muhammad-nadeem/>

12. IoT.Business.News. The Smart Home Revolution: How IoT Is Transforming Modern Living (2023). <https://iotbusinessnews.com/2023/11/10/89810-the-smart-home-revolution-how-iot-is-transforming-modern-living/>
13. Diachenko, R.: Top 7 Smart Home Protocols Compared. (2023). <https://lebergolutions.com/blog/smart-home-protocols-explained>
14. Ribeiro Serrenho, T., Bertoldi, P.: Smart home and appliances: State of the art, EUR 29750 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-03657-9. <https://doi.org/10.2760/453301>. JRC113988. (2019)
15. Amar, Y., Haddadi, H., Mortier, R., Brown, A., Colley, J., Crabtree, A. An analysis of home IoT network traffic and behavior (2018). arXiv preprint [arXiv:1803.05368](https://arxiv.org/abs/1803.05368)
16. Li, Y., Tu, W.: Traffic modelling for IoT networks: a survey. In: Proceedings of the 10th International Conference on Information Communication and Management, pp. 4–9 (2020)
17. Statista. Penetration rate of the smart homes market worldwide from 2019 to 2028 [Graph]. In Statista. Retrieved 18 Apr 2024. <https://www.statista.com/forecasts/887636/penetration-rate-of-smart-homes-in-the-world>
18. Mazhar, M.H., Shafiq, Z.: Characterizing smart home iot traffic in the wild. In: 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), pp. 203–215, IEEE. (2020)

Chapter 17

Augmented Reality in Education—Prototype for the Undergraduate System in Romania



Andreea-Cristina Stroe

Abstract Education 4.0, as the new age of education, offers a variety of technologies that can be used in a blended learning environment. One of the possibilities that has gained more and more visibility in recent years is unquestionably augmented reality. Even though Romania is still in its infancy regarding the technologization of the teaching act, the community proves to be more and more open to new initiatives. Thus, this paper has the goal to provide an easy-to-be-implemented prototype of AR application that could be used as an adjuvant to a K-12 lesson. Firstly, this study will briefly give an introduction on what augmented reality means, its evolution throughout time, as well as some advantages and disadvantages. Secondly, the paper will present an easy and cost-free way to develop a simple prototype built on the principles of augmented reality. Finally, the paper will provide an example of AR application that can be suitable for the undergraduate education system, in the field of foreign languages. Overall, the contribution of this paper to the literature consists in revealing a simple and cost-free way of how augmented reality can support the teaching act in Romania. Moreover, it offers more visibility to the subject of augmented reality in the context of K-12 education.

17.1 Introduction

The digitalization of the Romanian education has been a priority both for the educators and the authorities for a very long time. Education 4.0 has been a familiar term for all the actors in the educational field. Thus, the purpose of Romanian education has become to prepare the new “digital natives” for the labor market [1]. A proper way to adapt the education to the new requirements was to ensure new usages of

A.-C. Stroe (✉)
University of Economic Studies, Bucharest, Romania
e-mail: andreea.stroe@csie.ase.ro

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
C. Ciurea et al. (eds.), *Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)*, Smart Innovation, Systems and Technologies 426,
https://doi.org/10.1007/978-981-96-0161-5_17

191

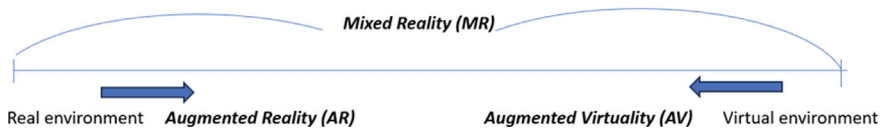


Fig. 17.1 Representation of the reality-virtuality (RV) continuum, re-drawn from Milgram et al.

technology that would contour a more engaging and adaptive way of learning. So, new ways of using the current reality have been explored.

Through technology and with the use of multimedia resources, people managed to enhance one's perceptions of reality [2]. In this way, the concept of mediated reality has been brought into attention once again. This is not a new concept, as Steve Mann defined it in 1999 as being a way to "implement a visual filter" [3]. But it gained more and more popularity in recent years when new "virtual filters" have emerged with the development of computers and mobile devices. New nuances of mediated reality have appeared and they can be easily categorized based on the taxonomy Milgram et al. developed in 1994 [4, 5]. Therefore, terms such as mixed reality, augmented reality, augmented virtuality, or virtual reality are more and more used by the specialists in education. Figure 17.1 below shows a brief representation of the taxonomy.

This taxonomy presents the most common two nuances that emerged in the educational act: the virtual reality and the augmented reality. Virtual reality can briefly be reduced to the simple definition of "human immersion in a synthetic world" [6]. Thus, VR consists in creating a completely new artificial world that will replace the reality [7], which the user will emerge in through a headset [8]. Whereas augmented reality is positioned in the other part of the spectrum, more closely to the real environment. An inclusive definition of the concept of AR is that it represents "a wide spectrum of technologies that project computer-generated materials, such as text, images, and video, onto users' perceptions of the real world" [9]. So, the user will not leave the real world, but it will be enriched with virtual objects adequate to the purpose of the model [10]. An option that has been investigated is also the Mobile Augmented Reality (Mobile AR), that represents a version of augmented reality with additional challenges brought by the device it is used on, the mobile phone in the present case [2].

The focus of the present paper will be exploring augmented reality. The paper will be structured in multiple sections. The first section has the goal to present the way augmented reality was introduced in Romanian society, with a particular focus on the educational field. The second section will present a short taxonomy of augmented reality, as well as some types of applications this concept was integrated in. The third section will focus on presenting the steps that were undertaken in order to build a prototype that can assist a French lesson for the undergraduate education system. The paper will conclude with presenting the advantages and challenges of the prototype, attempting to determine whether this prototype can be feasible in the current curriculum of the Romanian undergraduate education system.

17.2 Augmented Reality in Romanian Education

The tendency toward digitalization, especially towards mobile devices technology, has sustained, in recent years, a direction of the Romanian Education for exploring new technologies that can assist the educational act [11]. One of the new trends that were explored was unquestionably augmented reality. Over the last decade, the interest in this new era has grown and actions towards understanding how it works and what benefits it can bring to the Romanian education were taken. This tendency was observed both in the higher education and in the undergraduate education system. The Romanian researchers have come up with different prototypes for different levels of education, but they also wanted to understand, via the way the teachers and students perceived this trend, if this was a viable direction for our system. Nonetheless, at the moment of writing this paper, the integration of augmented reality in the Romanian education system remains only as a prototype, since no projects from the ministry were found.

A study was undertaken in 2018 in order to prove if augmented reality can enhance higher education experience [12]. So, the researchers concluded that the interest for AR among higher education students is significant, especially in the areas of engineering, management and humanities [12]. The study also showed that the majority of the participants (61%) claimed that the Ministry of Education should take into consideration recognizing such technologies as part of the teaching process [12].

In the sphere of undergraduate education system, in the last decade, multiple prototypes have been proposed. Nevertheless, there is no data recorded that they were actually implemented in class more than just for research, to collect some relevant data about the perception among students. The researchers explored different fields (such as physics [13], chemistry [14]), but also different types of AR applications.

An example of a developed prototype is an AR book created with the aid of ZooBurst [13]. Researchers have come up with an interactive book containing 3D pop-ups the students can interact with. The field in which the application can be used is the science, since it can help students learn about the solar system [13]. Whereas this book clearly enhances students' motivation, it also helps in preparing the new era of digital nomads for the necessary skills required in the new digital labor market.

A second prototype that was investigated by Romanian researchers is ARChemistry Learning [14]. It represents an application built with the aid of Vuforia and Unity, which contains several modules. The distinctive aspect of this application is that it can be used both in the teaching part of the process, when students acquire knowledge about the chemical substances, and in the building-up part, once they are already familiar with different concepts and they want to strengthen their knowledge. The experience was proven to be more pleasant both for students and teachers as the collaboration between them was improved. Moreover, students perceive this experience as less stressful [14].

17.3 Research Methodology

17.3.1 *Topology of Augmented Reality Approaches*

Creating a classification of augmented reality applications have been a preoccupation of the research community. Based on what capabilities the applications explore, they can be divided into three categories: emphasizing roles, emphasizing locations, and emphasizing tasks [15]. Nonetheless, this categorization is not extremely strict and possible overlaps can occur between the categories [16].

When considering applications with an accent on roles, students are, in fact, assigned different roles from the AR ecosystem, which is a dynamic environment [17]. Thus, the students are, in fact, matched to the characters of a game that needs to collaborate, to interact with each other in order to benefit from the learning experience. These applications have three directions the students can be included into: participatory simulations, role playing and jigsaw approach. Whereas in the first two, each student can simulate individually a role, in the jigsaw approach, each student receives a specific role in the environment and, in order to complete the learning experience, they need to unite every part of every role each student plays in the whole puzzle [15].

The applications that emphasize on locations take advantage of the physical capabilities of the devices that mediate the AR experience, such as the geolocation. There are two subtypes of applications, place-based and location-based [15] and both of them assume that the student which is engaged in an AR experience will explore the immediate physical ecosystem.

When the emphasis is on tasks, thus on what the student is supposed to do in the AR ecosystem so that the learning experience to happen [15], multiples subtypes of applications are possible: game-based, problem-based and studio-based [15]. Game-based applications can easily be considered as different games situated in the real world where the students' experience is mediated by electronic devices [18]. The problem-based applications expect students to gain knowledge by finding a solution to a problem through an AR experience [15]. The most interesting type of applications are undoubtedly the studio-based ones because they assume the students being the creators of their own AR applications [15].

Another topology of the augmented reality application is based on the mechanism used for displaying the 3D content. So, applications can be marker-based or markerless-based [19]. Marker-based applications, as their name suggests, needs a marker that is processed in order for the 3D content to be displayed on the screen. The whole process of transforming the marker into AR content needs three stages or elements: a booklet (that has the role to obtain the information of the marker), the gripper (very similar to a converter, taking the information the booklet obtains and convert it to other types of data) and the cube (that transforms the data into 3D content) [20]. Due to the difficulty of this mechanism, the marker-based applications are not very widespread. Meanwhile, markerless-based applications utilize a particular tracking system (the most often ones are the GPS or the compass) in order to

guide the user through the AR experience [20]. Due to its easiness for implementation (they are based on the capabilities of the device and no additional content is needed in order to trigger the AR application), the markerless-based ones have known more applicability and more interest manifested by the scientific community.

In the next section, a prototype of a marker-based application is discussed. The prototype can also be considered an application that emphasizes on the role of a game-based task-oriented application.

17.3.2 Building an Augmented Reality-Based Prototype

This section will focus on presenting a prototype of an augmented reality project that can be used as an adjuvant for a lesson in the undergraduate education system. It will be based on a marker-based approach and the target group will be students in primary school or any child with an A1 French level. It will cover the teaching part of the learning process and will integrate free resources a teacher with no programming experience can use in order to create an augmented reality experience. The main purpose of this section is to prove that each teacher with a minimum set of digital skills can create their own augmented reality-based application to use in any subject, free of charge.

The chosen subject will be French. This choice is based primarily on the number of pupils that are learning French in Romania as it is proven that French is one of the second most studied foreign language in the Romanian undergraduate education system. It was reported that there are approximately 240 thousand pupils who were learning French as their first foreign language, whereas over 1 million pupils were learning it as their secondary foreign language [21]. Moreover, the students are quite reluctant regarding acquisition of vocabulary in a foreign language since the majority of teachers use the traditional, non-interactive approach. Thus, an augmented reality adjuvant can enrich their experience by making it more interactive and motivating.

The topic will be *My house*. It will cover acquiring basic vocabulary about the main rooms of a house (the entryway, the living room, the kitchen, the bedroom, and the bathroom) and some common objects from each room. It will simulate a gamification experience, as the screens could be compared to the levels of a game. The pupil will need to navigate the screens one by one, as if they are visiting a real house. The pupil, through 3D images of every room, will be able to interact with the objects. Some objects from the rooms are labeled with their reference and, at one's click, the pronunciation of the word in French will be available. Thus, the application allows children not only to acquire the reference of the word but also to get in touch with a real French pronunciation. In this way, the learning experience will combine multiple senses (sense of touch, sense of hearing, sense of seeing) and will become more valuable and more interactive. The progression through screens will be done via button click and it will not be restricted by listening to all available words from a previous screen.

The development of the augmented reality application will not need any programming experience and it will be mediated by the designer that one of the products of Zappar (<https://www.zappar.com/>) offers. It is called Zapworks and it is an easy tool, the content being drag-and-dropped. It offers a free trial of two weeks and some available 3D content. Nonetheless, the free content available in the designer is not so offering, but rather limited. So, the author of the application should create their own content.

In order to create the custom-made content, a few free resources can be easily accessed. For obtaining the 3D content used for the screens, two sites can be used: poly.pizza (<https://poly.pizza/explore>) and Sketchfab (<https://sketchfab.com/feed>). Poly.pizza is the recreated version of Poly, the website created by Google that was shut down in 2021 [22], and it offers free downloadable 3D objects. An account can be created to access the content, but it is not mandatory since there is free content available without any account. Sketchfab is another site that offers free 3D content, but the majority of its content is not free of charge. It also offers two yearly plans, for better quality and more options. For obtaining real French pronunciation of the words, Forvo (<https://forvo.com/>) and YouTube (<https://www.youtube.com/>) can be used. Forvo offers various types of pronunciations, from different French-speaking countries, so the content can be customized according to different needs.

The prototype is called *Ma maison*, and it is a marker-based AR application. After the screens are created in the designer and the application is published, a QR-code will be generated. Once a student is provided the QR-code, the application will be opened in a mobile browser. So, the experience created with Zapworks is actually a WebAR experience.


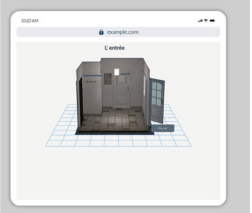
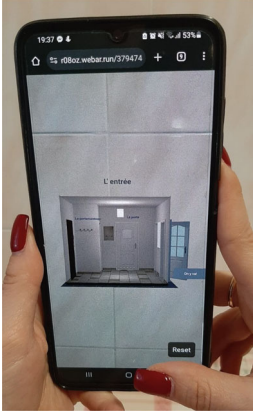
Table 17.1 provides a brief overview of all the screens of the application. The first column of the table contains a view from the designer. It is available for every screen of the application. The second column represents the view in an actual test usage of the application. It is available only for the screens that represent actual rooms of the house (screens 2–6). This decision was based on the fact that these are the screens that includes 3D content and the intention was to show how it is displayed in an actual usage. The last column includes explanations that accompany each screen (for instance, what the available options for the user are and what the significance of the screen is).

In the next section of this paper, the strengths and challenges of developing such a prototype will be covered.

17.4 Discussions

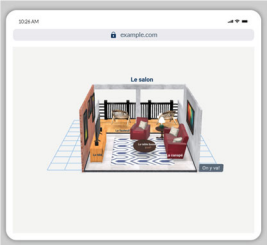
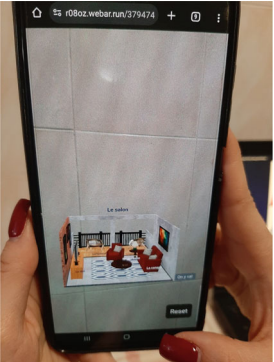

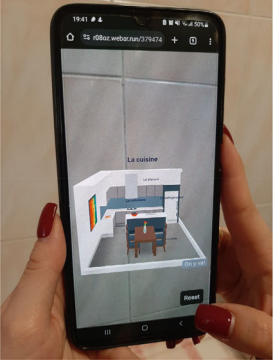
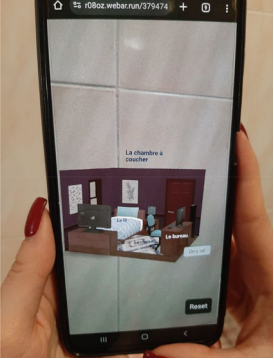
The prototype presented in the previous section is a very efficient tool that can be integrated in the educational act. Its main characteristic is the easiness, both to develop it and to use it. In order to develop it, one needs only some basic digital skills. For usage, the needed infrastructure is minimal, only a device that can scan a

Table 17.1 Description of screens of the prototype

No.	View in designer	View in actual usage	Explanations
1.		<p>Not necessary/no impact</p>	<p>After utilizing the provided QR-code, the students need to start the AR experience. Since the application is a marker-based application, the first screen of the application represents the trigger (the image of a house). The image will be provided by the teacher (either by projecting it on a board or by printing it and distributing to each pupil). Once the experience starts, the pupil will listen to an instruction in French that will mimic an invitation to enter the house. On a click on the door, the pupil will be redirected to the second screen.</p>
2.			<p>The second screen is a 3D representation of an entryway. The user can interact with the content, making it bigger or smaller, rotating it so that all the objects can be visualized. Moreover, the user can listen to the pronunciation of two main objects from the room (the door and the coat stand). At button click, the user will be redirected to the next screen.</p>

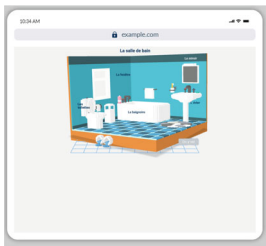
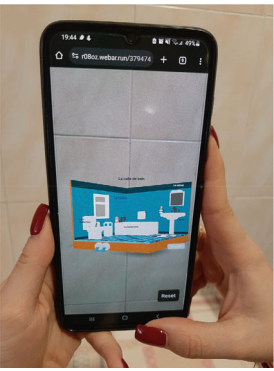
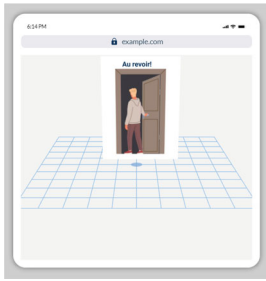
(continued)

Table 17.1 (continued)

No.	View in designer	View in actual usage	Explanations
3.			<p>The next screen is a 3D representation of a living room. The user can again interact with the content at their convenience. Moreover, the user can listen to the pronunciation of four main objects from the room (the television, the couch, the armchair, and the coffee table). At button click, the user will be redirected to the next screen.</p>
4.			<p>The next screen is a 3D representation of a kitchen. The user can again interact with the content at their convenience. Moreover, the user can listen to the pronunciation of four main objects from the room (the cupboard, the table, the fridge, and the stove). At button click, the user will be redirected to the next screen.</p>
5.			<p>The next screen is a 3D representation of a bedroom. The user can again interact with the content at their convenience. Moreover, the user can listen to the pronunciation of three main objects from the room (the bed, the chair, and the writing desk). At button click, the user will be redirected to the next screen.</p>

(continued)

Table 17.1 (continued)

No.	View in designer	View in actual usage	Explanations
6.			<p>The next screen is a 3D representation of a bathroom. The user can again interact with the content at their convenience. Moreover, the user can listen to the pronunciation of five main objects from the room (the toilet, the bathtub, the window, the sink, and the mirror). At button click, the user will be redirected to the next screen.</p>
7.		<p>Not necessary/no impact</p>	<p>The AR experience will end. There is no interaction of the user with this screen. It contains only a 2D picture and its purpose is to draw an end to the experience through the application.</p>

QR-code and an internet connection. This section will further cover the advantages and the challenges that come with this prototype.

First of all, the prototype is based on the augmented reality principles. It contains 3D objects that the user can interact with. In this way, multiple senses (such as the sense of touching and the sense of seeing) are combined. In this way, the overall experience is more interactive, thus, user’s engagement and motivation will undoubtedly increase. By using this prototype in a French class, the association between the referent (an object from the house) and the reference (the word to name the object) becomes stronger and the knowledge can persist for a longer period.

Second of all, the development of such a prototype does not imply advanced programming skills. Any teacher with a minimum of digital skills, a clear purpose, and an Internet connection can easily create an augmented reality experience. This paper explored the capabilities of Zapworks Platform, one of the products of Zappar technology, a very powerful tool in the field of AR since 2011 [23]. The tool that the platform offers is called Zapworks Designer and it operates on the drag-and-drop principle. In this way, one can rapidly create a WebAR application to be used on any modern mobile phone. They also have at one’s disposal a very thorough documentation and multiple tutorials any novice in the sphere of developing AR

experiences should consult [24]. Moreover, while using the designer, one can create an AR experience both for handheld devices and for headset devices, according to the needs. The marker is also diverse, being possible to attach content to surfaces, images, or faces.

Another advantage of the prototype is that free content and free tools were used to develop it. The 3D content is entirely customized and based on free resources. In addition, it is of good quality and extremely diverse. The designer can also be used free of charge, although an account needs to be created to access the page. Authentication with a Google account is also possible, therefore, the experience should be extremely easy. The free usage of Zapworks Designer is limited to a 2-week free trial, afterwards additional costs are implied. Nonetheless, the time of free usage is enough to build a complete WebAR experience for a one-time usage in the classroom.

The biggest benefit of the prototype is unquestionably the gamification of the experience. The progression through the screens is sequential, thus they can be associated to modules of a game. Moreover, the fact that the user can interact with the 3D content, by rotating it, making it bigger or smaller in order to analyze it, will have a positive impact on the learning experience. In this way, the user will feel more engaged, more motivated to continue since each screen will bring new content. It is indeed a better approach than the traditional one, being based on the constructivism pedagogy. The pupil is the one who discovers the content and progresses through the lesson at their pace, whereas the teacher is there only to assist them. Therefore, this prototype encourages a student-centered approach.

The prototype has some challenges that need to be mentioned. Firstly, the fact that the author of the prototype needs to create the whole content might be challenging. Even though the content can be fully customized according to the needs of the lesson, finding the appropriate content at a good quality may be difficult. There are not many free resources that offer 3D content so that it can be integrated into an augmented reality application. The quality of the content is also important since it will have a direct impact on the learning experience. Therefore, the majority of time will be dedicated in finding the resources that are free and that offer high-quality content.

After finding the resources, it might be difficult to manipulate the content. Although minimal digital skills are implied and tutorials are at one's disposal, one needs to take into account that the application needs to be responsive. In this way, every student, regardless of the device they are using, will experiment the same learning act. So, the biggest challenge will be to combine the content and to fit it in the screen in order to create an easy-to-use application.

One last challenge that applies to the Romanian undergraduate education system is the need for personal devices (both for students and teachers). Currently, the majority of teachers and students from the public sector are not provided with devices (such as mobile phones, tablets, or laptops) that they can use within the learning act. So, the teachers need to have their personal device to create the application, whereas students need to have a personal device to use in class. This approach implies costs on a personal side and some teachers or students might not afford to have such experiences.

17.5 Conclusions

Augmented reality unquestionably brings valuable experience in the education act, both for the higher education and for the undergraduate system. Through this technology, the student is able to directly interact with the content, that is emerged in the real environment. In this way, connections are building more easily, more efficiently due to the combination of multiple senses.

This paper firstly offers an insight into the terminology of augmented reality. It presents the taxonomy Milgram et al. [4] developed and explains the differences between mixed, virtual, and augmented reality. Then, it offers a summarization of the experience of the Romanian education system with AR systems. It was proven that, at the moment of writing this article, there is no legislative attempt to include AR in the teaching act even if studies have shown that the community, the actors (teachers and students) see this movement as a beneficial one [12]. Moreover, multiple prototypes of AR applications and their impact on the learning act were researched in different fields. Nonetheless, the focus was on sciences.

Secondly, the focus of this paper shifted to presenting a topology of augmented reality applications. Two criteria were taken into consideration: the emphasis of the application and the presence of a physical trigger in order to start the application. Therefore, applications can be divided as being focused on roles, tasks, or locations [15], with the possibility of overlapping [16]. The applications can also be categorized as marker-based or markerless-based.

The next section of the paper focused on presenting a prototype of a marker-based AR application that can be used as an adjuvant of a French lesson for primary school children. It is a gamification-based experience where the pupil acquires vocabulary related to rooms and items from a house while interacting with 3D content. The application was developed using only free resources.

This paper confirms that teachers with minimal digital skills can develop on their own AR experiences that can be brought in class. In this way, children can benefit from all the advantages of an augmented reality ecosystem. Nonetheless, there is no directive at the moment to regulate a permanent introduction of such technologies in the educational act. This paper contributes to the literature by presenting a prototype in a humanistic field. This approach has not been undertaken so far in the Romanian community of researching, the focus being primarily to the scientific fields.

References

1. Ossiannilsson, E.: Blended Learning. The state of nation, ICDE (2017)
2. Joan, R.: Enhancing education through mobile augmented reality. *J. Educ. Technol.* **11**(4), 8–14 (2015)
3. Mann, S.: Mediated reality: University of Toronto RWM Project. *Linux J.* (1999)
4. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: a class of displays on the reality-virtuality continuum. *Telemanipulator Telepresence Technol.* **2351**, 282–292 (1994)

5. FitzGerald, E., Ferguson, R., Adams, A., Gaved, M., Mor, Y., Thomas, R.: Augmented reality and mobile learning: the state of the art. *Int. J. Mob. Blended Learn.* (2013)
6. Seidel, R.J., Chatelier, P.R.: *Virtual Reality, Training's Future? Perspectives on Virtual Reality and Related Emerging Technologies.* Springer Science & Business Media, Berlin (1997)
7. Baratè, A., Haus, G., Ludovico, L.A., Pagani, E., Scarabottolo, N.: 5G technology for augmented and virtual reality in education. In: *International Conference on Education and New Developments* (2019)
8. Elmqaddem, N.: Augmented reality and virtual reality in education. myth or reality? *Int. J. Emerg. Technol. Learn. Emerg. Technol. Learn.* **14**(03), 234–242 (2019)
9. Yuen, S., Yaoyuneyong, G., Johnson, E.: Augmented reality: an overview and five directions for AR in education. *J. Educ. Technol. Dev. Exch.* **4**(1), 119–140 (2011)
10. Azuma, R.T.: A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* **6**(4), 355–385 (1997)
11. Chicioreanu, T., Amza, G.: Are Teachers ready for the augmented reality?. In: *Proceedings of the 17th International Scientific Conference eLearning and Software for Education, Bucharest* (2021)
12. Gheorghe, M., Caranica, C.-N., Platon, O. E.: Augmented reality—a future perspective for improving learning process in the Romanian higher education system. In: *Proceedings of the 14th International Scientific Conference eLearning and Software for Education, Bucharest* (2018)
13. Crăciun, D., Bunoiu, M.: Augmented reality in Romanian science education pros and cons. In: *Proceedings of the 12th International Scientific Conference eLearning and Software for Education, Bucharest* (2016)
14. Macariu, C., Iftene, A., Gîfu, D.: Learn chemistry with augmented reality. *Procedia Comput. Sci.* **176**, 2133–2142 (2020)
15. Wu, H.-K., Lee, S.W.-Y., Chang, H.-Y., Liang, J.-C.: Current status, opportunities and challenges of augmented reality in education. *Comput. Educ. Educ.* **62**, 41–49 (2013)
16. Kljun, M., Geroimenko, V., Pucihar, K.C.: Augmented reality in education: current status and advancement of the field. In: *Augmented Reality in Education. A New Technology for Teaching and Learning*, Springer Nature Switzerland AG, pp. 3–22 (2020)
17. Klopfer, E., Squire, E.: Environmental detectives—the development of an augmented reality platform for environmental simulations. *Education Tech. Research Dev.* **56**, 203–228 (2008)
18. Squire, K.D., Jan, M.: Mad city mystery: developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *J Sci Edu Technol* **16**, 5–29 (2007)
19. Johnson, L., Levine, A., Smith, R., Stone, S.: *Simple augmented reality*, TX: The New Media Consortium (2010)
20. Lee, K.: Augmented reality in education and training. *TechTrends* **56**(2), 13–21 (2012)
21. Martie, S.L.: Ziua Internațională a Francofoniei. Peste 320 de milioane de oameni vorbesc franceza în lume, Edupedu, 20 March 2023. <https://www.edupedu.ro/20-martie-ziua-internati-onala-a-francofoniei-peste-321-de-milioane-de-oameni-vorbesc-franceza-in-lume/>. Accessed Apr 2024
22. Google, Poly was shut down on 30 Jun 2021. Google. <https://support.google.com/poly/answer/10192635>. Accessed Apr 2024
23. Z. Ltd.: About us, Zappar 2024. <https://www.zappar.com/about/>. Accessed Apr 2024
24. Z. Ltd.: Designer, Zappar 2024. <https://zap.works/designer/>. Accessed Apr 2024

Chapter 18

Empowering Sustainability: Upcycling Smartphones as the Future of IoT and Edge Computing in Emerging Economies



Ulrich Norbistrath , Perseverance Munga Ngoy,
Renato Perotto Machado , Huber Flores , and Mohan Liyanage 

Abstract Facing the accelerating issue of electronic waste (e-waste), with a record 62 million tons generated globally in 2022 and less than a quarter documented as formally collected and recycled, there is a pressing need for innovative solutions to mitigate their environmental and health repercussions. The rate at which e-waste is growing is five times faster than the rate of documented recycling, highlighting a widening gap in e-waste management. This burgeoning challenge not only poses severe environmental and human health risks but also signifies a loss of valuable natural resources, estimated at US 61 billion in 2022 due to inadequate recycling efforts. Moreover, the disparity in technological access continues to widen in emerging economies, often further hindered by the scarcity and elevated costs of modern hardware. This discrepancy exacerbates barriers to sustainable development, underscoring the urgent need for equitable and innovative technological solutions to bridge this gap. This paper introduces an innovative solution to these intertwined challenges: the repurposing of discarded smartphones. For the implementations as edge gateways for IoT applications, feasible on nearly all Android smartphones

Supported by the European Social Fund.

U. Norbistrath (✉) · P. M. Ngoy · R. P. Machado · H. Flores
Institute of Computer Science, University of Tartu, Tartu, Estonia
e-mail: ulrich.norbistrath@ut.ee

P. M. Ngoy
e-mail: perseverance.munga.ngoy@ut.ee

R. P. Machado
e-mail: renato.perotto.machado@ut.ee

H. Flores
e-mail: huber.flores@ut.ee

M. Liyanage
Computer Science Department, FH Dortmund, Dortmund, Germany
e-mail: mohan.liyanage@fh-dortmund.de

released later than 2014. By harnessing old Android phones that are readily available and perceived as not having apparent value, we propose a sustainable solution that not only addresses the issue of e-waste but also democratizes access to advanced technological resources.

18.1 Introduction

Electronic waste (aka e-waste) is the term used to describe discarded electrical or electronic devices that have become obsolete or broken. The ongoing challenge of how best to dispose of used and unwanted electronics is a long-standing issue and dates back to the 1970s¹. However, a lot has changed since then, particularly the number of discarded electronics today. E-waste contains toxic substances, such as lead, mercury, and cadmium, that can leach into the soil and water or release harmful emissions when burned [1]. It also represents a waste of valuable resources, such as metals, plastics, and rare earth elements, that could be reused or recycled. According to the World Health Organization, in 2019, an estimated 53.6 million tons of e-waste were produced globally, but only 17.4% was documented as formally collected and recycled². Recently, the magnitude of this challenge has just intensified, as evidenced by the latest Global E-waste Monitor 2024 report,³ which revealed that in 2022, a staggering 62 million tons of waste were generated globally. This volume is growing five times faster than the rate of documented e-waste recycling, highlighting a critical and widening gap in e-waste management efforts. Furthermore, only 22.3% of the e-waste generated in 2022 was documented as formally collected and recycled, leaving a vast amount of recoverable natural resources worth USD 91 billion, from which only USD 28 billion are recovered. The rest exacerbates pollution risks worldwide. With annual e-waste generation projected to reach 82 million tons by 2030, the challenge of e-waste management is becoming more acute, emphasizing the need for urgent action to enhance collection and recycling rates globally.

One of the primary sources of e-waste is the rapid growth of mobile phones, which are often replaced by newer models or discarded when damaged. Mobile phones are ubiquitous and have many features and functions that make them useful for communication, entertainment, and education. However, mobile phones have limited lifespan and durability and are challenging to repair or upgrade. As a result, millions of mobile phones end up in landfills, are just improperly stored away at home out of convenience, or are exported to developing countries, where they are often processed in unsafe and unprofessional ways⁴.

¹ <https://www.ewaste1.com/what-is-E-waste/>.

² <https://www.who.int/news/item/15-06-2021-soaring-E-waste-affects-the-health-of-millions-of-children-who-warns>.

³ <https://www.itu.int/en/ITU-D/Environment/Pages/Publications/The-Global-Ewaste-Monitor-2024.aspx>.

⁴ [https://www.who.int/news-room/fact-sheets/detail/electronic-waste-\(e-waste\)](https://www.who.int/news-room/fact-sheets/detail/electronic-waste-(e-waste)).

Of course, one of the ways out of this dilemma would be mobile phone repair, which is a growing industry that offers lots of benefits and opportunities but also faces its challenges. Lots of practices limit the consumers' choices and rights and create barriers and risks for the mobile phone repair industry. According to a study by iFixit, a website that provides repair guides and tools, the average repairability score of smartphones in 2020 was just 4.6 out of 10, which means that most smartphones are hard to repair⁵. Good practice would be to design the phones in a way that facilitates easy disassembly, and repair [2, 3].

However, not all phones need to be repaired. They could be just reused:

Reuse is an essential aspect of sustainability, as it extends the use of old mobile phones by giving or selling them to others [4] or giving them a new function. It can also happen when mobile phone manufacturers, retailers, and others resell or donate the phones collected through various schemes. Reuse can improve resource efficiency, reduce dependency on new materials, and provide access to affordable phones for people who cannot buy new ones [5, 6]. However, reuse is rarely one of many disposal methods adopted by users/consumers. Many store their old phones after replacing them, mainly due to technological obsolescence. It hinders the reuse potential as the value and functionality of the phones decrease over time, and their lithium-ion batteries might even degenerate to serious levels and become a potential fire hazard.

The main value proposal of this paper is to reuse an old phone as an edge gateway. An edge gateway serves as a point where data from various IoT devices converge, enabling preliminary data processing closer to the source, thus reducing latency and bandwidth usage as well as allowing disconnected operation if the internet is not constantly available. By repurposing old mobile phones as edge gateways, we leverage their processing power, connectivity options, and portability, transforming them into cost-effective, eco-friendly nodes within IoT networks. This approach extends the lifecycle of mobile devices, mitigating e-waste, and democratizes access to IoT technology, especially in emerging economies where cost and accessibility are major barriers. We will outline technical requirements, potential applications, and challenges to implementation. This research contributes to the burgeoning field of circular economy practices in information technology, offering insights into how societies can transition toward more sustainable consumption patterns of electronic devices.

Furthermore, this paper leverages the IoTempower teaching framework [7] and its comprehensive toolset that significantly simplifies the integration, compilation, and deployment of ESP-based devices as edge components within IoT ecosystems. IoTempower seamlessly installs essential tools such as Mosquito, Node-RED, and PlatformIO, providing pre-configuration for a wide range of ESP devices alongside extensive sensor and actor support. This facilitates a streamlined development process for IoT applications, from conceptualization to deployment, making it an invaluable asset for educational and practical purposes in the realm of IoT. IoTempower's framework services and gateways can be deployed on a multitude of devices, including

⁵ <https://www.ifixit.com/repairability/smartphone-scores>.

PCs, single-board computers like the Raspberry Pi, and smartphones, which makes it versatile to repurpose obsolete hardware.

The IoTempower framework not only demonstrates the practical feasibility of utilizing upcycled smartphones as edge gateways but also enhances the hands-on learning experience in IoT education. Through fostering collaboration and experimentation, IoTempower embodies a novel approach to IoT education, promoting a deeper understanding of IoT technologies and their potential applications. This is detailed further in our initial study on the framework, which presents a case study on leveraging devices as “things” for hands-on collaboration within IoT educational settings [8].

A common choice for a home IoT edge gateway is a modern Raspberry Pi which was around USD 1000 at the beginning of 2020 in Brazil due to chip shortage caused by the 2019 pandemic. Currently, a modern Raspberry Pi costs around USD 200 on Mercado Livre in Brazil and USD 100 in the Western world. A simple Android phone costs at the time of writing around USD 80, a sturdy used laptop that can be used as a server costs around USD 100 in the Western world and USD 200 in Brazil on Mercado Livre. In Nigeria, the price of a Raspberry Pi is approximately USD 200⁶. Similarly, in South Africa, the cost for a modern, well-specced Raspberry Pi is about USD 100*. These prices indicate that while there is a difference in cost due to regional economic factors, the increase is not excessively disproportional⁷. Initially, that would not hint at upcycling or reuse as being as important from an economic perspective. However, assuming that old phones are just aging and stored away and could be given a second life and that this boosts economic opportunity needs to be factored into our proposal. The potential of using basically free resources for this and avoiding waste at the same time should make it even more interesting—also for startups in Western communities.

At a university showcase, we garnered interest from an Ethiopian organization in developing a cost-effective power measurement device based on our setup. This system aims to allow electricity providers to accurately bill small apartment units, leveraging a smartphone as an edge gateway paired with a 16-channel ESP32-based energy monitor. This project exemplifies how upcycling smartphones can address specific economic and infrastructural challenges in emerging economies, providing affordable and scalable technology solutions.

18.2 Related Work

In this Section, we review the literature on mobile phone recycling and related work that provides insights into the challenges and opportunities in e-waste management, the integration of IoT devices in waste management, policy analysis for e-waste, and

⁶ <https://yaoota.com/en-ng/product/raspberry-raspberry-pi-4-2gb-ram-price-fromjumia-nigeria>.

⁷ It is important to note that these prices are subject to change and may vary based on the specific model and market fluctuations.

sustainable strategies for mobile phone lifecycle management. Most of the previous studies on the disposal of obsolete mobile phones have focused more on recycling than on reuse. Reuse was often a minor aspect or not considered at all. Reuse is a key element of the circular economy, and it can contribute to the environmental and economic sustainability of mobile phone consumption.

Repurposing old phones and laptops as servers is a sustainable and cost-effective approach to utilizing technology that might otherwise be discarded. Smartphones, with their compact size and efficient power usage, can be transformed into servers for various tasks, such as web hosting or local network ad-blocking [9]. Similarly, old laptops can serve as private servers, especially when equipped with lightweight Linux distributions that are resource-efficient and user-friendly [10]. Researchers at Princeton University have demonstrated the feasibility of constructing servers from old smartphones, highlighting the economic and environmental benefits of this practice [11]. These repurposed devices can provide data, storage, or computing power, functioning as the backbone for personal projects or small-scale server needs [12]. Alfaras Altes (2019) discusses sustainable alternatives for repurposing discarded smartphones, focusing on the environmental and economic viability of such practices. This work underscores the importance of considering the environmental and social impacts of upcycling processes, which is central to our study's objective of extending the lifecycle of smartphones in a sustainable manner [13]. Wanderley and Bonacin (2019) focus on the design of mobile and IoT solutions for sustainable smart cities, with an emphasis on electronic waste disposal. The discussed hackathon outcomes enhance smart city ecosystems, promoting design thinking to allow technological advancement to go hand in hand with sustainability efforts, mixing mobile phones, IoT, and e-waste management [14]. Integrating smartphones as IoT edge devices in emerging economies offers compelling economic opportunities, bridging the digital divide and fostering technological advancement. The widespread accessibility and utilization of upcycled smartphones for various services such as mobile banking, health, and education can significantly contribute to societal empowerment and economic development [15]. This approach is not only sustainable but also promotes job creation in the refurbishment sector, reduces dependence on imported materials, and mitigates CO₂ emissions by extending the lifecycle of mobile devices, aligning with the principles of a circular economy [16]. Furthermore, the role of smartphones in transforming livelihoods and creating new business models in resource-deficient environments showcases the transformative potential of mobile technology in developing countries [17]. Entrepreneurs in Ghana, for instance, are using smartphones to foster business growth despite barriers, demonstrating the devices' capacity to reshape economic opportunities through enhanced connectivity and innovation.

Such endeavors highlight the untapped potential of upcycled smartphones in emerging economies, underscoring the need for collaborative efforts in technology innovation, policy formulation, and entrepreneurship to harness this potential fully. This exploration into upcycling for IoT applications reveals a path toward not only environmental sustainability but also economic empowerment and technological inclusivity in regions that stand to benefit the most.

18.3 Solution Design

We explore three methods for deploying edge services on Android smartphones, evaluating them based on their complexity, potential pitfalls, and rewards. Additionally, we have also worked with alternative hardware solutions like Raspberry Pi and old laptops, but we focus here on old phones.

(Choice 1) Installing a Linux distribution, such as PostmarketOS, on an Android device offers a direct approach to repurpose it for edge computing. However, this method is often cumbersome and tends to be viable mainly for older phone models. The performance of these devices might be limited, potentially restricting their use to less demanding applications or cloudlets. This requires a high level of *skill set difficulty*. *Potential issues* are the compatibility with newer phone models and performance limitations. The *Reward* is moderate, considering the environmental benefit and the potential for easily being able to install a wide variety of edge computing tasks.

(Choice 2) Aftermarket Android Version with chroot (e.g., LineageOS) combined with a chroot environment provides a balance between maintaining an Android user interface and accessing Linux capabilities. This method's feasibility varies by phone model and involves a complex setup process. This requires a high level of *skill set difficulty*. *Potential issues* are device compatibility, complex setup, and system stability. The *Reward* is medium to high, offering a versatile environment with access to Android and Linux functionalities.

(Choice 3) Combining Termux and IoTempower represents the most accessible option, requiring only a factory reset of the device. IoTempower enhances Termux by streamlining the installation of essential IoT tools like Mosquitto, Node-RED, and PlatformIO, alongside a comprehensive set of libraries for device integration and development. This requires a low level of *skill set difficulty*. *Potential issues* are the autostart capabilities and battery management without root access; however, external battery regulation can be achieved with IoT setups using relay switches and Node-RED. The *Reward* is very high, due to the ease of setup and the broad range of supported devices and applications.

The IoTempower installation, as shown in the next section, facilitates a seamless setup, installing necessary components and providing a local web-based orchestration environment. Therefore, we are going to elaborate here on Choice 3.

18.4 Implementation

For the implementation any Android phone from 2014 onwards should work here. Connect the phone to the internet, install F-Droid⁸, and from there Termux and Termux:Boot. Upgrade packages, acquire termux-wake-lock and storage access, and install openssh, git, curl, and termux-api. Perform an SSH key exchange to enable

⁸ <https://fdroid.org>.

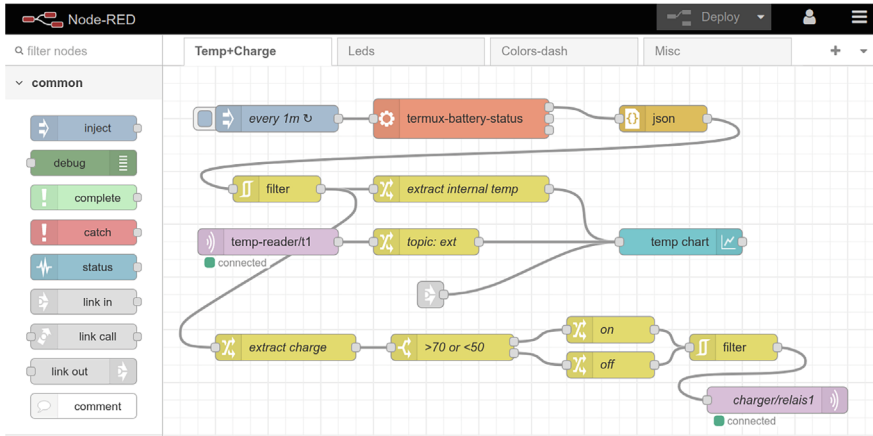


Fig. 18.1 Node-RED flow for temperature and charging setup

secure connections without password prompts and start the ssh daemon. Activate the phone's hotspot feature and establish an SSH connection to it. Install IoTempower and its IoT edge services with its one-line installer as described here⁹. To ensure seamless operation following device reboots, make sure to autostart termux-wake-lock, sshd, and IoTempowers services mqtt starter, and web starter, basically enabling Mosquitto, Node-RED, and access to IoTempower's PlatformIO compilation environment. Ensure to turn off battery and energy optimizations to keep the services running. Consider enabling a reverse proxy for public internet access and installing a local database like influxdb.

If the battery is still functional (not swollen and holds at least an hour's charge), it can be managed effectively through Node-RED using an exec node and calling termux-battery-status, which returns JSON formatted battery status with health, percentage, plugged, status, and temperature. Figure 18.1 shows our Node-RED flow for temperature monitoring and charging management. To prolong the lifespan of a battery that is still in acceptable condition, it is advisable to maintain its charge level at around 50% to prevent fully charging or discharging that both harm the condition of the battery.

If the battery is swollen or not working anymore, try to remove the Lithium part of the phone's battery and bypass it with wall-power. Start by determining the precise voltage (typically between 3.7 V and 4.2 V) and current required by consulting the phone's specifications or its original battery's label. Next, use specialized tools like a spudger to open the phone and access the battery connector on the motherboard, noting the orientation and pin configuration. Disconnect the battery, prepare your power supply and step-up-down converter, ensuring the correct output voltage, and add a fuse safety. Solder the power supply's wires to the corresponding terminals of the battery connector, possibly omitting the third pin used for temperature

⁹ <https://github.com/iotempire/iotempower/blob/master/doc/installation.rst>.

monitoring or communication unless essential for your setup. Finally, secure all connections to prevent short circuits. Modifying a phone’s power system to serve as an IoT edge gateway presents a complex challenge, posing several technical hurdles and requiring specialized skills. Ensuring stable voltage regulation, managing the device’s peak current draw, protecting the circuit from overcurrent and short circuits, and overcoming space constraints within the phone’s form factor require a lot of attention to detail. Moreover, the process demands precision soldering, circuit analysis, proficient use of a multimeter, and mechanical dexterity to disassemble and reassemble the components safely. Given these risks, we would, in general, advise against this approach.

We use IoTempower to configure actors and sensors, facilitating a streamlined integration of those. For our demonstration case, we connect a DS18x20 (Dallas) temperature sensor and a WS2811 12 V LED strip to an ESP32-based M5StickC and an ESP8266-based Wemos D1 Mini, respectively. The programming for each component is straightforward and can be done with the one-line per device strategy of IoTempower: *(node 1)* `dallas(temp);`, *(node 2)* `rgb strip bus(strip, 10, F GRB,NeoEsp8266Uart1800KbpsMethod, D4);`, and *(node 3)* `output(relais1, RELAIS1).off(); input(button1, BUTTON1);` This way, we can monitor and control the actors, sensors, and even the smart phone’s battery health.

Caveat: There are still a few problems with compiling microcontroller code with PlatformIO directly on Termux following the IoTempower installation on 32bit phone architecture. Despite PlatformIO’s reliance on vendor toolchains that are incompatible with Termux’s Bionic libc, a workaround involving glibc-based Python and glibc-runner has been partially successful. Efforts by the IoTempower community are anticipated to resolve these issues in the very near future. Of course, we can also use IoTempower on any other Linux-like platform, like native Linux distributions, WSL, or MacOS, and compile the firmware there. The integration and edge functionality work well now with the described method above.

As seen in Fig. 18.2, we installed the setup in our in-class “demo-box,” which we use to teach students how to build their own IoT access control system. Therefore, it shows not only the Sonoff relay switch in the lower left corner, the phone as a gateway in the center (also active as a Node-RED dashboard¹⁰), the external temperature sensor to the left of the gateway, and the led-strip on the top, but also RFID card reader, solenoid lock, and small notification screen for simulating a full access control system.

¹⁰ Old phones and tablets can serve as excellent wall-mounted dashboards for status display or control interfaces, further extending the utility of repurposed devices. The described battery hacks might still apply.



Fig. 18.2 The phone-based IoT edge gateway in our “demo-box”

18.5 Evaluation, Discussion, and Future Work

Throughout our experiments, we encountered various challenges and insights: (1) Maintaining the electronic components and battery of a smartphone in a functioning state presented considerable difficulties. (2) Using PostmarketOS or chroot environments in aftermarket phone firmware is rewarding but requires significant system skills. (3) The implementation of Termux combined with the IoTempower framework stood out for its simplicity, making the upcycling process accessible even to those with limited technical expertise.

The goal of this project extends beyond cost savings; it encompasses e-waste reduction and fostering community engagement. The power company use case has garnered interest from both students and organizations, highlighting the practical applications of our approach. Experiences in Brazil demonstrated the potential for enabling artists and citizen science projects. One author prototyped an edge gateway on a phone in Brasilia for Carnival 2019, creating a mobile gateway for a distributed costume. This setup showed that IoT edge gateways using Node-RED and a control dashboard work well on Android smartphones. Testing showed that Node-RED and Mosquitto operate effectively on smartphones, with a hotspot supporting up to 10 WiFi clients. While initially perceived as gimmicky for industrial applications, the real value of this project lies in its impact on local communities and its potential to inspire collective action. Our initiative aims to prevent e-waste, stimulate educational opportunities, and support startups in emerging economies, with promising interest from stakeholders in the energy sector. Moreover, this approach holds significant promise for reducing e-waste in developed countries. We advocate for the donation

of unused phones to local hackerspaces and maker labs, encouraging the adoption of such programs to extend the life of electronic devices.

Looking forward, we aim to (1) resolve compatibility in the offered libc environment, (2) broaden device testing to a more diverse array of smartphone models, (3) conduct international workshops to disseminate knowledge and foster global community involvement in smartphone upcycling, (4) integrate this initiative with local hackerspace programs to facilitate smartphone upcycle subgroups, and (5) partner with industry in the Western world to pilot and evaluate upcycling solutions, exploring their potential in e-waste reduction and sustainable development.

Our study highlights the crucial role that emerging economies can play in advancing sustainable technological solutions. By demonstrating the feasibility and benefits of repurposing discarded smartphones, we call for a paradigm shift in the global perspective on e-waste, envisioning a future where e-waste is seen as a treasure trove of opportunity for sustainability, inclusivity, and innovation in the digital era.

References

1. Yu-Gong, T., Wu, Y.F., Zhe-Tan, Lei-Lv: Recent development of recycling lead from scrap crts: a technological review. *Waste Manag.* **57**, 176–186 (2016). <https://doi.org/10.1016/j.wasman.2015.09.004>, <https://www.sciencedirect.com/science/article/abs/pii/S0956053X15301136>
2. Cordella, M., Alfieri, F., Clemm, C., Berwald, A.: Durability of smartphones: a technical analysis of reliability and repairability aspects. *J. Clean. Prod.* **286**, 125388 (2021). <https://doi.org/10.1016/j.jclepro.2020.125388>
3. Hankammer, S., Jiang, R., Kleer, R., Schymanietz, M.: Are modular and customizable smartphones the future, or doomed to fail? A case study on the introduction of sustainable consumer electronics. *CIRP J. Manuf. Sci. Technol.* **23**, 146–155 (2018). <https://doi.org/10.1016/j.cirpj.2017.11.001>
4. Rahmani, M., Nabizadeh, R., Yaghmaeian, K., Mahvi, A., Yunesian, M.: Estimation of waste from computers and mobile phones in Iran. *Resour. Conserv. Recycl.* **87**, 21–29 (2014). <https://doi.org/10.1016/j.resconrec.2014.03.009>, <https://www.sciencedirect.com/science/article/pii/S0921344914000676>
5. Markov, T., Fishman, T., Chertow, M., Blass, V.: What affects the secondhand value of smartphones: evidence from Ebay. *J. Ind. Ecol.* **23** (2018). <https://doi.org/10.1111/jiec.12806>
6. Riisgaard, H., Mosgaard, M., Zacho, K.: Local circles in a circular economy: the case of smartphone repair in Denmark. *Eur. J. Sustain. Dev.* **5**(1), 109–124 (2016)
7. Norbirsath, U.: IoTempower Community: IoTempower: A framework and environment for making the internet of things (Iot) accessible for everyone. <https://github.com/iotempower/iotempower> (2024). Last Accessed 08 Apr 2024
8. Machado, R.P., Norbirsath, U., Jubeh, R.: Iot educational framework case study: devices as things for hands-on collaboration. *J. Eng. Educ. Transform.* **37** (2024)
9. Skipo, P.: How to re-purpose your old android phone by running Linux on it—medium.com. Last Accessed 16 Apr 2024
10. Turn your old laptop or pc into your own private server devto. Last Accessed 16 Apr 2024
11. Using old smartphones to build computer servers—research.princeton.edu. <https://research.princeton.edu/news/using-old-smartphones-build-computer-servers> (2024). Last Accessed 16 Apr 2024

12. Old phones get new life in high-powered computer servers—princeton.edu. <https://www.princeton.edu/news/2017/10/24/old-phones-get-new-life-high-powered-computer-servers> (2024). Last Accessed 16 Apr 2024
13. Altes, E.A.: *Economia circular i reciclatgedes components electronics* (2019)
14. Wanderley, A.R.M.C., Bonacin, R.: Designing mobile and Iot solutions for sustainable smart cities. In: Streitz, N., Konomi, S. (eds.) *Distributed, Ambient and Pervasive Interactions*, pp. 212–226. Springer International Publishing, Cham (2019). <https://doi.org/10.1007/978-3-030-21935-2>
15. Toufe, H.I.: Empowering communities: how mobile is transforming development. *Harv. Int. Rev.* **34**(3), 33 (2013)
16. Rizos, V., Bryhn, J., Alessi, M., Campmas, A., Zarra, A.: Identifying the impact of the circular economy on the fast-moving consumer goods industry: Opportunities and challenges for businesses, workers and consumers—mobile phones as an example. Study (2019)
17. Coffman, C.: From individuals to organizations: how smartphones are shaping entrepreneurship in rural Africa (2019)

Chapter 19

Quantum Computing and Cybersecurity: Threat or Opportunity?



Alexandru Butoi

Abstract Quantum computing perspective is that it will revolutionize the way we implement and execute highly intensive processing algorithms or large space solution search heuristics. So algorithms of polynomial or exponential complexity will deliver their results in matters of hours, rather than days and the implications for classic and de-facto security algorithms are disruptive. We explore some possibilities for leveraging quantum computing power to secure data stored and exchanged in non-quantum environments while there is enough evidence in the existing literature that quantum algorithms may have the potential to level up the security level of a hybrid computational ecosystem consisting of a few quantum nodes and several non-quantum nodes. We explore a possible way of how we can use quantum algorithms to secure data stored in non-quantum environments in front of quantum threats.

19.1 Introduction

Quantum computing development and market availability of quantum computers would be by far a breakthrough in computer science and industry. Quantum computing perspective is that it will revolutionize the way we implement and execute highly intensive processing algorithms or large space solution search heuristics. So algorithms of polynomial or exponential complexity will deliver their results in matters of hours, rather than days and the implications for classic and de-facto security algorithms are disruptive. For example, in May 2019 MIT Technology Review released an article on “how a quantum computer could break 2048-bit RSA encryption in 8 h” citing an arXiv article [1]. Shor’s Algorithm could be used to efficiently factor these encryption keys when deployed on a powerful enough quantum computer [1].

The current paper explores this field of quantum computing and several possibilities for leveraging quantum computing power to secure data stored and transported

A. Butoi (✉)

Business Information Systems Department, Babeş-Bolyai University, Cluj-Napoca, Romania
e-mail: alex.butoi@econ.ubbcluj.ro

in non-quantum environments. We provide a relevant literature review related to quantum security and post-quantum security, cryptography approaches, key management, and quantum security risk assessment; we also explore the standardization efforts of NIST and Cloud Security Alliance to quantum computing security.

We explore how Fault Trees can be applied in quantum computing security risk modeling and how quantum computing creates the opportunity to overcome the main disadvantage of using large and complex fault trees on classical computation infrastructure due to the state space explosion. In the same section, we shortly envision a quantum fortress model, a conceptual model for architecting the security of non-quantum clusters or computational systems to protect them from quantum threats. The last section explores how quantum homomorphic encryption algorithms can enable new ways of post-quantum, quantum-reliable data encryption, and secure key management protocols.

19.2 Literature Review

At a very fundamental and simplified level, a security algorithm is designed to be hard to reverse engineer via brute force attack techniques. For example, a cryptographic algorithm uses advanced mathematical techniques that are very hard to reverse engineer and take too much time to decode via a trial-and-error brute-force strategy using classical computers. Quantum computers are more powerful than the classic ones and trial-and-error strategies are not so expensive and time-consuming as Shor's algorithm proved in 1994 [2, 3]. The implications of quantum computing technologies advancement for the non-quantum cybersecurity algorithms are tremendous and disruptive in this sense. The classic encryption algorithms will not become deprecated overnight, but as we build more powerful quantum computers that can compute a higher number of qubits, the more deprecated actual encryption mechanisms may become. In this sense, there is a solid track on GitHub for open-source projects on Post-Quantum cryptography consisting of several repositories and libraries that aim to secure cryptographic algorithms against potential future quantum computers, most of them being experimental projects at the moment [4]. Also, according to American Scientist [5], current quantum computers are still error-prone and able to compute a limited number of qubits but we can see from [1], that knowledge advancement in quantum algorithms may significantly improve the performance of these heuristics requiring significant fewer qubits and resources to run. For example in [1], a combination of different techniques and algorithms showed that a quantum computer can decode a 2048-bit RSA integer using 20 million qubits compared to ten years ago in 2015 when scientists showed that we would need billions of qubits to do the calculation reliably [3].

The post-quantum cryptography is also a project supported by NIST that aims to standardize security and encryption algorithms that are efficient against quantum attacks. As part of this project, they announced in July 2022, the first four Quantum

Resistant algorithms for Public Key Encryption and Key-establishment, as well as for Digital Signature Algorithms [6].

Cloud Security Alliance is conducting 3 research projects on the quantum security track related to the Transitioning of Blockchains to Quantum-Safe, Quantum Security Controls, as well as Quantum and AI [7]. In the Blockchain to Quantum-Safe track, authors state that Grover's algorithm is a threat to hash functions, while for the cryptographic functions, the biggest threat is represented by Shor's algorithm (pg. 17, 18).

Ye et al. [8] present a hybrid protocol (SQKD) for key distribution between quantum users and classical users. The security level can be the same as the full quantum protocol. We can speculate that the quantum presence in a hybrid computational ecosystem can increase the level of the system's security and the postulate that the security level of the system can be reduced to the weakest node may not apply here in specific setups. The quantum nodes seem to be enforcing the overall security level of the protocol despite that the non-quantum nodes are limited in performance. A similar approach is presented in [9], where a Quantum Key Distribution Protocol is used for sharing keys between parties to better secure confidential data.

In terms of risk and security modeling in the quantum computing world [10], tackle privacy protection for users having nothing or limited quantum capabilities and being threatened by users with more performant quantum capabilities. Quantum entanglement and uncertainty principle are used to ensure third-party security measurement. Also [11], tackles the problem of data privacy protection by transforming classic data representation into unique, unclonable quantum states, and due to the probabilistic nature of the states data is secured by design. A security mechanism aims to minimize the probability of a security breach happening, and quantum computing probabilistic states can be leveraged in this direction of modeling security risks and probabilities through quantum states. In [12], authors are leveraging information theory to encrypt large text content with a small encryption key that has at least the same entropy as the entire content measured entropy and in quantum computing, key multiplication to the exact size of the content proves to be faster than applying the same technique in non-quantum computers.

19.3 Models of Hybrid Quantum-Non-quantum Security

19.3.1 *The Fault Tree Quantum Model*

Fault trees are a theoretical and practical instrument used in safety and reliability engineering to map different events of fault in a complex system together with their associated probabilities. Using a set of logical gates we can compute the probabilities of complex sequences of events to happen that can affect the reliability or availability of a system. If we model the security risk of a complex computational system as a probability of a set of security-related events so the system's security is fully

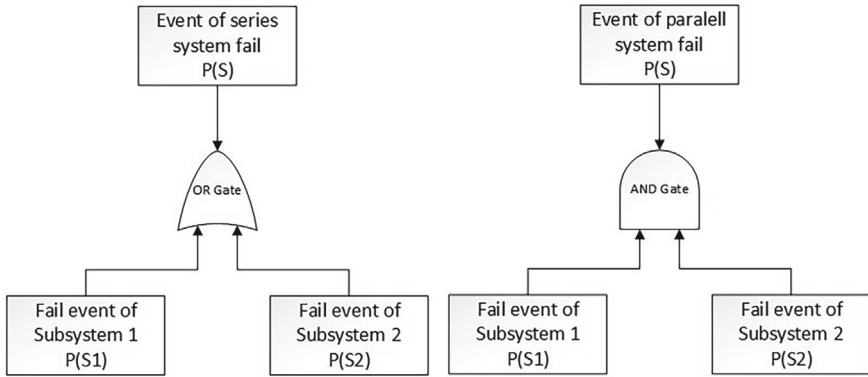


Fig. 19.1 Basic fault tree representation and gates [13]

compromised, we could employ fault tree models as a mathematical tool to measure the system’s security risk.

Given the probabilistic approach of the event mapping and event sequence computation, fault tree models might be a suitable instrument for developing quantum and post-quantum security algorithms deployable on both quantum and non-quantum terminals (Figs. 19.1 and 19.2).

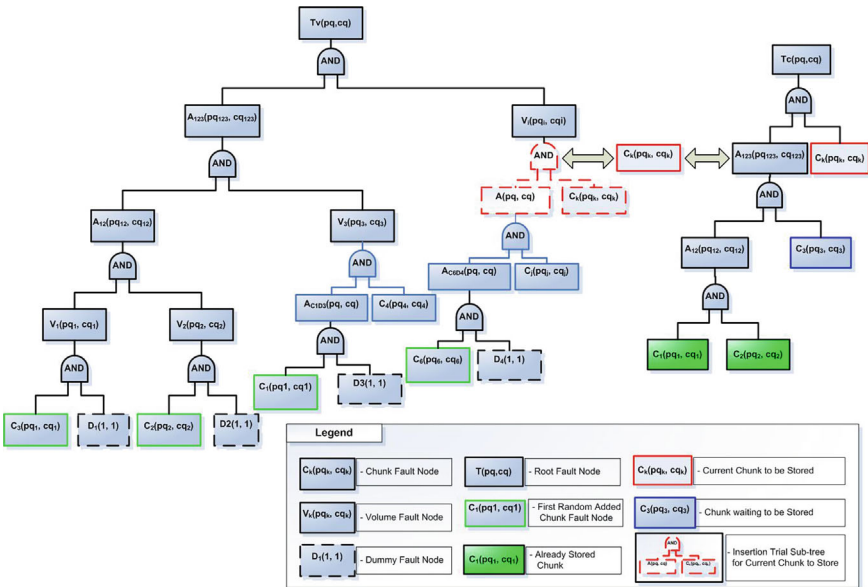


Fig. 19.2 Example of a fault tree for securely distributing segments of data into a finite volume set of virtualized storage in the cloud [16]—a relatively simple problem produces a 5-level depth tree that can grow exponentially

These basic structures can take the form of a complex tree that is computed from bottom to top using formulas of independent or dependent events probability calculation.

A major downside of using these fault trees on non-quantum terminals is that for complex systems they become a representation of a large space of possible events and the space explosion problem appears, increasing the computation costs exponentially. Roth et al. [14] argue that for complex systems, the space of states grows rapidly as the complexity of the system grows, while Li et al. [15] argue that a large-scale fault tree processing problem can be an NP-hard problem that is not feasible to be deployed on non-quantum nodes but can be effective to deploy and execute on quantum terminals.

Silva et al. [17, 18] propose a method to deploy quantum fault trees into quantum computers by mapping every fault node (that represents a probability) to a qubit and using different quantum gates it computes the minimal cut set (the probability of root node—the total failure probability). The limitation of this approach would be that the large explosion of state space would mean a higher number of qubits, the number of states/fault nodes equals the number of qubits necessary to run the model on a quantum computer. Employing a single qubit to handle multiple probabilistic states would optimize the approach. Also, another downside of the current work is that a large and complex number of quantum gates are needed to implement the fault tree.

In conclusion, using fault trees to resolve post-quantum security problems could be a feasible approach due to the NP-hard nature of the space explosion problem that is difficult to compute on non-quantum computers. So if we model a space/set of possible threats and vulnerabilities with their probabilities of apparition in a classical or quantum complex system and then map them into a quantum fault tree we could obtain promising results for total and partial security risk calculation.

19.3.2 Quantum Fortress Model

In the following section, we propose a conceptual model for architecting the security of non-quantum clusters or classic computational systems to protect them from quantum threats. We call this model the Quantum Fortress Model in which one or more quantum computers have the single responsibility of running quantum algorithms that aim to protect the non-quantum nodes from threats both non-quantum and quantum-originated. For reliability reasons, we employ three quantum computers to protect a large cluster or network of non-quantum devices.

Figure 19.3 depicts a hybrid topology in which quantum computers are securing non-quantum devices in a reliable and redundant setup. The functions of the quantum nodes would be to: run firewall filtering quantum algorithms, load balancing between the three nodes, consensus mechanisms, security, and data risk assessment in an open and connected to the Internet ecosystem; for example, Fault tree security models could be deployed on these quantum nodes for in-time security risk assessments.

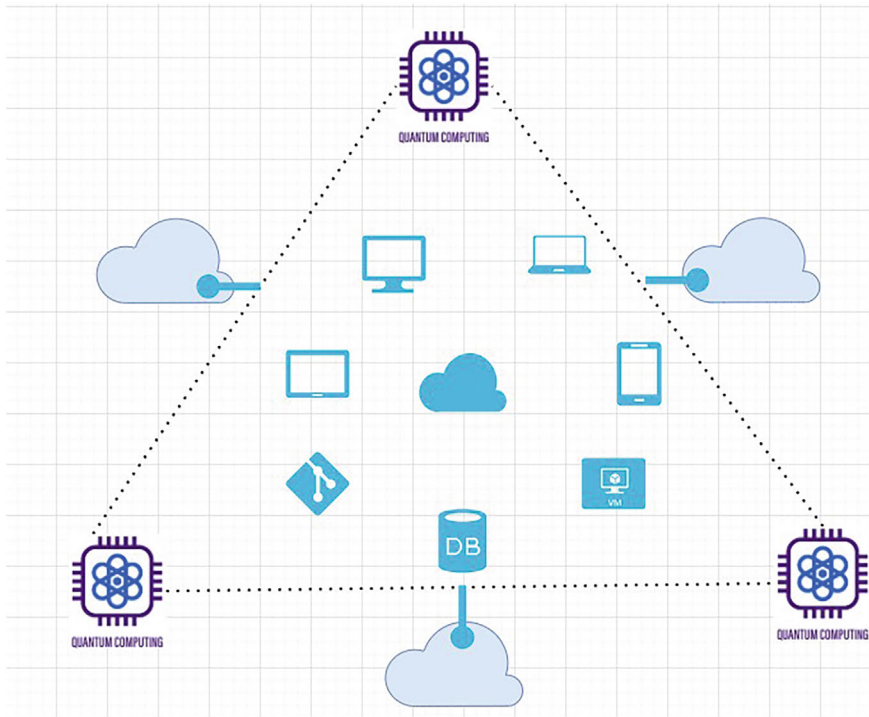


Fig. 19.3 Quantum fortress model—quantum computer establishing intranet virtual limits

Post-quantum encryption, post-quantum identity algorithms, and encryption key management could be deployed on quantum nodes for better security against the quantum and non-quantum-originated threats in the outer and inner ecosystems.

The setup should be redundant, all 3 quantum nodes should be redundant nodes that would have the ability to run in parallel all algorithms that other quantum nodes deploy. The sole objective would be to maximize the qubit utilization and minimize errors. Specific protocols are required to ensure communication with the outer and inner non-quantum devices and networks.

19.3.3 Homomorphic Encryption and Key Management

Tan et al. [19] present a quantum private-key homomorphic encryption protocol with a broad class of computation. Homomorphic encryption allows data processing to be performed without knowing the encryption key or decrypting the data. In 2009 Gentry proposed a fully homomorphic encryption algorithm using lattices [20], that opened the road to secure homomorphic encryption of data, the processor having

zero knowledge of the data. In [19], encrypts data in such a way that encryption affects only spatial modes of the qubit particles and leaves intact the internal states.

This sort of encryption could provide a way of quantum encrypting non-quantum encrypted data. Before sending the data to the destination through the open internet, the quantum fortress nodes would encrypt the data without affecting its integrity; then the data would be decrypted by the quantum fortress destination node and then passed to the non-quantum destination node to be decrypted and consumed or processed in a classical algorithmic manner. The same technique could be applied to securely delivering keys over the Internet in a post-quantum secure manner.

Classic RSA public keys could be encrypted with quantum homomorphic encryption before sharing them with other actors to protect them from malicious quantum de-factorization using Shor’s algorithm. To tackle this, we would need a conversion algorithm that knows how to represent byte-encoded data into qubits and vice-versa. If we consider a public key as a succession of bytes from 0 to 255, there are ways to encode the byte representation of the key into qubits: as exemplified in [21] we transform every byte in the byte-key representation in its binary representation and set the state of the i -th qubit as the value of i -th bit. A more advanced model to represent integers in quantum states by combining sub-partitions and quantum integers is presented in [22].

Figure 19.4 depicts the process of quantum homomorphic encryption of a public key before sending it over to a public network.

When the destination is inside a quantum fortress the homomorphic encrypted key (Quantum Public Key—QPK) can be validated without decrypting it and if it is valid can be decrypted and it can be converted back to its byte representation to be processed by a non-quantum destination device.

When the destination is a non-quantum computer that has no access to a quantum computer to decrypt it, the source quantum computer can add an extra step and

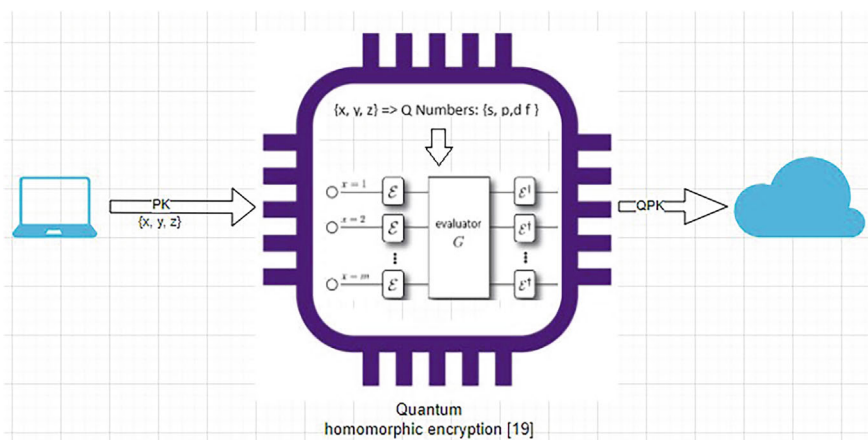


Fig. 19.4 Homomorphic public key encryptions

transform the encrypted quantum key into an equivalent byte representation of the QPK. Then it has to factor a private-key from the new QPK byte representation and provide it to the source device. In this scenario, the byte representation of the QPK will be the public key in the communication between the source device and the destination.

19.4 Conclusion

The current paper aims to identify some state-of-the-art aspects of the debate around the advancements of quantum computing putting at risk of deprecation the current security algorithms and encryption schemes. It is also focused on existing actions to enable post-quantum encryption schemes that would be reliable enough in front of quantum computing threats. On these premises, we presented some opportunities and a conceptual framework for tackling the post-quantum security problem, not only from the angle of data encryption but also from the perspective of security risk assessment and modeling using quantum fault trees. We advocate for a hybrid approach where quantum security algorithms would engage in securing data and RSA keys from and for non-quantum devices, actively protecting them from quantum and non-quantum security threats.

Further development of the current paper would be the experimental implementation of the Quantum Fortress Model using quantum algorithms on simulated and non-simulated quantum infrastructure, aiming to explore the optimization of qubits usage in quantum fault trees.

References

1. Gidney, C., Eker, M.: How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. <https://arxiv.org/abs/1905.09749> (2019)
2. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.* **41**(2), 303–332 (1999)
3. Technology Review.com.: <https://www.technologyreview.com/2019/05/30/65724/how-a-quantum-computer-could-break-2048-bit-rsa-encryption-in-8-hours/> (2024). Last Accessed 19 Apr 2024
4. Online Resource: <https://github.com/topics/post-quantum-cryptography> (2024). Last Accessed 19 Apr 2024
5. Denning, E.: Dorothy: Is Quantum Computing a Cybersecurity Threat? <https://www.americanscientist.org/article/is-quantum-computing-a-cybersecurity-threat> (2024). Last Accessed 19 Apr 2024
6. National Institutes for Standards and Technology: Post-Quantum Cryptography—Selected Algorithms for 2022. <https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022> (2024). Last Accessed 19 Apr 2024
7. Cloud Security Alliance: https://cloudsecurityalliance.org/research/working-groups/quantum-safe-security#current_projects (2024). Last Accessed 19 Apr 2024

8. Ye, C.Q., Li, J., Chen, X.B., et al.: Security and application of semi-quantum key distribution protocol for users with different quantum capabilities. *EPJ Quantum Technol.* **10**, 21 (2023). <https://doi.org/10.1140/epjqt/s40507-023-00180-3>
9. Sundar, K., Sasikumar, S., Jayakumar, C., et al.: Quantum cryptography based cloud security model (QC-CSM) for ensuring cloud data security in storage and accessing. *Multimed Tools Appl* **82**, 42817–42832 (2023). <https://doi.org/10.1007/s11042-023-15463-1>
10. Li, J., Wang, Z., Ye, C., et al.: A semi-quantum private comparison with high-level security third party. *Int. J. Theor. Phys.* **62**, 196 (2023). <https://doi.org/10.1007/s10773-023-05452-3>
11. Hafiz, M.W., Hwang, S.O.: A probabilistic model of quantum states for classical data security. *Front. Phys.* **18**, 51304 (2023). <https://doi.org/10.1007/s11467-023-1293-3>
12. Temel, M.H., Škorić, B.: Entropically secure encryption with faster key expansion. *Quantum Inf. Process.* **23**, 119 (2024). <https://doi.org/10.1007/s11128-024-04330-z>
13. Haimes, Y. Y.: *Risk Modeling, Assessment, and Management*. John Wiley & Sons (2011)
14. Roth, M., Hartoyo, A., Liggesmeyer, P.: Efficient reachability graph development for qualitative analysis of state/event fault trees. In: 2015 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), pp. 144–151 (2015)
15. Li, Z., Ren Y., Liu, L., Wang, Z.: Improving parallel modularization algorithm of large complex fault trees. In: 2016 Annual Reliability and Maintainability Symposium (RAMS), pp. 1–6 (2016)
16. Butoi, A., Tomai, N.: Secret sharing scheme for data confidentiality preserving in a public-private hybrid cloud storage approach, utility and cloud computing (UCC). In: 2014 IEEE/ACM 7th International Conference on utility and cloud computing, pp. 992–997. <https://doi.org/10.1109/UCC.2014.163>, <https://doi.org/10.1109/UCC.2014.163> (2014)
17. Silva, G.S.M., Droguett, E.L.: Quantum fault trees and minimal cut sets identification (2024). arXiv preprint [arXiv:2404.05853](https://arxiv.org/abs/2404.05853)
18. Silva, G.S.M., Parhizkar, T., Droguett, E.L.: Quantum fault trees (2022). arXiv preprint [arXiv:2204.10877](https://arxiv.org/abs/2204.10877)
19. Tan, S.H., Kettlewell, J., Ouyang, Y., et al.: A quantum approach to homomorphic encryption. *Sci. Rep.* **6**, 33467 (2016). <https://doi.org/10.1038/srep33467>
20. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing, STOC '09, pp. 169–178, ACM, New York, NY, USA (2009)
21. Online source: <https://quantumalgorithms.org/chap-classical-data-quantum-computers.html>. Last Accessed 19 Apr 2024
22. Khrennikov, A., Klein, M., Mor, T.: Quantum integers. In: AIP Conference Proceedings, American Institute of Physics 1232(1) (2010)

Chapter 20

New Real-Time Encryption Mechanism for Financial Transactions



Ionuț-Eduard Matei , Costin-Radu Boldea , and Bogdan-Boldea 

Abstract There is a need for high security of information in financial transactions. This security is assured using various encryption algorithms such as RSA, DES, AES, and, in last time, with hybrid encryption algorithms. The problem arises when financial transactions take place in the banking sector involving large amount of data, involving the need for very fast, even real-time, and strong secured encryption, transmission and decryption algorithms. In this paper we proposed a new cryptography system for sensitive financial (bank or stock exchange) data that uses a combination of logistic map and sigmoid function for generating a secret key and a version of ElGamal cryptosystem based on elliptic curves for the transfer of the secret key online. This approach assures higher efficiencies in term of fast and strongly secured performance.

20.1 Introduction

An efficient encryption mechanism for financial transactions in present time needs to use strong symmetric key encryption algorithms, such as AES (Advanced Encryption Standard [1, 2]) or TDES (Triple Data Encryption System [3]) to ensure the confidentiality and integrity of the data being transmitted. The encryption mechanism should also incorporate key management techniques to securely generate, distribute, and store encryption keys.

I.-E. Matei · Bogdan-Boldea
University of Timisoara, Timisoara, Romania
e-mail: ionut.matei@e-uvt.ro

Bogdan-Boldea
e-mail: bogdan.boldea@e-uvt.ro

C.-R. Boldea (✉)
University of Craiova, Craiova, Romania
e-mail: cboldea@inf.ucv.ro

Furthermore, the encryption mechanism should be optimized for performance to minimize overhead and latency in encrypting and decrypting the data during transactions. This can be achieved by using hardware-accelerated encryption processors or implementing efficient encryption algorithms that are well-suited for fast processing.

In addition, the encryption mechanism should also support secure communication protocols such as TLS (Transport Layer Security [4]) to establish secure connections between clients and servers, as well as ensure data authenticity and protection against man-in-the-middle attacks.

Overall, an efficient encryption mechanism for financial transactions should prioritize security, performance, and scalability to ensure the safe and swift transmission of sensitive financial information.

Currently, the most used cryptography system for encrypted communication of sensitive financial data are the asymmetric key encryption models based on public and private key, as the RSA [5] or ECC (Elliptic Curve Cryptography, [6]). Protecting the public key becomes crucial in a public key cryptography system because anybody can use it, leaving information that could be exploited by an attacker. In order to provide more secure communication, related researchers have proposed new techniques for protecting public keys and signature keys, arriving to hybrid key encryption algorithms that use first a symmetric key encryption of data and, after, a secured transmission of the key using asymmetric key (public/private) methods (see [7] for a survey).

These days, a lot of strong secured encryption is done using chaos-based cryptography [8, 9]. It is thought that chaos is a safe way to introduce randomness into consistent data. Sensitivity to the initial condition, replication, ergodicity, non-periodicity, and pseudo-randomness are all features of chaotic systems. Furthermore, it is possible to quickly and reliably produce chaotic sequences [10] useful for strong and efficient encryption algorithms.

Although very random and secure keys are provided by chaotic maps, the use of secret keys is equally essential for a strong cryptosystem. The cryptosystem's resilience is guaranteed by the unpredictability of its output data and its defense against cryptographic intrusions. Numerous encryption algorithms with longer execution times and poor security exist [11, 12].

In this work we propose a new chaotic hybrid cryptography system, based on combination of logistic map and sigmoid function for random secret key generator for encryption of Brut financial data, and the ElGamal cryptosystem based on elliptic curves for the transfer of the secret key between the users in the network.

The rest of the paper is organized as follows: Sect. 20.2 introduces some fundamental concepts; Sect. 20.3 shows the encryption structure proposed by us, with the description of the new encryption logistic-sigmoid algorithm; Sect. 20.4 presents the comparative experimental analysis between our method and some other classical encryption algorithm from the point of view of vulnerabilities and speed; and last Section concludes with a statement.

20.2 Preliminaries

20.2.1 Key Exchange Across the Network Using Elliptic Curves

Lets $p \geq 3$ a prime number. An elliptic curve E defined over \mathbb{F}_p represents a set of points of the form $(x, y) \in E(\mathbb{F}_p)$, which satisfy the following equation

$$E_{a,b} : y^2 = x^3 + ax + b(\text{mod } p) \tag{20.1}$$

The set of points on the elliptic curve E , are defined as follows

$$E_{a,b}(\mathbb{F}_p) = \{(x, y) \in \mathbb{F}_p^2 | y^2 = x^3 + ax + b(\text{mod } p)\} \cup \{\mathcal{O}\} \tag{20.2}$$

where \mathcal{O} represents the neutral element of the elliptic curve, called the point at infinity. Elliptic curves are used in cryptography, for asymmetric encryption systems when $|E(\mathbb{F}_p)| = h \cdot q$, where q is a very large prime number and $1 \leq h \leq 4$ (Table 20.1).

ElGamal scheme based on elliptic curves, shown in the table above is used to send encrypted information over the network. To improve the encryption speed and thus the transfer speed of financial data, we will use the ElGamal encryption scheme based on elliptic curves to transfer the secret key for the proposed encryption system based on chaos theory. Elliptic curve based asymmetric encryption systems provides increased security and are based on the Elliptic Curve Discrete Logarithm Problem (ECDLP) [13].

Moreover, the keys used in the encryption process are much smaller than in the case of asymmetric cryptosystems such as RSA [14]. To transfer the secret key m used for encryption-decryption process, it is necessary to encode the binary message m as a point on the elliptic curve. For the process of encoding or compressing data in the form of points on an elliptic curve, we will use Koblitz’s algorithm, which we will describe below.

Table 20.1 Encryption/decryption using EC ElGamal scheme [13]

Public parameters	E an elliptic curve defined over \mathbb{F}_q , with G a base point
Bob private key	$s \in [1, E(\mathbb{F}_q) - 1]$
Bob public key	$P_B = sG$
Alice want to send to bob a message which is a point of E , P_m	
Encryption	Alice chooses an integer k and send to Bob the pair $(kG, P_m + kP_B)$
Decryption	Bob compute $P_m + kP_B - skG = P_m$

Koblitz's Algorithm [13]

Require: m a binary message, \mathbb{F}_p , a, b, k

Ensure: the pair (x, y) that represents binary message as point of $E_{a,b}: y^2 = x^3 + ax + b(\text{mod}p)$

1. $s \leftarrow 1$
2. **while** $s \leq k$ **do**
 - a. $\tilde{x} \leftarrow m \cdot k + s$
 - b. express \tilde{x} with l digits, as m_{l-1}, \dots, m_1, m_0
 - c. compute $x = \sum_{i=0}^{l-1} m_i \cdot g_i \in \mathbb{F}_p$, where $g \in \mathbb{F}_p$ is a generator
 - d. evaluates $y^2 = f(x) \equiv x^3 + ax + b(\text{mod}p)$ and compute $\sqrt{f(x)}$
 - e. if there exists such y , that satisfy the equation $y^2 = f(x)$, return $P_m = (x, y)$
 - f. $s \leftarrow s + 1$
3. **end while**

Koblitz's algorithm is a probabilistic algorithm, whose success rate depends on the value chosen for k . For a very small error probability for encoding binary messages, it is required that $k \geq 30$. The above algorithm computes $\tilde{x} = m \cdot k + s$, for all $s = 1 \dots k$, and checks whether $f(x)$ is quadratic residue modulo p . If $f(x)$ is a quadratic residue, it is computed $y \equiv \sqrt{f(x)}(\text{mod}p)$ and return the point $P_m(x, y)$ corresponding to the binary message m . If the correct value for $\sqrt{f(x)}$ is not determined, the algorithm increments the value of s and continues the process until it finds $y \equiv \sqrt{f(x)}(\text{mod}p)$. If $s \geq k$, the algorithm failed to encode the binary message as a point on the elliptic curve, and we will choose a higher value for k . After changing the value of k , the algorithm will restart the while loop for $s = 1$. The decoding of the binary message m , taking into consideration the point $P_m(x, y)$ representing the binary message encoded as a point on the elliptic curve, is performed by calculating [15], [16] $m \equiv \frac{x-1}{k}(\text{mod}p)$.

20.2.2 Logistic-Sigmoid Composed Map

In this section, we propose a combination of the logistic map and the sigmoid function, to obtain a pseudo-random sequence generator, used for the secret key derivation process, corresponding to the presented stream cipher. We define logistic map [17] for $\lambda \in (3,4)$ and for initial parameter $x_0 \in (0,1)$, as follow

$$x_{n+1} = \lambda x_n(1 - x_n) \quad (20.3)$$

It can be seen that $x_n \in (0,1)$, for every $n \geq 0$. Further, we define the sigmoid function [18] $s: \mathbb{R} \rightarrow (0,1)$, such that

$$s(x) = \frac{1}{1 + e^{-x}}, \text{ for } x \in \mathbb{R} \quad (20.4)$$

Thus, we will consider the compound Logistic-Sigmoid function, denoted by G , defined as follows

$$G(s) = \lambda s(1 - s), \text{ where} \quad (20.5)$$

$$s(x) = \frac{1}{1 + e^{-x}} \in (0, 1) \quad (20.6)$$

First, to be able to analyze the chaotic effect of this combination of functions, we will calculate the bounds for $x = \pm\infty$, so we have

$$\lim_{x \rightarrow -\infty} G(s(x)) = \lim_{x \rightarrow -\infty} \frac{\lambda e^{-x}}{(1 + e^{-x})^2} = 0 \quad (20.7)$$

$$\lim_{x \rightarrow \infty} G(s(x)) = \lim_{x \rightarrow \infty} \frac{\lambda e^{-x}}{(1 + e^{-x})^2} = 1 \quad (20.8)$$

To prove that the orbit $(s_n)_{n \geq 0}$ is defined in $(0,1)$, we assume that $s = \lim_{n \rightarrow \infty} s_n$ exists, with $s < \infty$ and we have

$$s_{n+1} = \lambda s_n(1 - s_n) \quad (20.9)$$

$$s = \lambda s(1 - s) \quad (20.10)$$

$$s = \frac{\lambda - 1}{\lambda} \quad (20.11)$$

Knowing that $\lambda \in (3,4)$, one can see that $s \in (0,1)$, therefore $(s_n)_{n \geq 0}$ is defined in $(0,1)$, for $s_0 = s(x)$, where x is the initial parameter of the system. To analyze the chaotic effect that the function G presents, we will first compute the fixed points of the equation

$$G(s) = s \quad (20.12)$$

$$\lambda s(1 - s) = s \quad (20.13)$$

$$-\lambda s^2 + (\lambda - 1)s = 0, \text{ with solutions} \quad (20.14)$$

$$s_1 = 0 \text{ and } s_2 = \frac{\lambda - 1}{\lambda} \quad (20.15)$$

To check the stability of points s_1 , respectively s_2 , we use the derivative of the function G , given as follows

$$G'(s) = \lambda(1 - 2s) \tag{20.16}$$

Since $\lambda > 3$, it follows that $G'(s_1) = \lambda > 1$, which means that the point s_1 is unstable. For the second point s_2 , we calculate $G'(s_2) = \lambda(1 - 2 \cdot \frac{\lambda-1}{\lambda}) = \lambda - 2\lambda + 2 = 2 - \lambda > 1$. The fixed points of the function G , for $\lambda \in (3,4)$, are unstable, therefore the composite map is chaotic, with infinite period. To be able to determine that the composite function G is chaotic, in the software implementation, we use the Lyapunov exponent λ_L , computed with a very good accuracy, as follow

$$\lambda_L = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \ln|G'(s_i)| \tag{20.17}$$

$$= \ln\lambda + \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \ln|1 - 2s_i| \tag{20.18}$$

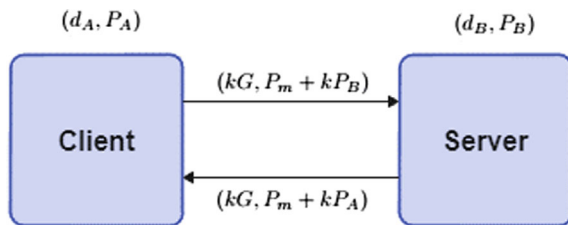
20.3 Proposed Encryption Scheme

In this section, we present a formalism for the transfer of secret keys used for encrypting financial transactions, using elliptic curves. Additionally, we present a new chaos based encryption scheme and describe how it can be implemented.

20.3.1 Secure Key Exchange Across the Network

Asymmetric encryption systems are generally slower in terms of execution time compared to symmetric ciphers. For network communication, symmetric ciphers cannot be used because there is no means to protect the keys used in the encryption-decryption process. Therefore, the use of a hybrid encryption scheme necessary. Thus, for network communication, public key encryption systems, especially those based on elliptic curves, can be used to transfer the secret keys, necessary to the information encryption (Fig. 20.1).

Fig. 20.1 Secure key exchange over the network



Lets $p \geq 3$ an odd prime and an elliptic curve E defined over \mathbb{F}_p . For the parameters $a, b \in \mathbb{F}_p$, the equation of the elliptic curve E is given as

$$E_{a,b} : y^2 = x^3 + ax + b(\text{mod } p) \quad (20.19)$$

In the above figure, the client generates $d_A \in [1, |E(\mathbb{F}_p)| - 1]$ and calculates the public key $P_A = d_A G$. Similar to the client, the server generates $d_B \in [1, |E(\mathbb{F}_p)| - 1]$ and obtains the public key $P_B = d_B G$, where G is the generator point on the elliptic curve E . Both entities shares their public keys. To send a generated secret key m , the client will encode the message m as a point $P_m(x_m, y_m)$ on the elliptic curve, using the algorithm 2. For a very low probability of error, we will set the value of the parameter $k = 64$, required for encoding the binary message m using the Koblitz algorithm. On the client side, a pseudo-random integer k is generated in the range $[1, |E(\mathbb{F}_p)| - 1]$, and then kG is computed, where G represents the generator point on the elliptic curve E .

Using the ElGamal encryption scheme presented in Sect 2.1, the client encrypts the point P_m and obtains $P_m + kP_B$, then sends the pair $(kG, P_m + kP_B)$ to the server. After receiving the encrypted message from the client given by the pair $(kG, P_m + kP_B)$, the server calculates $C_m = P_m + kP_B - d_B(kG) = P_m + kP_B - k(d_B G) = P_m$ and obtains the message encoded as a point on the elliptic curve. For increased security, the server stores the cipher binary message m encoded as a point C_m on the elliptic curve. To obtain the binary message m , the calculation is done as $m \equiv \frac{x_m - 1}{k} (\text{mod } p)$. If the client requests the server for the secret key used in the encryption scheme for the encryption-decryption process, the server encodes the binary message m as a point P_m on the elliptic curve. Then, the server thus generates an integer $k \in [1, |E(F)_p| - 1]$ and sends the pair $(kG, P_m + kP_A)$ to the client. As with the server, the client then decrypts the received secret key m $P_m = C_m - d_A(kG) = P_m + kP_A - k(d_A G)$ and then recover secret key m by decoding the point P_m . Considering that we use 256-bit secret key for data encryption, we recommend for implementation, the use of an elliptic curve defined over a prime field encoded on at least 384 bits.

For the software implementation of the secured transfer scheme for key exchange over the network, we use the NIST P-384 [16] standardized elliptic curve.

20.3.2 Our Encryption Algorithm

In this section, we present the proposed algorithm for the encryption-decryption process of financial data. Our encryption scheme is a chaotic stream cipher, which uses a combination of the logistic map and the sigmoid function, necessary in the subkey generation process. The use of this combination of non-linear functions contributes to the generation of pseudo-random sequences, which have strong cryptographic properties. In the figure below, the entire workflow of the encryption scheme is presented (Fig. 20.2).

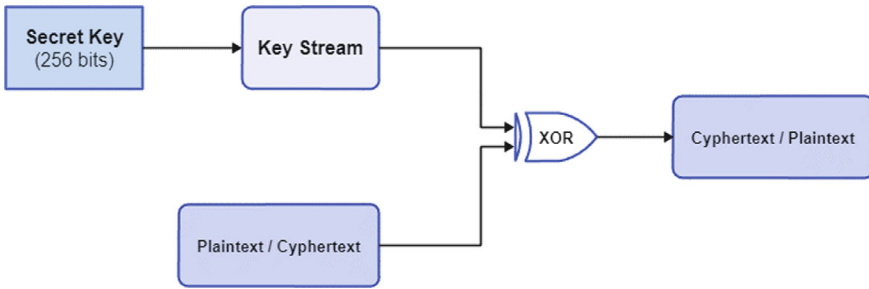


Fig. 20.2 Proposed encryption scheme

In order to achieve the encryption and decryption of financial transactions, we will group the software implementation in three main stages:

1. Secret key preprocessing
2. Key derivation mechanism
3. Encryption/Decryption.

20.3.2.1 Secret Key Preprocessing

In order to be able to use the secret key as the initial parameter x_0 , used for the key derivation mechanism, an appropriate transformation is required. Let k be the symmetric key used for the encryption-decryption process. If k is encoded on a power-of-two number of bits of the form 2^j , where j is an even integer, we obtain the pair of integers (k_1, k_2) formed by binary segmentation of the key. Next, we determine the parameter τ and compute x_0 by using the sigmoid function [18], as follows

$$\tau = \left\lfloor \frac{k_1}{k_2} \right\rfloor \tag{20.20}$$

$$x_0 = \frac{1}{1 + e^{-\tau}} \tag{20.21}$$

20.3.2.2 Key Derivation Mechanism

The proposed mechanism for the key derivation process uses the logistic map, defined as follows

$$x_{n+1} = rx_n(1 - x_n) \tag{20.22}$$

where the parameter $r \in (3,4)$ and the initial value $x_0 \in (0,1)$, that was computed previously.

To obtain a pseudo-random sequence that has strong cryptographic properties, we set the value $r = 3,915$. Using the parameters r and x_0 for the key derivation mechanism, with the help of the logistic map, we determine subkeys of variable length in bytes as follows

$$y_i = ry_{i-1}(1 - y_{i-1}) \quad (20.23)$$

$$SK_{i-1} = (y_i \cdot 10^{18})(\text{mod } 256) \quad (20.24)$$

for $i = 1, \dots, l$, where $y_0 = x_0$ and SK_l represents a subkey of length l bytes. Next, we will present the algorithm used for both encryption and decryption, due to the symmetry of the encryption scheme.

20.3.2.3 Encryption/Decryption

In the previous subsections, we described how the initial symmetric key is transformed and presented the subkey derivation mechanism, used in the encryption-decryption scheme.

Encryption/Decryption algorithm

Require: *data* a data block of length l given in bytes.

Ensure: *Encrypted/Decrypted data* m

1. Compute τ, x_0 using key preprocessing described in 3.2.1
2. $seed \leftarrow x_0, r \leftarrow 3,915$
3. **for** $i \leftarrow 0$ **to** $l - 1$ **do**
 - a. $seed \leftarrow r \cdot seed \cdot (1 - seed)$
 - b. $unit \leftarrow seed \cdot 10^{18}(\text{mod } 256)$
 - c. $m_i \leftarrow data_i \oplus unit$
4. **return** m

The above algorithm uses the key derivation mechanism, presented in 3.2.2, to construct symmetric sub-keys required in the encryption-decryption process. To achieve increased computing speed, the algorithm generates iterative the byte blocks of the key and then encrypts and decrypts the information using the XOR operation.

20.4 Experiments and Analysis

In this section, we present a comparison between our encryption scheme and some existing methods used for financial transactions encryption. We implemented our encryption scheme for financial transactions in C# programming language, using

the.NET 8 framework. For implementation, we used an AMD Ryzen 5 3600 6-Core computer (3.59 GHz) with 16 GB of RAM (Table 20.2).

For comparison, we used the block ciphers AES [2], BlowFish [15], Triple DES [3] and the RSA public key encryption system [5]. For the comparison with existing encryption algorithms, implemented in the.NET 8 Framework, we use the public dataset, taken from the Bucharest Stock Exchange [19]. In the Table 20.2, the first column describes the name of encryption scheme and in the second the length in bits of the keys used in the encryption-decryption process.

To improve the accuracy of the tests performed, for each of the implemented algorithms, we determine the calculation time, expressed in the total number of seconds, and then calculate the speed expressed in MB/s as follows

$$\text{Speed} = \frac{M}{T \cdot 2^{20}} \quad (20.25)$$

where M represents the encrypted binary message and T is the real number, that express the total number of seconds corresponding to the test.

On the last column, we used Shannon's function to evaluate the entropy of the ciphertexts. Due to the fact that the presented method presents an entropy value close to the upper bound, on the interval (0,8), it demonstrates resistance against chosen-plaintext attack. For a given ciphertext, denoted C of length l , we calculate the entropy using Shannon's function, as follows

$$H(C) = \sum_{i=0}^l C_i \cdot \log_2(C_i) \quad (20.26)$$

In the case of the **AES** block cipher, no major vulnerabilities have yet been found, same as is the case with the **3DES** and **BlowFish** ciphers. However, cryptanalytic studies of the **AES** cipher can lead to the development of differential cryptanalysis methods that can compromise encrypted data. The Triple DES (**3DES**) block cipher provides lower security than the symmetric **AES** algorithm and our algorithm, due to the maximum secret key size, that is bounded at 192 bits.

Although entropy is not best metric to measure the security level of an encryption system, in the case of the proposed algorithm, the use of the ElGamal cryptosystem based on elliptic curves for the transfer of the secret key in the network considerably increases the security level of the proposed stream cipher. Also, the very large period

Table 20.2 Evaluation of existing encryption schemes

Method	Key size	Encryption speed (MB/s)	Entropy
AES	256	72.05	3.87500000
RSA	15,360	2.15	7.99999786
BlowFish	256	4.974	7.99999534
3DES	192	5.82	7.99999520
Logistic-sigmoid	256	6.18	7.99999648

of the pseudo-random sequence generator, that uses a combination of logistic map and sigmoid function proves that our symmetric algorithm provides high security.

In the Table 20.2, it can be seen that our algorithm (**Logistic-Sigmoid**) outperforms both the **RSA** public key encryption system and the Triple DES (**3DES**) symmetric cipher in terms of encryption speed. Therefore, our symmetric cipher based on chaos theory can represent, an alternative encryption system that can be used to encrypt financial transactions.

20.5 Conclusion and Future Work

In this paper, we proposed a new chaos-based encryption scheme that uses a combination of logistic map and sigmoid function. We also presented a formalism for the transfer of secret keys, necessary for encrypting of financial transactions, using the ElGamal cryptosystem based on elliptic curves. Compared to the AES block cipher, our algorithm does not provide hardware acceleration, therefore the encryption speed is lower.

Possible improvements to the encryption scheme can be made, including the partial implementation in assembly language of the routines used for atomic operations on real numbers in double precision, or the use of SIMD instructions, for parallel data encryption.

References

1. Daemen, J., Rijmen, V.: The Design of Rijndael: AES—The Advanced Encryption Standard. Springer Science & Business Media (2002)
2. Dworkin, M.J., Barker, E.B., Nechvatal, J.R., Fodi, J., Bassham, L.E., Roback, E., Dray Jr, J.F.: Advanced encryption standard (AES). Federal Information Processing Standards (NIST FIPS), National Institute of Standards and Technology, Gaithersburg, MD (2001). <https://doi.org/10.6028/NIST.FIPS.197>
3. De Cannière, C.: Triple-DES, in Encyclopedia of Cryptography and Security. Boston, MA, USA: Springer (2005). https://doi.org/10.1007/0-387-23483-7_437
4. Heinrich, C.: Transport layer security (TLS). In: van Tilborg, H.C.A., Jajodia, S. (eds.) Encyclopedia of Cryptography and Security. Springer, Boston, MA (2011). https://doi.org/10.1007/978-1-4419-5906-5_234
5. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**(2), 120–126 (1978)
6. Miller, V.S.: Use of elliptic curves in cryptography. In: Conference on the Theory and Application of Cryptographic Techniques, pp. 417–426. Berlin, Heidelberg. Springer Berlin Heidelberg (1985)
7. Mousavi, S.K., Ghaffari, A., Besharat, S., Afshari, H.: Security of internet of things based on cryptographic algorithms: a survey. Wirel. Netw. **27**(2), 1515–1555 (2021). <https://doi.org/10.1007/s11276-020-02535-5>
8. Kocarev, L.: Chaos-based cryptography: a brief overview. IEEE Circuits Syst. Mag. 1(3):6–21 (2001)

9. Wu, W., Peng, H., Wen, L., Liu, Y., Fenghua Tong, F., Lixiang Li, L.: A secure and efficient data transmission method with multilevel concealment function based on chaotic compressive sensing. *IEEE Sens. J.* **23**(17), 19823–19841 (2023)
10. Hanouti, I.E., Fadili, H.E., Souhail, W., Masood, F.: A lightweight pseudo-random number generator based on a robust chaotic map. In *Proceedings of the 4th International Conference on Intelligent Computing Data Sciences (ICDS)*, pp. 1–6 (2021)
11. Munir, N., Khan, M., Shah, T., Alanazi, A.S., Hussain, I.: Cryptanalysis of nonlinear confusion component based encryption algorithm. *Integration* **79**, 41–47 (2021)
12. Churcher, A., Ullah, R., Ahmad, J., Rehman, S., Masood, F., Gogate, M., Alqahtani, F., Nour, B., Buchanan, W.J.: An experimental analysis of attack classification using machine learning in IoT networks. *Sensors* **21**(2), 446 (2021)
13. Koblitz, N.: *A course in number theory and cryptography*, vol. 114. Springer Science & Business Media (1994)
14. Saho, N.J.G., Ezin, E.C.: Comparative study on the performance of elliptic curve cryptography algorithms with cryptography through RSA algorithm. In: *CARI 2020-Colloque Africain sur la Recherche en Informatique et en Mathématiques Appliquées* (2020)
15. Schneier, B.: Description of a new variable-length key, 64-bit block cipher (Blowfish). In: *International Workshop on Fast Software Encryption*, pp. 191–204. Berlin, Heidelberg. Springer Berlin Heidelberg (1993)
16. Chen, L., Moody, D., Regenscheid, A., Randall, K.: Draft NIST special publication 800–186 recommendations for discrete logarithm-based cryptography: elliptic curve domain parameters. National Institute of Standards and Technology, Gaithersburg, MD, USA (2019)
17. May, R.M.: Simple mathematical models with very complicated dynamics. *Nature* **261**(5560), 459–467 (1976)
18. Han, J., Moraga, C.: The influence of the sigmoid function parameters on the speed of back-propagation learning. In: *International Workshop on Artificial Neural Networks*, pp. 195–201. Berlin, Heidelberg. Springer Berlin Heidelberg (1995)
19. Bursa de valori Bucuresti-Date MiFiD II. <https://www.bvb.ro/TradingAndStatistics/Trading/MiFIDIIData>. Last accessed 30 Mar 2024

Chapter 21

Enhanced Blockchain-Based e-Voting System Using Zero-Knowledge Proofs



Marcela Lupu and Iulian Aciobăniței 

Abstract In the rapidly evolving digital landscape, electronic voting (e-voting) systems have the potential to offer improved accessibility and enhanced security. This paper introduces a novel e-voting solution deployed on the Ethereum blockchain, leveraging the strengths of zero-knowledge proofs (ZKPs). This approach ensures voter anonymity, prevents double voting, and verifies voter eligibility. Our publicly available implementation uses the zk-SNARK protocol, a form of ZKP, to securely verify voter eligibility across both local and national elections. To contextualize our contribution, we conducted a comparative analysis against comparable solutions. This comparison focused on gas costs associated with blockchain operations and the range of supported functionalities. Our system stands out by incorporating a “voting sandbox” feature, which features the ability for a user to prove eligibility for the vote session, through different criteria, while also maintaining anonymity.

21.1 Introduction

Traditional voting systems are still the preferred choice of communities around the world because of the trust they have in the current system, which is considered more secure in terms of voting integrity and anonymity of participants. This option implies trust in the electoral authorities and operators. On the other hand, organizing elections is costly and requires a lot of financial and human resources, as well as a longer period of time to obtain the final results.

Over time, various solutions have been proposed to implement e-voting, using a variety of cryptographic methods, such as homomorphic encryption that allows votes to be counted without decrypting the data, blind signatures, which offer the possibility

M. Lupu (✉) · I. Aciobăniței
Faculty of Information Systems and Cyber Security, Military Technical Academy Ferdinand I,
Bucharest, Romania
e-mail: marcela.lupu@mta.ro

I. Aciobăniței
e-mail: iulian.aciobanitei@mta.ro

of authenticating voters without exposing the content of the vote, or mixnets, designed to mask the source of a vote in order to eliminate any direct link between the voter's identity and their choice [1]. In order to facilitate the systematization of elections and to increase voter turnout, there has been interest from governments in addressing electronic voting systems. Standards and legislative frameworks on e-voting have also been developed, such as CM/Rec(2017)5 which contains a set of best practices developed to serve governments wishing to adopt such systems [2].

Estonia is the first country to introduce such a system, which is still in use today, noting in recent years an increase in participants using the online platform. In order not to abandon the traditional method, it has considered a mixed approach to elections, with voting being possible both online and physically at a polling station. Moreover, users can change their choice made until the end of the online voting session, even physically by showing up at the polling station, and the final results are only transmitted after the election is concluded. The Estonian system allows users to authenticate and cast votes based on government-issued electronic ID cards and a digital signature. Some of the main features of the system include user eligibility, data auditing using digital certificates, and the confidentiality and anonymization of voters by removing the digital signature from the ballots before they reach the electoral committee for counting [3]. An analysis was conducted on the system during the 2013 elections, and critical issues were discovered in its architecture. One of them is that the solution does not ensure end-to-end (E2E) verifiability [4].

Although it is a controversial topic, research in the field of electronic voting has increased substantially, and attention has been directed toward another technology that has gained prominence in recent years, namely blockchain. This is highlighted by a recent study of 252 articles that shows significant interest in the development of electronic voting systems using blockchain [5]. A critical analysis was performed on online voting systems, arguing the security risks involved, as well as the problems that integrating blockchain brings. They also proposed possible solutions, among which is the use of a new innovative concept called zero-knowledge proof. Additionally, there are nations that have conducted smaller experiments on their systems implemented with blockchain, such as the United States, Russia, Japan, and Switzerland [6].

The applicability of zero-knowledge proof in various systems, both centralized and decentralized, is still a topic under research and has not yet been integrated into any system for national use. Many of the existing works on ZKP provide abstract and theoretical analyses of the concepts but do not present details of implementation or technical constraints on the practical use of tools that ensure the development of applications based on ZKP. This paper aims to implement an electronic voting application based on blockchain technology that explores and validates the feasibility of ZKP concepts through the zk-SNARK protocol.

21.2 State of the Art

Existing papers published over the years have described various approaches to implementing an e-voting solution, based on classical cryptographic methods, blockchain, and more recently, they have also incorporated zero-knowledge proof protocols. Researchers have focused on decentralized solutions that ensure transparency and integrity, which, on the other hand, do not guarantee the confidentiality of the voter or the vote, as noted in works [5, 7, 8]. The integration of ZKP has contributed to solving problems introduced by blockchain and has been explored in various solutions that will be described in the exemplification of the following works, highlighting their applicability and relevance for electronic voting.

1. **Open Vote Network** [9] is one of the first electronic voting solutions built with blockchain technology using smart contracts. It stands out for its ability to operate without relying on a trusted authority for calculating the final results or for protecting the identity of the users. ZKP is used in the first stage to verify the secret associated with each voter's private key. In the second stage, it is used to confirm the eligibility of the votes (checks if their value is zero or one).
2. **ElectAnon** is a solution that stands out by using the Semaphore¹ protocol for implementing ZKP. It ensures that a voter is eligible to vote and that they have voted only once, without revealing any other information about their identity. Furthermore, the author of the work proposed an additional level of security by generating proofs constructed with the Merkle Forest extension [10].
3. **S. Panja and B.K. Roy** propose an electronic voting system implemented with blockchain and zero-knowledge proofs to guarantee transparency, integrity, and confidentiality of votes. ZKP is used to demonstrate the validity of votes without revealing the identity of the candidates, thus ensuring the anonymity of the electoral process [11].
4. **EthVote** proposes a solution based on the Ethereum blockchain and combines Paillier homomorphic encryption to encrypt and count votes. The system proposes an algorithm for verifying the integrity of votes after their decryption, which, however, the author did not implement due to the limitations of the Zokrates tool regarding the language and libraries it offers [12].
5. **Murtaza** proposed another solution that stands out as it is built on the fundamental idea of Zerocoin to achieve anonymous transactions by "burning" coins, ensuring the anonymity of voters. This approach uses blockchain and ZKP to provide transparency and integrity, as well as to allow the separation of votes from the identity of election participants [13].

Other electronic voting systems have complemented blockchain and ZKP technologies with an additional homomorphic encryption mechanism, such as Hääl [14]. The work [15] introduces a method of identification that uses the zk-SNARK protocol to certify the identity of users without exposing personal data.

¹ <https://docs.semaphore.pse.dev/>.

Following the analysis of various specialized works, it was concluded that they use similar architectures and cryptographic schemes. Although blockchain technology is still a debated topic, its features make it a good choice for addressing certain security requirements. A large part of researchers have used zero-knowledge proof protocols to provide confidentiality and anonymity for voters or votes, as well as to verify the correctness of the votes. Additionally, some approaches have used homomorphic encryption to ensure vote confidentiality.

The electronic voting platform proposed in this work is based on the same principles. The solution ensures security criteria by applying zk-SNARK and additionally introduces a hybrid storage model to reduce transactional costs. Another relevant feature of the system is its flexibility regarding voting sessions, which are dynamic in nature. In this way, the platform allows for the conduct of electoral scrutinies at both local and national levels.

21.3 Proposed Solution

The voting system was designed to meet the minimum security requirements necessary for building an electronic voting platform, as follows:

1. The system shall allow user authentication, ensuring the uniqueness of each voter in the system and their eligibility to participate in the voting process.
2. Access to a voting session shall be allowed only to users eligible to participate in the voting session (within local elections).
3. The user shall be able to cast a vote within a voting session only once, without disclosing their identity or any personal information.
4. The system shall allow users to check and validate their vote and that it has been correctly recorded in the system.
5. The counting of votes shall be carried out without the need for intervention by a third party or external entity, and the results will be displayed after the election is completed.
6. The system ensures through its infrastructure that user votes are not modified or corrupted.
7. The system shall allow users to obtain the necessary proofs to be able to vote. These proofs will be issued by the electoral authority (which aims to generate the proof necessary for registering to a voting session), as well as by users to demonstrate their eligibility after registration.
8. The system shall allow the creation of elections and polls from different categories: both local or within an electoral district, as well as national. The management of electoral elections must be carried out by an administrator.

The foundation of our solution is the blockchain technology and the zk-SNARK protocol, to which is added a web server developed using the ASP.NET framework and a conventional SQL database for the persistent storage of voting contexts and the metadata of electoral sessions.

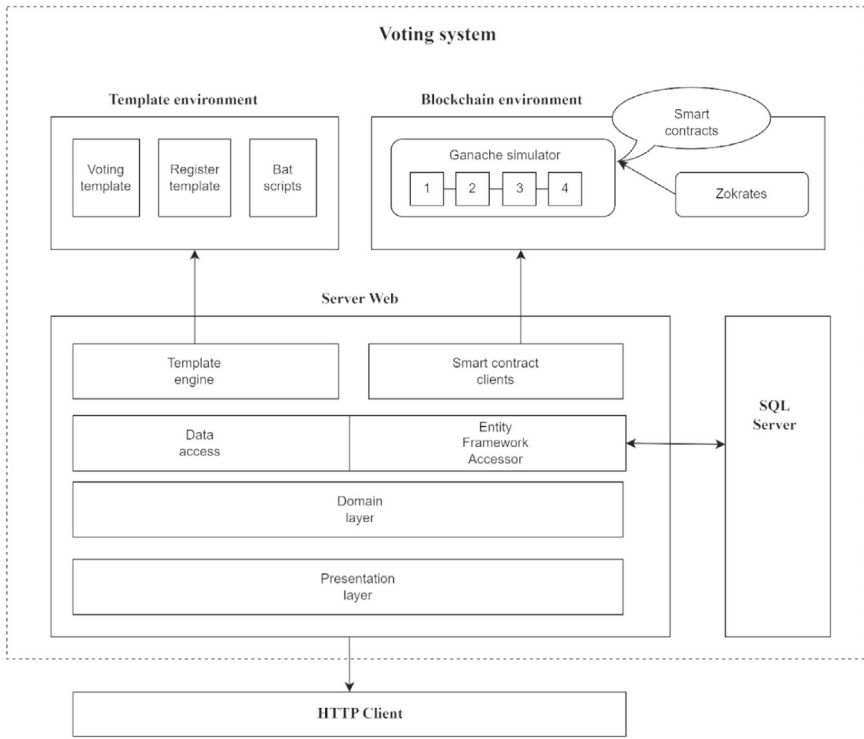


Fig. 21.1 System architecture

Figure 21.1 presents the proposed architecture of the solution, highlighting the key components and associated sub-components, as well as how they interact with each other.

The platform is structured on individual components with unique responsibilities. The core of the voting system are the smart contracts responsible for managing the logic of the voting process such as adding candidates, casting the vote, and calculating the results. The web server acts as an orchestrator for the rest of the components, managing both the creation of evidence for ZKP and the interaction with the smart contracts.

The blockchain environment is the infrastructure where the information needed for voting will be stored. The template environment is a module of the application that contains the scripts needed to generate the context of a voting session, generate metadata for initiating contracts on Ethereum, and create ZKP proofs. In addition, it also contains the rest of the elements needed to create the contexts for voting and registration.

Each voting session is characterized by dedicated smart contracts, specific data, and files or scripts used within the ZKP verification and proof creation mechanisms.

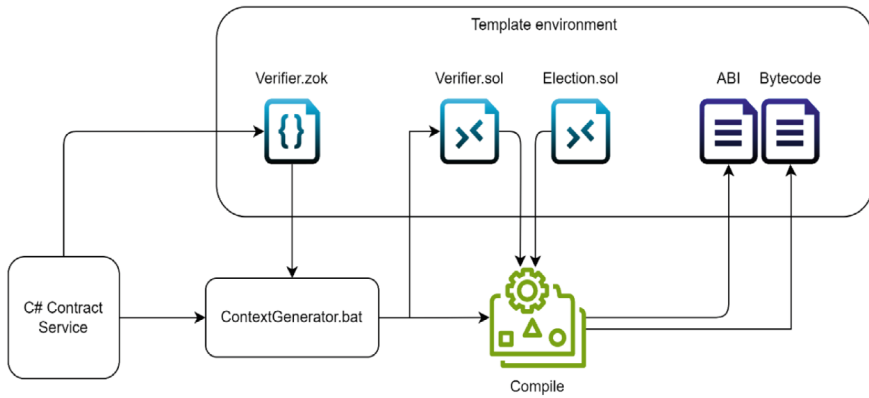


Fig. 21.2 Contract compilation stage

The operation flow of the smart contracts used in relation to the context of a voting session is divided into two stages: compilation and deployment.

The compilation stage is highlighted in Fig. 21.2. The C# client dynamically creates the Zokrates program. This is used as input for the script responsible for generating the smart contract for proof verification. The procedure concludes with the compilation of the smart contracts, resulting in the creation of the Application Binary Interface (ABI) and the associated bytecode. The second phase of the process consists of the actual deployment of the smart contract. In this way, the contract will be integrated into the Ganache testing environment and will set the foundation for future interactions with the smart contract.

ZKP was used in two stages to ensure the confidentiality and anonymity of the voter, as follows: in the registration stage, the ZKP proof was used to demonstrate that the user is of legal age to vote, and also that they belong to the relevant electoral district if applicable. In the second stage, ZKP is used as an additional layer of security to demonstrate that the user is eligible to vote by verifying that they are on the list of eligible users.

The implementation of blockchain technology was carried out using the Solidity language for creating smart contracts, the Ganache simulator for testing, and Truffle for compiling and migrating contracts to the local blockchain network. For the implementation of zk-SNARK, Zokrates was used, a set of tools that allows the definition of programs in a high-level language, simplifying the process of integrating ZKP with the rest of the application. With its help, proofs were also generated and verified to demonstrate the authenticity of users. Zokrates helps improve gas costs through several key mechanisms such as proof compaction, reducing on-chain interactions, and optimizing the logic of smart contracts. Integration with .NET for Ethereum was achieved with the help of the Nethereum library, which was used for managing and interacting with smart contracts.

21.3.1 Methods Description

The main methods of the proposed framework can be observed in Table 21.2, where we also present a comparative analysis with two other solutions, in regard to gas consumption. For a better understanding, we describe the two main operations of each method proposed.

Register. Voters are registered for voting. This method involves the following operations:

- Verification of the proof presented by a voter to register for a new voting session (verification involves validating if the user is of legal voting age and if the election is for a specific county, to check if the user belongs to the corresponding county). The registration proof is assumed to be received from a trusted electoral authority.
- Mark the attribute designed to guarantee the uniqueness of the proof and the fact that a user has already registered by setting the *usersRegistered*[< user *RegistrationProofSha* >] dictionary.

Vote. Actually casting the vote. This method also can be divided into multiple actions:

- Verify the candidate, which involves checking if it is a part of the candidates list.
- Verify if the candidate has already voted.
- Prevent double voting by marking *usersVoted*[< user *VotingProofSha* >] dictionary.
- Verify the ZKP proof. As an observation, this tends to be the most costly operation in the entire framework.

21.4 Results

Following the analysis of various systems in the specialized literature, it was noted that the minimum requirements a voting system considers are confidentiality, anonymity, transparency, and immutability, as well as the verifiability of the vote or voter. The electronic voting platform proposed in this work is founded on the same principles, ensuring these security criteria through the application of blockchain and ZKP. Table 21.2 provides a comparison between the solution proposed in this work and two other relevant studies in the field of interest, focusing on the requirements that a voting system shall meet, with an emphasis on security aspects.

Based on the results presented by these papers, a comparison was made between them and the work of this paper in terms of gas consumed for a number of 10 candidates and 40 voters. The costs for solution [9] were obtained from Table 21.1 within the paper, and for solution [10] were obtained from Table 4. To obtain the gas consumption for the current work, the Ganache tool was used. We used the values for the executed transactions (the Gas Used value from Ganache). Each smart contract method call represents a transaction in Ganache.

Table 21.1 Representative solutions comparison—functionalities

Implemented functionalities	[9]	[10]	Proposed solution
Unicity	✗	✓	✓
Eligibility	✓	✓	✓
Verifiability	✓	✓	✓
Correctness	✓	✓	✓
Anonymity	✗	✓	✓
Vote confidentiality	✓	✓	✓
Immutability	✓	✓	✓
Multiple roles	✗	✓	✓
Multiple regions	✗	✗	✓
Platform	Ethereum	Ethereum	Ethereum

Compared to the two studies, our solution is distinguished by a much lower gas consumption than that proposed by McCorry [9], but it is not as efficient as ElectAnon [10]. The registration and voting phases themselves represent the transactions with the highest costs, but which can be optimized by more careful use of the instructions used in Solidity. It is important to emphasize that all system testing was carried out in a local development environment, using Ganache, and was conducted using a single account.

It is important to understand the reason for the higher cost of the *Vote* method, in comparison to [10]. The most time and gas-consuming operation of the *Vote* method is actually the verification of the ZKP proof generated with Zokraes, which is done on-chain. On the other hand [10] uses a Merkle tree to define the list of eligible users and for checking the proof they use the Semaphore protocol. In short, we check on-chain the eligibility of the voter and this drives a higher cost (Table 21.2).

Following the analysis of transaction costs, it was concluded that the hybrid approach of outsourcing as much of the metadata of an election session as possible

Table 21.2 Representative solutions comparison—gas consumption

Method	[9]	[10]	Proposed solution
Deploy	6.215.811	3.430.754	4.751.148
Add voters	2.153.461	113.963	–
Add candidates	–	286.040	50.224
State change	–	71.877	23.960
Propose	–	42.681	–
Register	763.118	–	268.988
Commit	70.112	312.856	–
Vote	2.490.412	105.140	341.070
Tally	746.485	48.937	–
Total	145.381.870	20.812.181	29.727.558

outside of the blockchain helped reduce costs compared to if the entire system had been built on the blockchain. On the other hand, the total cost is still relatively high for the system to be used for large-scale elections.

From a security point of view, the ZKP implementation ensures the anonymity of the user, without revealing information that could associate the vote with the user's identity.

21.4.1 Discussion

Current section dives into a discussion regarding storing data on-chain and off-chain. Using everything on-chain clearly has the advantage of decentralization, hence leveraging the main blockchain security features. On the other hand, storing data off-chain brings the advantage of obtaining a feasible solution in regard to cost and processing time.

Storing data offline adds complexity to the system and can potentially introduce vulnerabilities, requiring extra measures for data protection. For instance, ensuring the secure transmission of data to the blockchain is essential. This involves implementing encryption protocols and using secure channels to protect sensitive information during transfer.

Moreover, because off-chain data is an additional part of the system, another challenge is to keep the blockchain with the relational database data synchronized. Otherwise, this can result in inconsistencies within the system. Another interesting topic is the file hierarchy generated based on templates. This might constitute a potential vulnerability. Although hash mechanisms or signatures can be used to ensure their integrity after generation, this can introduce implementation challenges or issues.

21.5 Conclusions

In this paper, we propose a new e-voting system based on the Ethereum blockchain and zero-knowledge proofs (ZKPs). We implemented the main requirements that each e-voting system shall fulfill. The paper details the implementation aspects, and the entire solution is accessible on Github.² Our comparative analysis further emphasizes the efficiency and robustness of our implementation, particularly in terms of operational gas costs and the breadth of functionalities it supports. A distinctive feature of our system is the “voting sandbox,” enabling tailored voting sessions for individual counties or nationwide.

Based on our measurements, we conclude that, despite its benefits, the current cost structure makes national-level implementation prohibitively expensive, necessitating

² <https://github.com/marcilupu/BlockchainVotingApp/tree/feature/Extend-ZKP>.

further optimizations. Moreover, our findings indicate that leveraging ZKPs not only enhances privacy but also significantly reduces blockchain storage requirements, leading to lower operational costs.

References

1. Mursi, M., Assassa, G.M.R., Abdelhafez, A., Samra, K.: On the development of electronic voting: a survey. *Int. J. Comput. Appl.* **61**, 1–11 (2013). <https://doi.org/10.5120/10009-4872>
2. Council of Europe adopts new Recommendation on Standards for E-Voting: https://www.coe.int/en/web/portal/news-2017/-/asset_publisher/StEvosr24HJ2/content/council-of-europe-adopts-new-recommendation-on-standards-for-e-voting. Accessed 09 Feb 2024
3. <https://e-estonia.com/solutions/e-governance/e-democracy/>
4. Springall, D., Finkenauer, T., Durumeric, Z., Kitcat, J., Hursti, H., MacAlpine, M., Halderman, J.A.: Security analysis of the estonian internet voting system. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14), Association for Computing Machinery, New York, NY, USA, 703–715 (2014). <https://estoniaevoting.org/>
5. Hajian Berenjestanaki, M., Barzegar, H.R., El Ioini, N., Pahl, C. : Blockchain-based E-voting systems: a technology review. *Electronics* **13**, 17 (2024). <https://doi.org/10.3390/electronics13010017>
6. Park, S., Specter, M., Narula, N., Rivest, R.L: Going from bad to worse: from Internet voting to blockchain voting. *J. Cybersecur.* **7**(1), tyaa025 (2021). <https://doi.org/10.1093/cybsec/tyaa025>
7. Lahane, A., Patel, J., Pathan, T., Potdar, P.: Blockchain technology based e-voting system. In: ITM Web of Conferences, vol. 32, p. 03001 (2020). <https://doi.org/10.1051/itmconf/20203203001>
8. Shah, A., Sodhia, N., Saha, S., Banerjee, S., Chavan, M.: Blockchain enabled online-voting system. In: ITM Web of Conferences (2020)
9. McCorry, P., Shahandashti, S.F., Hao, F.: A smart contract for boardroom voting with maximum voter privacy. In: Kiayias, A. (Ed.) *Financial Cryptography and Data Security*. Springer International Publishing, Cham, pp. 357–375 (2017)
10. Onur, C., Yurdakul, A.: ElectAnon: a blockchain-based, anonymous, robust, and scalable ranked-choice voting protocol. *Distrib. Ledger Technol.* **2**, 3, Article 19, 25 (2023). <https://doi.org/10.1145/3598302>
11. Panja, S., Roy, B.K.: A secure end-to-end verifiable e-voting system using zero knowledge based blockchain. *IACR Cryptol. ePrint Arch.* **466** (2018)
12. Mols, J., Vasilomanolakis, E.: ethVote: towards secure voting with distributed ledgers. In: International Conference on Cyber Security and Protection of Digital Services (Cyber Security), Dublin, Ireland, pp. 1–8 (2020)
13. Murtaza, M.H., Alizai, Z.A., Iqbal, Z.: Blockchain based anonymous voting system using zkSNARKs. In: 2019 International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 2019, pp. 209–214
14. Osorio, E.: Anonymous electronic voting system on public blockchains (2018) <https://github.com/eddieoz/haal7>
15. Mukherjee, A., Majumdar, S., Kolya, A., Nandi, S.: A privacy-preserving blockchain-based e-voting system (2023)

Part IV
Quantitative Economics

Chapter 22

The Environmental Kuznets Curve for Deforestation in Romania: An ARDL-Based Evaluation



Irina Georgescu , Ionuț Nica , and Nora Chiriță 

Abstract Regional economic development and innovation play a crucial role in sustaining sustainable economic growth and promoting social inclusion. Our study examines the long-run and short-run relationship between forest area growth (FAG) as the dependent variable and its determinants: Gross Domestic Product (GDP), energy consumption (EUSE), and urbanization (URB). The data used covers the period 1990–2021 and is extracted from the World Bank database, with the methodology applied being AutoRegressive Distributed Lag (ARDL). The results of the long-run estimations indicate a positive and significant influence of GDP on FAG, while urbanization and energy consumption have negative and significant effects. Additionally, using Python programming language, we constructed the Environmental Kuznets Curve for deforestation in Romania, which is highlighted by the inverted U-shaped relationship between GDP and FAG. Our findings suggest that urbanization can contribute to the expansion of FAG through urban planning practices that incorporate green spaces and forests in urban areas. Moreover, energy consumption can have a negative effect on FAG, being associated with industrial and energy production activities that generate pollution and deforestation.

22.1 Introduction

Deforestation stands out as one of the most pressing issues of the contemporary world, with profound consequences on the environment and society as a whole. In an effort to understand and address this issue, our study focuses on evaluating the relationship between deforestation and economic and social factors in the context of Romania. Over the past decades, Romania has witnessed significant economic and social transformation, and understanding the impact of this transformation on its

I. Georgescu · I. Nica (✉) · N. Chiriță
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: ionut.nica@csie.ase.ro

forests is essential for the development of sustainable environmental conservation policies and practices.

By employing a robust economic methodology and comprehensive data spanning from 1990 to 2021, our study aims to analyze how economic variables such as GDP, energy consumption, and urbanization influence the forest area in Romania. The results of this analysis can provide a solid foundation for the development of strategies and policies aimed at sustainable management of forest resources and biodiversity conservation.

Therefore, the primary objective of this study is to make a significant contribution to understanding the dynamics of deforestation in the specific economic and social context of Romania, offering relevant perspectives and recommendations for forest conservation practices and the country's sustainable development.

Additionally, this study will explore the concept of the Kuznets Curve in the context of deforestation in Romania. Although initially proposed to examine economic inequality, the Kuznets Curve offers an intriguing perspective on the relationship between economic growth and environmental impact. By applying this concept to the context of deforestation, we aim to investigate whether there is a critical stage in Romania's economic evolution where economic growth may begin to mitigate the negative effects of deforestation.

The Environmental Kuznets Curve (EKC) is a conceptual framework that explores the relationship between environmental degradation and economic development, graphically expressed as a U-inverted curve [5]. It suggests that, initially, as a country's economic growth advances, environmental degradation worsens. Once a specific income threshold is surpassed, EKC implies that environmental quality begins to improve due to heightened affluence, advancements in technology, and changes in consumption patterns. The EKC hypothesis is named after the economist Simon Kuznets [9] who initially proposed a similar idea regarding income inequality and economic development. The EKC for deforestation is called forest transition hypothesis (FTH) Mather [11] and Grainger [12].

FTH suggests a distinctive pattern in which a region initially experiences a decline in forest area as it undergoes economic development and urbanization. After reaching a certain level of economic development, the region attains an increase in forest cover. FTH highlights the dynamic relation between economic development and forest cover. In the long term, when a society progresses, it can experience a transition from net deforestation to net reforestation [13].

Our study is divided into several sections. Section 22.2 provides a review of the state of knowledge in the analyzed field, followed by Sect. 22.3 where we present the methodological framework we will apply to analyze the proposed main objective. Section 22.4 is dedicated to presenting the data we will model, as well as the ARDL framework. Section 22.5 is dedicated to presenting the obtained results, and the final section, Sect. 22.6, presents the conclusions of our research, the limitations of the study, and future research directions.

22.2 Literature Review

Multiple studies discuss the Environmental Kuznets Curve (EKC) for Romania. Simionescu [31] explores GHG emissions in CEE countries, including Romania, revealing an inverted N-shaped relationship between GDP and GHG emissions and a U-shaped Renewable Energy Kuznets Curve (RKC). Shabhzaz et al. [10] confirm the EKC in Romania from 1980 to 2010 using the ARDL bounds test. Jubenot [34] finds initial stages of the Kuznets curve in Romanian agriculture. Simuț and Pețicaș [35] show the EKC did not hold in Romania pre-1990. The paper by [32] examines the EKC hypothesis for the 42 Romanian counties during 2000–2014 using spatial panel econometric methods. The findings reveal an inverted EKC, suggesting that an increase in residential built-up areas is associated with higher wealth levels. Furthermore, the analysis indicates that the expansion of built-up land in Romania primarily mirrors urbanization processes like sprawl or suburbanization, which could lead to adverse environmental and social impacts.

Simionescu et al. [33] discover an N-shaped GHG-GDP relationship in CEE countries. Nuroglu and Kunst [36] examine EKC in similar countries, noting a trade-off between environmental health and social equity. Radulescu et al. [19] support the Kuznets curve for deforestation in Romania. The studies by [26, 27] focus on the integration of advanced digital twin technology within the framework of business model-oriented energy communities. They delved into the holistic view of how energy communities can be structured and managed effectively through innovative business models, while also exploring the role of digital twins in enabling better coordination and efficiency. They highlighted the potential of merging cutting-edge digital tools with strategic business planning to enhance the sustainability and effectiveness of energy communities avoiding deforestation.

22.3 Methodology

EKC has been modeled with both time series and panel data, under various forms. There are several expressions of EKC such as linear, quadratic, cubic and quartic, with or without logarithms. According to Brown and McDonough [14], EKC pertains to a long-term conceptual framework. For a comprehensive discussion of the EKC equation one should refer to [8].

The cubic ECK equation has the form [16]:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \beta_4 z_t + \varepsilon_t \quad (22.1)$$

where t denotes the time, y denotes the environmental indicator, x denotes the income, z represents other indicators having an impact on environmental degradation and ε is the error. There are seven cases and the signs of the constants of Eq. (22.1).

- (i) $\beta_1 = \beta_2 = \beta_3 = 0$, it follows that y does not depend on x ;
- (ii) $\beta_1 > 0$ and $\beta_2 = \beta_3 = 0$, it follows that y and x are in a linear or monotonically increasing relation;
- (iii) $\beta_1 < 0$ and $\beta_2 = \beta_3 = 0$, it follows that y and x are in a monotonically decreasing relation;
- (iv) $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 = 0$, it follows that y and x are in an inverted U-shaped relation (this is EKC);
- (v) $\beta_1 < 0$, $\beta_2 > 0$ and $\beta_3 = 0$, it follows that y and x are in a U-shaped relation;
- (vi) $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$, it follows that y and x are in an N-shaped relation;
- (vii) $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 < 0$, it follows that y and x are in an inverted N-shaped relation.

Efforts have been undertaken to calculate the EKC across various model specifications. Grossman and Krueger [5] conducted their estimation using a cubic model in its level form. Harbaugh et al. [30] carried out their EKC estimation using both logarithmic and level forms of a cubic specification. In this paper we will focus on the quadratic ECK equation and we will attempt to determine the shape of the EKC, in case it exists. When examining the correlation between economic growth and forest resources, several indicators are used: deforestation, forest area growth, forest stock volume [18]. In our case the environmental indicator is the forest area growth. GDP per capita is commonly used to gauge economic growth [17].

22.4 Data and ARDL Framework

This study examines the long-run and the short-run causality between FAG as a dependent variable and its determinants: GDP, EUSE and URB. The data was collected from World Bank for Romania for the period 1990–2021 and is summarized in Table 22.1.

In this study, we utilize the ARDL methodology pioneered by [1], along with [2]. This approach is widely acknowledged as the most appropriate econometric

Table 22.1 Variables and sources

Variable	Acronym	Measurement unit	Source
Forest area	FA	% of land area	World bank
GDP per capita	GDP	Constant 2015\$USD	World bank
Energy use	EUSE	Kg of oil equivalent per capita	World bank
Urbanization	URB	%	World bank

technique for situations in which variables have a combination of 0 and 1 orders of integration.

The formulation of the ECK for deforestation has the following form:

$$FAG_t = a_0 + a_1GDP_t + a_2GDP_t^2 + a_3EUSE_t + a_4URB_t + \varepsilon_t \tag{22.2}$$

where ε_t represents the error term.

The ARDL regression is:

$$\begin{aligned} \Delta FAG_t = & a_0 + \sum_{k=1}^n a_1 \Delta FAG_{t-k} \\ & + \sum_{k=1}^p a_2 \Delta GDP_{t-k} + \sum_{k=1}^q a_3 \Delta GDP_{t-k}^2 + \sum_{k=1}^r a_4 \Delta EUSE_{t-k} + \\ & + \sum_{k=1}^s a_5 \Delta URB_{t-k} + \lambda_1 GDP_{t-1} + \lambda_2 GDP_{t-1}^2 + \lambda_3 EUSE_{t-1} + \lambda_4 URB_{t-1} + \varepsilon \end{aligned} \tag{22.3}$$

In Eq. (22.3) Δ is the first difference, and n, p, q, r and s are lag lengths.

The error correction model (ECM) in case there is cointegration is represented in Eq. (22.4):

$$\begin{aligned} \Delta FAG_t = & a_0 + \sum_{k=1}^n a_1 \Delta FAG_{t-k} \\ & + \sum_{k=1}^p a_2 \Delta GDP_{t-k} + \sum_{k=1}^q a_3 \Delta GDP_{t-k}^2 + \sum_{k=1}^r a_4 \Delta EUSE_{t-k} + \\ & + \sum_{k=1}^s a_5 \Delta URB_{t-k} + \Gamma ECM_{t-1} + \varepsilon_t \end{aligned} \tag{22.4}$$

Γ describes the short-run dynamics of ECM.

Following the assessment of both long-run and short-run estimations, the model’s robustness is examined through the application of the Breusch-Godfrey serial correlation LM model, ARCH heteroscedasticity test, and the normality test. According to Brown’s (1975) methodology, the stability of the model is assessed through the examination of CUSUM and CUSUMSQ plots.

22.5 Empirical Results and Discussion

The Augmented Dickey-Fuller (ADF) unit root test [4] is used to check the stationarity of the data. From Table 22.2 one obtains that all variables are integrated of order 1.

Table 22.2 ADF unit root test

Variables	Level	First difference	Order of integration
	T-statistics	T-statistics	
FAG	- 2.09 (0.248)	- 4.04*** (0.004)	I (1)
GDP	1.64 (0.999)	- 4.24*** (0.002)	I (1)
GDP2	2.27 (0.999)	- 4.29*** (0.002)	I (1)
EUSE	- 2.00 (0.283)	- 4.66*** (0.001)	I (1)
URB	- 0.24 (0.922)	- 3.37** (0.020)	I (1)

*, **, *** Indicate the significance of variables at 10%, 5%, and 1% levels, respectively

Table 22.3 details the lag structure for the ARDL model, indicating that the optimal number of lags to use is 3. LR is a sequential modified LR test statistic performed at a 5% significance level, FPE represents Final Prediction Error, AIC stands for Akaike Information Criterion, SC denotes Schwarz Information Criterion, and HQ signifies Hannan-Quinn Information Criterion.

Table 22.4 presents the results of the ARDL cointegration bounds test, where the computed F-statistic is 7.28. This value surpasses the critical upper bounds for I (1) at both the 5% and 10% significance levels. The chosen model for this analysis is ARDL (2, 2, 3, 3).

Table 22.5 presents the estimated long-run coefficients, showing a significant positive impact of GDP on Forest Area Growth (FAG). Specifically, a 1% increase in GDP results in a 0.002% increase in FAG. This relationship can be attributed to various interconnected factors. As GDP grows, there is heightened awareness and concern for environmental conservation, leading to the implementation of policies

Table 22.3 VAR Lag order selection criteria

Lag	LogL	LR	FPE	AIC	SC	HQ
0	673.40	N/A	4.19×10^{20}	61.67	61.92	61.73
1	- 548.10	182.24	4.93×10^{20}	52.55	54.04	52.90
2	- 508.90	39.20*	2.07×10^{16}	51.26*	53.99*	51.90*

* Indicates the lag order selected by the criterion

Table 22.4 ARDL cointegration bounds test

Test statistic	Value	K (number of regressors)
F-statistic	7.28	3
Critical value bound (n = 35)		
Significance	I (0)	I (1)
10%	2.46	3.46
5%	2.94	4.08
1%	4.09	5.53

Table 22.5 ARDL long-run estimated coefficients

Variables	Coefficient	t-statistics	Prob.
GDP	0.002	2.53	0.027**
GDP2	$- 2.46 \times 10^{-7}$	- 2.54	0.027***
URB	1.69	3.71	0.003***
EUSE	- 0.004	- 4.26	0.001***
C	- 88.49	- 3.55	0.004***

and practices that support FAG. Economic growth allows for more investments in environmental projects, technological advancements for sustainable logging, and greater public awareness about forest conservation.

The study also identifies a negative impact of GDP2 on FAG, indicating an inverted U-shaped relationship between GDP and FAG for Romania. Initially, GDP growth increases FAG, but after reaching a certain threshold, this growth stagnates. This finding aligns with [19], who observed a similar pattern using the VECM model for Romania over 22 years.

Additionally, Table 22.5 reveals that a 1% increase in urbanization (URB) leads to a 1.69% increase in FAG. Urbanization often results in rural-to-urban migration, leaving agricultural lands to revert to natural vegetation and forests. As remarked by [20, 21], modern urban planning practices emphasize the inclusion of green spaces and forests within urban areas. Modern urban planning in Romania includes creating urban forests and parks, contributing to FAG. Urban planning also promotes environmental awareness and biodiversity through the development of green spaces and wildlife corridors. Similar actions of urban planning, building urban ecological networks and eco-logical corridors were mentioned in the study of [22] in case of Zhuhai city, China and in the study of Wang et al. [7] for 30 Chinese provinces. Conversely, a 1% increase in energy use (EUSE) results in a 0.004% decrease in FAG. The negative impact of EUSE on FAG in Romania stems from historical reliance on wood for energy and industrial practices that degrade forests. Energy production from non-renewable sources contributes to pollution, soil degradation, and GHG emissions, negatively affecting forest ecosystems. As energy demand grows, forest lands may be repurposed, leading to biodiversity loss. Similar trends were noted by [23] in rural Iranian communities, where increased fossil fuel consumption drove deforestation. Next, we will determine the turning point of the EKC, according to [28].

The ECM equation is

$$FAG = 0.0029GDP - 2.46e - 07GDP2 + 1.694URB - 0.004EUSE - 88.49 \tag{22.5}$$

We compute the first derivative of Eq. (22.5) with respect to GDP and set it to 0. Then we find the optimal point denoted GDP^* :

$$GDP^* = 5894.31 \tag{22.6}$$

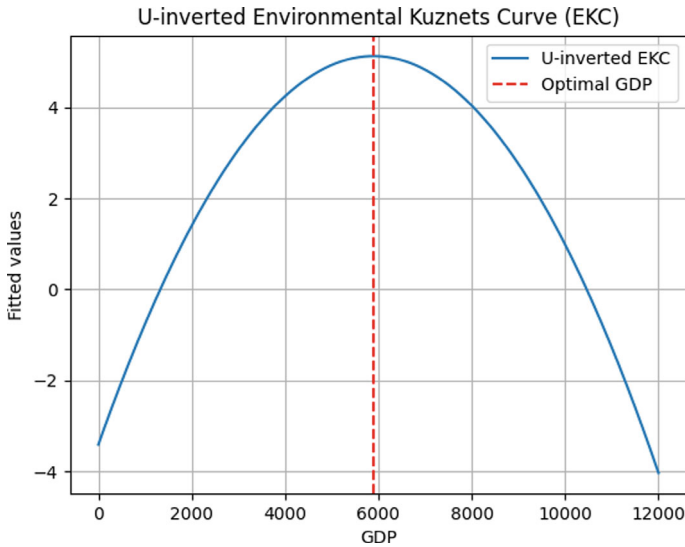


Fig. 22.1 EKG for Romania

At the initial values of GDP, the level of FAG rises, but as GDP progresses beyond the optimal point of 5894.31, the level of FAG declines. Using the `<<matplotlib.pyplot>>` and `<<numpy>>` libraries, the U-inverted EKC curve was implemented in PyCharm using the Python programming language. The U-inverted EKC had the graph according to Fig. 22.1.

Table 22.6 shows the short-run coefficients and the error correction term (ECT), which is -0.53 and statistically significant. This indicates that the system returns to long-run equilibrium at a 53% adjustment rate after a shock. The small short-run coefficients of the ARDL model suggest that changes in GDP, urbanization (URB), and energy use (EUSE) have minimal immediate effects on forest area growth (FAG). This implies that FAG is relatively stable or resilient to short-term economic and developmental changes. For policymakers, this indicates that efforts to influence forest area growth through short-term measures may be ineffective, emphasizing the need for long-term planning and sustainability. These findings suggest a possible balance or decoupling between economic development and environmental impact in the short term, as noted by [19], possibly due to effective environmental policies or natural forest resilience.

Table 22.7 reports four diagnostic tests along with their null hypotheses. The acceptance of all these null hypotheses indicates that the ARDL-ECM model is robust.

The stability of the parameters is assessed by using the CUSUM and CUSUMSQ tests as proposed by [3]. The two straight red dashed lines are drawn on the CUSUM and CUSUMSQ plots in Fig. 22.2a and b representing critical bounds at 5% confidence level. Since the CUSUM and CUSUMSQ lines stay within these bounds, the

Table 22.6 Short-run ARDL approach

Variables	Coefficient	t-statistics	Prob.
D (GDP)	- 0.001	- 3.79	0.003***
D (GDP (- 1))	- 0.001	- 3.38	0.006***
D (EUSE)	- 0.0004	- 1.64	0.127
D (EUSE (- 1))	0.001	6.50	0.000***
D (GDP2)	9.31×10^{-8}	3.55	0.004***
D (GDP2 (- 1))	1.01×10^{-7}	3.04	0.011**
CointEq (- 1)	- 0.53	- 7.97	0.000***
R-squared	0.82		
Adjusted R-squared	0.76		

*, **, *** Indicate the significance of variables at 10%, 5%, and 1% levels, respectively

Table 22.7 Diagnostic tests

Diagnostic test	H ₀	Decision statistics [p-value]
χ^2 SERIAL	There is no serial correlation in the residuals	Accept H ₀ 1.99 [0.188]
χ^2 ARCH	There is no autoregressive conditional heteroscedasticity	Accept H ₀ 0.39 [0.539]
χ^2 JARQUE BERA	Normal distribution	Accept H ₀ 1.00 [0.604]
χ^2 RAMSEY	Absence of model misspecification	Accept H ₀ 3.56 [0.118]

null hypothesis of coefficient stability is not rejected, indicating that the coefficients of the ARDL model are stable over time.

22.6 Conclusions and Recommendations

Our study contributes to understanding the relationship between economic development and forest area growth in Romania. We find evidence supporting the existence of an inverted U-shaped Environmental Kuznets Curve (EKC) for deforestation. This suggests that as GDP initially increases, forest area grows, indicating a positive correlation between economic development and forest conservation efforts. However, beyond a certain threshold, further economic growth leads to a decline in forest area. This finding underscores the importance of balancing economic development with environmental conservation efforts to ensure sustainable growth.

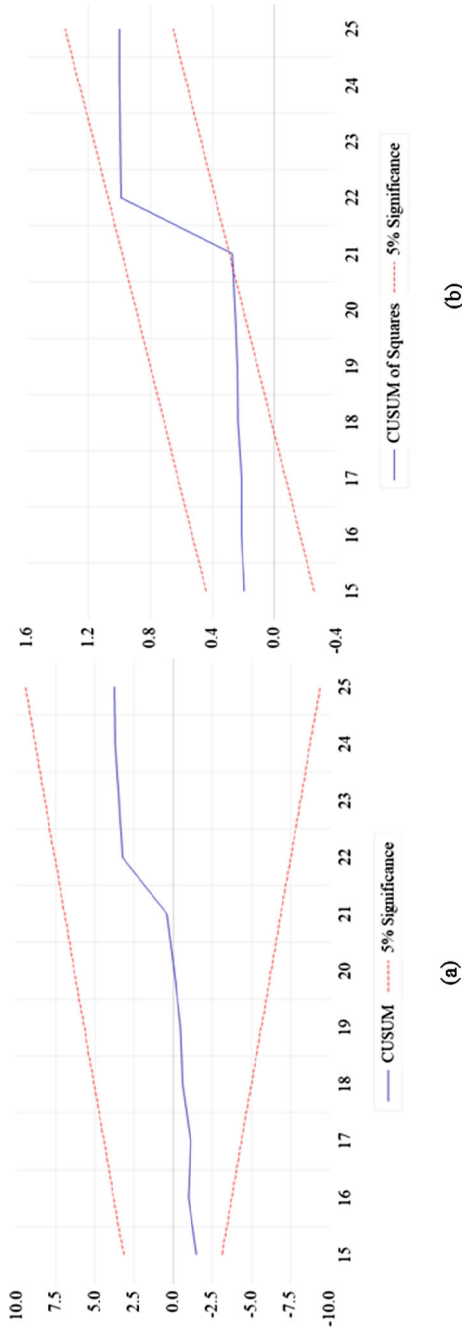


Fig. 22.2 Plot of **a** CUSUM at 5% level of significance; **b** CUSUM of squares at 5% level of significance

Also, our analysis reveals several significant long-run effects of economic and urban factors on forest area in Romania. Specifically, we find a positive and significant impact of GDP on forest area, indicating that higher GDP levels are associated with increased forest coverage. This suggests that economic growth can provide resources for environmental conservation initiatives. Additionally, urbanization has a positive effect on forest area, highlighting the role of urban planning in preserving and expanding forested areas within urban environments. However, energy consumption has a negative impact on forest area growth, indicating the importance of transitioning towards sustainable energy sources to mitigate deforestation. Our short-run analysis suggests that changes in GDP, urbanization, and energy consumption have minimal immediate effects on forest area in Romania. This implies that forest ecosystems exhibit a degree of resilience to short-term economic fluctuations and urban development patterns. Therefore, policies aimed at promoting forest conservation and expansion may require long-term planning and sustainability measures. Additionally, our findings suggest a potential decoupling between economic development and environmental impact in the short term, indicating the need for targeted policies to ensure that economic growth is environmentally sustainable.

Our study may have some limitations. Firstly, the latest available data were from 2021; in future research directions on our topic, these data can be extended when they become available. Another limitation could be the generalization of results. Our case study focuses on the local economy of Romania, and therefore, the results obtained could be specific to the context and conditions of this country.

Based on our results, policymakers in Romania should prioritize sustainable economic development strategies that promote forest conservation and expansion. This includes investing in green technologies, promoting urban planning practices that incorporate forested areas within urban environments, and implementing policies to reduce reliance on fossil fuels and promote renewable energy sources. Furthermore, efforts to raise public awareness about the importance of forest conservation and biodiversity should be enhanced to garner support for environmental initiatives.

Acknowledgements This work was funded by the EU's NextGenerationEU instrument through the National Recovery and Resilience Plan of Romania—Pillar III-C9-I8, managed by the Ministry of Research, Innovation and Digitization, within the project entitled “*Place-based Economic Policy in EU's Periphery—fundamental research in collaborative development and local resilience. Projections for Romania and Moldova (PEPER)*”, contract no. 760045/23.05.2023, code CF 275/30.11.2022.

References

1. Pesaran, M.H., Shin, Y.: An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch, Centennial Symposium* **31**, 371–413 (1998). <https://doi.org/10.1017/CCOL521633230.011>
2. Pesaran, M.H., Shin, Y., Smith R.J.: Bounds testing approaches to the analysis of level relationships. *J. Appl. Econ.* **16**(1), 289–326 (2001). <https://doi.org/10.1002/jae.616>

3. Brown, R.L., Durbin, J., Evans, J.M.: Techniques for testing the constancy of regression relationships over time. *J. R. Stat. Society. Ser. B (Methodol.)* **37**(2), 149–192 (1975). <https://doi.org/10.1111/j.2517-6161.1975.tb01532.x>
4. Dickey, D., Fuller W.: Distribution of the estimator for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **74**, 427–431 (1979). <https://doi.org/10.2307/2286348>
5. Grossman, G.M., Krueger, A.B.: Economic growth and the environment. *Q. J. Econ.* **110**(2), 353–377 (1995). <https://doi.org/10.2307/2118443>
6. Ojaghlou, M., Ugurlu, E., Kadlubek, M., Thalassinou, E.: Economic activities and management issues for the environment: an environmental Kuznets curve (EKC) and STIRPAT analysis in Turkey. *Resources* **12**, 57 (2023). <https://doi.org/10.3390/resources12050057>
7. Wang, J., Li, Z., Wu, T., Wu, S., Yin, T.: The decoupling analysis of CO₂ emissions from power generation in Chinese provincial power sector. *Energy*. **255**, 124488 (2022). <https://doi.org/10.1016/j.energy.2022.124488>
8. Hasanov, F.J., Hunt, L.C., Mikayilov, J.I.: Estimating different order polynomial logarithmic environmental Kuznets curves. *Environ. Sci. Pollut. Res.* **28**, 41965–41987 (2021) <https://doi.org/10.1007/s11356-021-13463-y>
9. Kuznets, S.: Economic growth and income inequality. *Am. Econ. Rev.* **45**, 1–28 (1995)
10. Shabbaz, M., Mutascu, M., Azim, P.: Environmental Kuznets curve in Romania and the role of energy consumption. *Renew. Sustain. Energy Rev.* **18**, 165–173 (2013). <https://doi.org/10.1016/j.rser.2012.10.012>
11. Mather, A.S.: The forest transition. *Area* **24**(4), 367–379 (1992). <https://www.jstor.org/stable/20003181>
12. Grainger, A.: The forest transition: an alternative approach. *Area* **27**(3), 242–251 (1995). <https://www.jstor.org/stable/20003580>
13. Culas, R.J.: Deforestation and the environmental Kuznets curve: an institutional perspective. *Ecol. Econ.* **61**(2–3), 429–437 (2007). <https://doi.org/10.1016/j.ecolecon.2006.03.014>
14. Brown, S.P.A., McDonough, I.K.: Using the environmental Kuznets curve to evaluate energy policy: some practical considerations. *Energy Pol.* **98**, 453–458 (2016)
15. Leal P.H., Marques A.C.: The evolution of the environmental Kuznets curve hypothesis assessment: a literature review under a critical analysis perspective. *Heliyon* **8**(11):e11521 (2022). <https://doi.org/10.1016/j.heliyon.2022.e11521>. PMID: 36406679; PMCID: PMC9668524
16. Karadaş, H.A., Işık, H.B.: Environmental Kuznets Curve (EKC): is it valid for Turkey? *Economic and Management Issues in Retrospect and Prospect*, pp. 285–286 (2018)
17. Caravaggio, N.: Economic growth and the forest development path: A theoretical reassessment of the environmental Kuznets curve for deforestation. *Policy Econ.* **118**, 102259–102311 (2020). <https://doi.org/10.1016/j.forpol.2020.102259>
18. Wu, X., Liu, G., Bao, Q.: Impact of economic growth on the changes in forest resources in Inner Mongolia of China. *Front. Environ. Sci.* **11**, 1241703 (2023). <https://doi.org/10.3389/fenvs.2023.1241703>
19. Rădulescu, C.V., Bran, F., Ciuvăț, A.L., Bodislav, D.A., Buzoianu, O.C., Ștefănescu, M., Burlacu, S.: Decoupling the economic development from resource consumption: implications and challenges in assessing the evolution of forest area in Romania. *Land* **11**, 1097 (2022). <https://doi.org/10.3390/land11071097>
20. Nita, M.R., Nastase, I.I., Badiu, D.L., Onose, D.A., Gavrilidis, A.A.: Evaluating urban forests connectivity in relation to urban functions in Romanian cities. *Carpathian J. Earth Environ. Sci.* **13**(1), 291–299 (2018). <https://doi.org/10.26471/cjees/2018/0>
21. Grigorescu, I., Geacu, S.: The dynamics and conservation of forest ecosystems in Bucharest Metropolitan Area. *Urban For. & Urban Green.* **27**, 90–99 (2017). <https://doi.org/10.1016/j.ufug.2017.04.012>
22. Cui, N., Feng, C.-C., Wang, D., Li, J., Guo, L.: The effects of rapid urbanization on forest landscape connectivity in Zhuhai city China. *Sustainability* **10**, 3381 (2018). <https://doi.org/10.3390/su10103381>
23. Abbaspour, M., Ghazi, S.: An alternative approach for the prevention of deforestation using renewable energies as substitute. *Renew. Energy* **49**, 77–79 (2013). <https://doi.org/10.1016/j.renene.2012.01.062>

24. MMAP: The National Forest Strategy 2030; Ministry of Environment, Waters and Forests. Bucharest, Romania (2022). <http://www.mmediu.ro/categorie/strategia-nationala-a-padurilor-2022-2031/386>. Accessed on 3 Nov 2022
25. Giurca, A., Nichiforel, L., Stăncioiu, P.T., Drăgoi, M., Dima, D.-P.: Unlocking Romania's forest-based bioeconomy potential: knowledge-action-gaps and the way forward. *Land* **2022**, 11 (2001). <https://doi.org/10.3390/land11112001>
26. Băra, A., Oprea, S.V.: A holistic view on business model-oriented energy communities. *Kybernetes* **54**(2), 871–894 (2023). <https://doi.org/10.1108/K-07-2023-1235>
27. Băra, A., Oprea, S.-V.: Enabling coordination in energy communities: a digital twin model. *Energy Policy* **184**, 113910. <https://authors.elsevier.com/a/1iAus14YGgpcuM> (2024)
28. Adeleye, B.N., Akam, D., Inuwa, N. et al.: Does globalization and energy usage influence carbon emissions in South Asia? An empirical revisit of the debate. *Environ. Sci. Pollut. Res.* **30**, 36190–36207 (2023). <https://doi.org/10.1007/s11356-022-24457-9>
29. Georgescu, I., Kinnunen, J., Nica, I.: Assessing forest conservation for Finland: an ARDL-based evaluation. *Sustainability* **16**(2), 612 (2024). <https://doi.org/10.3390/su16020612>
30. Harbaugh, W., Levinson, A., Wilson, DM.: Re-examining the empirical evidence for an environmental Kuznets curve. *Rev. Econ. Stat.* **84**(3), 541–551 (2002). <https://doi.org/10.1162/003465302320259538>
31. Simionescu, M.: Revised environmental Kuznets Curve in CEE countries. Evidence from panel threshold models for economic sectors. *Environ. Sci. Pollut. Res.* **28**, 60881–60899 (2021). <https://doi.org/10.1007/s11356-021-14905-3>
32. Pontarollo, N., Serpiri, K.: Testing the Environmental Kuznets Curve hypothesis on land use: the case of Romania. *Land Use Policy* **97**, 104695 (2020). <https://doi.org/10.1016/j.landusepol.2020.104695>
33. Simionescu, M., Păuna, CB., Niculescu, M.-DV.: The relationship between economic growth and pollution in some new European union member states: a dynamic panel ARDL approach. *Energies* **14**(9), 2363 (2021). <https://doi.org/10.3390/en1409236>
34. Jubenot, M.N.: Applying the Kuznets curve in case of Romania. *The Annals of the University of Oradea, Oradea, Romania*, pp. 106–115 (2013)
35. Simuț, R.S., Perțicaș, D.C.: Testing Kuznets curve for the socialist and capitalist period in Romania. *Res. J. Econ. Bus. ICT* **8** (2013)
36. Nuroglu, E., Kunst, R.M.: Kuznets and environmental Kuznets curves for developing countries. In: Yülek, M. (eds.) *Industrial Policy and Sustainable Growth. Sustainable Development*. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-5741-0_12

Chapter 23

Modeling Financial Contagion: Exploring Interconnectedness and Systemic Risk Through Agent-Based Simulation



Ștefan Ionescu , George Tătaru, Gabriel Dumitrecu, Nicole Bianca Stanică,
and Liviu-Adrian Cofas 

Abstract Our study delves into the dynamics of financial contagion within the network of 32 commercial banks in Romania through the application of agent-based modeling using NetLogo. What distinguishes our study is the incorporation of government intervention into the model when default occurs in the banking network to assess its effects. Our simulations demonstrate that the level of interconnectivity between banks plays a crucial role in determining the extent of systemic risk and the spread of financial contagion within the banking network. Higher levels of interconnectivity increase the risk of contagion, highlighting the importance of understanding and monitoring interbank linkages. Additionally, our results underscore the significance of government intervention in mitigating the effects of financial contagion. We observed that higher levels of government intervention lead to a reduction in the number of banks entering default, indicating the effectiveness of policy measures in stabilizing the banking system during crises. Banks that do not default, especially those without government intervention, demonstrate resilience to financial stress. Their ability to maintain stability without external assistance emphasizes the importance of sound risk management practices and capital buffers in safeguarding against systemic shocks. The insights gained from our simulations have significant implications for policymakers and regulators. Understanding the dynamics of financial contagion and the factors influencing systemic risk can inform the development of more robust regulatory frameworks and risk management strategies aimed at enhancing the stability and resilience of the financial system.

Ș. Ionescu (✉) · G. Tătaru · N. B. Stanică · L.-A. Cofas
Department of Economic Informatics and Cybernetics, Bucharest University of Economic
Studies, Bucharest, Romania
e-mail: stefion09@gmail.com

G. Dumitrecu
Department of Accounting and Audit, Bucharest University of Economic Studies, Bucharest,
Romania

23.1 Introduction

In a world marked by the complexity and interconnectedness of financial systems, understanding the phenomenon of contagion becomes a crucial priority for establishing and maintaining global economic stability. In this era of technology and innovation, agent-based modeling represents a powerful tool in elucidating the dynamics and effects of financial contagion.

Contagion in financial markets refers to the transmission of shocks or disturbances from one financial institution, market, or asset class to others, leading to a widespread disruption or amplification of the initial shock [1]. It can occur due to various channels, such as interconnections between financial institutions, herding behavior among investors, or common exposure to risk factors [2]. Understanding and managing financial contagion is essential for maintaining the stability of financial systems and preventing systemic crises.

Agent-based modeling in analyzing the effects of financial contagion represents a promising approach in understanding the complex dynamics of interactions among financial entities within a system. This method involves creating virtual agents that represent various financial institutions or market entities and are capable of interacting with each other based on defined rules and behaviors. Through simulations conducted with these agents, researchers can explore how events or shocks in one part of the financial market can propagate and affect the entire system, thereby highlighting the channels and mechanisms through which financial contagion can spread. This approach provides a detailed perspective on complex interactions and can assist in identifying vulnerabilities in the financial system and developing more effective risk management strategies.

Our study explores the financial contagion effects encountered in a banking network using agent-based modeling. The novelty of our study lies in introducing government intervention as an influencing parameter in a network where systemic default risk spreads to observe how this intervention influences the behavior of banks in the network. The structure of this work is as follows: the next section presents some significant studies from the literature, Sect. 23.3 outlines the methodological flow and mathematical premises underlying the development of the agent-based model, while Sect. 23.4 presents the results obtained from 1000 simulations for each formulated scenario. Finally, Sect. 23.5 presents the research conclusions, potential limitations of our study, and future research directions.

23.2 Literature Review

Financial contagion, a phenomenon characterized by the rapid spread of financial distress from one market to others, has been a focal point of research in economics and finance. Understanding the mechanisms and implications of financial contagion is essential for safeguarding the stability of global financial systems.

In recent years, there has been a growing body of literature dedicated to unraveling the complexities of financial contagion, driven by the increasing interconnectedness of global financial markets and the recurrence of systemic crises. Researchers have employed various methodologies, including empirical analyses, theoretical models, and simulation techniques, to explore the dynamics of contagion and its impact on different sectors and regions.

Allen and Gale [3] describe financial contagion as a phenomenon that can affect the entire economy, with this effect being modeled as an equilibrium-reaching phenomenon in which banks determine a level of interregional claims to protect themselves against liquidity preference shocks, and the degree of vulnerability to contagion depending on the completeness of the structure of these claims.

Bae et al. [4] propose a new approach to evaluate the contagion effect in financial markets. They analyze the correlation between extreme return shocks among countries within a region and across regions, considering these shocks as an extension of the contagion phenomenon, along with its economic significance and determining factors. They employ a multinomial logistic regression model to address these contagion effects, concluding that the effect is stronger for extreme negative returns than for extreme positive returns, with mixed evidence.

Camera et al. [2] explore the policy interventions during extreme shocks using a random matching framework where the financial contagion process is explicit. The authors concluded that ordinary policy alone cannot sufficiently prevent or mitigate contagion, while adding a lockdown or individual closures can do so and improve social welfare if the initial shock is severe but not widespread. In another study conducted by Lai and McNelis [5], three methods for assessing the contagion effect among ten Globally Significant International Banks are compared, based on machine learning and network analysis.

Nica et al. [6] conduct a bibliometric analysis providing a holistic perspective on the concepts of financial contagion and the shadow banking sector. The authors concluded that their paper offers a detailed insight into the progress, gaps, and future directions in understanding the complex interplay between the shadow banking system and financial contagion.

Events occurring in the environment can impact economic systems, given the cybernetic and complex adaptive nature of all systems. The COVID-19 pandemic represented a shock both socially and economically. Akhtaruzzaman et al. [7] analyzed the pandemic context from the perspective of financial contagion. Thus, the authors examined how the contagion effect occurs during the COVID-19 pandemic among financial and non-financial firms in China and the G7 countries (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States of America). The authors observed that listed firms in these countries experienced a significant increase in correlations of their stock returns.

A recent study authored by Pallant et al. [8] extends the Schumpeter-correlated Keynesian model to analyze the evolution of the interbank network in the money market and implement a micro-prudential policy based on Network Valuation (NEVA). The authors explored methods including agent-based modeling and showed

that the economic system can benefit from financial stability without a trade-off between this and macroeconomic performance.

One of the significant researches works underlying our study is the one conducted by Gai and Kapadia [9]. In their study, the authors developed an analytical model of contagion in financial networks with arbitrary structure, exploring how the probability and potential impact of contagion are influenced by aggregate and idiosyncratic shocks, changes in network structure, and asset market liquidity.

Thus, we observe that studies and research in the field of financial contagion highlight the importance of understanding this complex phenomenon for safeguarding the stability of global financial systems. These research efforts explore the mechanisms and implications of financial contagion using various analytical methods and modeling techniques, including empirical analyses, theoretical models, and simulations.

23.3 Methodology

Since the advent of computers, progress in the natural sciences and engineering has been enormous, computer simulations have allowed understanding of physical and astronomical interactions, description of chemical properties, and design of safer and more energy-efficient vehicles [10]. The use of computers is ubiquitous in daily activities, from offices and administration to financial transactions and our communication.

To explore the interconnectedness and systemic risk underlying the modeling of financial contagion effects, we will employ agent-based modeling (ABM) in our study. This approach involves creating virtual entities (agents) representing various financial institutions, banks, or market participants [11, 12]. These agents are programmed to react to changes in the environment and the actions of other agents in a realistic manner, based on defined rules and behaviors. By simulating interactions among these agents, ABM can help identify the channels and mechanisms through which financial contagion spreads and assess its impact on the entire financial system. Additionally, ABM can be used to evaluate the effectiveness of various policies and systemic risk management strategies, thus providing valuable insights for the formulation of policies and regulations in the financial domain [12–14].

Our study will build upon the approach developed by Gai and Kapadia [15], where they considered a financial network in which “ n ” financial intermediaries (such as banks) are randomly linked together by their reciprocal claims. Essentially, each node of the formed network will be represented by a bank, and the interbank exposures of bank “ j ” define the connections to other banks. These connections are directed and weighted, reflecting the fact that interbank exposures include assets as well as liabilities, and that the size of these exposures is important for contagion analysis. As defined by Gai and Kapadia [15], we will consider the same hypothesis, namely that each bank has two main types of assets: interbank assets Γ_{IA_i} and illiquid external assets Γ_{IEA_i} , such as mortgages. We will assume that each bank evenly distributes

its interbank assets among all its incoming links and that this is independent of the number of links it has. Even when banks efficiently share their risks, it is possible for a crisis to spread widely throughout the entire financial network. We will consider that interbank liabilities Ψ_{IL_i} are endogenously determined since each interbank asset represents a liability for another bank. Additionally, we will assume the presence of another component of a bank's liabilities, such as customer deposits Φ_i , which will also be given exogenously. Therefore, the condition for bank i to be solvent is:

$$(1 - \xi) * \Gamma_{IA_i} + p * \Gamma_{IEA_i} - \Psi_{IL_i} - \Phi_i > 0 \quad (23.1)$$

In Eq. (23.1), ξ denotes the proportion of banks with liabilities to bank "i" that have defaulted, while "p" represents the residual price of the illiquid asset. Assuming that a connected bank defaults, bank i will forfeit all interbank assets associated with that bank. Hence, the solvency condition can also be stated as:

$$\xi < \frac{\gamma_i - (1 - p) * \Gamma_{IEA_i}}{\Gamma_{IA_i}}, \Gamma_{IA_i} \neq 0 \quad (23.2)$$

In Eq. (23.2), γ_i represents the bank's capital (Eq. (23.3)).

$$\gamma_i = \Gamma_{IA_i} + \Gamma_{IEA_i} - \Psi_{IL_i} - \Phi_i \quad (23.3)$$

To model the dynamics of contagion, we assume that all banks are initially solvent and that a single bank defaults at the beginning of the model. This disruption may result from a rare idiosyncratic shock or an aggregate shock that particularly affects a specific financial institution.

Let's denote q_i as the number of incoming links for bank i . Considering that each connected bank forfeits a fraction $\frac{1}{q_i}$ of its interbank assets in the event of a single counterparty's default, it's evident from Eq. (23.2) that the spread of default is contingent upon the presence of neighboring banks for each bank i with an in-degree q_i :

$$\frac{\gamma_i - (1 - p) * \Gamma_{IEA_i}}{\Gamma_{IA_i}} < \frac{1}{q_i} \quad (23.4)$$

We will further define banks exposed in this sense (according to Eq. (23.4)) to the default of a single neighbor as vulnerable, while other banks as safe. Given that the capital buffer is considered a random variable, a bank with an in-degree "q" is vulnerable with a probability:

$$\varphi_i = P \left[\frac{\gamma_i - (1 - p) * \Gamma_{IEA_i}}{\Gamma_{IA_i}} < \frac{1}{q} \right], \forall q \geq 1 \quad (23.5)$$

Moreover, the likelihood of a bank having an in-degree q , an out-degree of k , and being vulnerable is denoted as p_{qk} , where p_{qk} represents the joint degree distribution of in and out-degree.

Based on the mathematical reasoning described above, in our study, we will employ an agent-based modeling approach to simulate the effects of financial contagion in the banking system of Romania. We will consider that the nodes of the network consist of the 32 financial banks [16] existing at the level of Romania. Agent-based modeling [17, 18] offers an efficient approach to understand the behavior of economic agents and the factors that shape it. By employing computational modeling, it delves into the complex interactions among these agents, providing a nuanced and realistic perspective on real-world processes, distinct from equation-based models. The first step in agent-based modeling is defining the properties of agents within the network. Therefore, agents exhibit the following characteristics [12, 19–23]: heterogeneity, emergence, complexity, social interaction, adaptability, learning, adaptiveness, feedback, realistic behavior, learning, interdependence, and exploration of the possibility space.

To model the financial contagion effect using agent-based modeling, we will use the NetLogo computer program because it is the most widely used programming environment for simulating interactions among individual agents in a virtual environment [24].

23.4 Results

Commercial banks, approached and understood as cybernetics systems, must be understood from the perspective of the functions and services they offer, as well as the interconnections between their subsystems. In our study, we will explore how the risk of default can propagate systemically in the network formed by the 32 commercial banks in Romania to investigate the effects of financial contagion that may arise. Essentially, we will simulate the interaction between banks and their environment through agent-based modeling to understand both the behavior of autonomous agents, the banks, and the behavior of the entire banking system when government interventions occur.

Regarding the proposed NetLogo model, we have considered the parameters outlined in Table 23.1 for the simulations. Table 23.1 provides an overview of various aspects considered in simulating financial contagion, including the total number of banks involved in the simulation, changes due to government intervention, and other relevant parameters. It serves as a useful guide for configuring and conducting the simulation, providing the necessary data for a detailed analysis of the effects of financial contagion in a simulated environment.

Thus, we conducted a total of 1000 simulations for the following combined scenarios:

Table 23.1 Parameters and variables for financial contagion simulation

Parameters and variables	Description	Range of value	Parameters and variables	Description	Range of value
Banks	Total number of banks (agents) against which the simulation will run	[1;32]	Default smallest bank	The bank with the smallest size will be picked as the starting point for the financial contagion	[1; number of banks with smallest size]
Up-to-change-for-given connectivity	The chance to achieve the desired connectivity degree, that can be less or equal than the value this variable holds, for each agent individually	[0;100]	Default biggest bank	The bank with the biggest size will be picked as the starting point for the financial contagion	[1; number of banks with biggest size]
Mean of interbank asset distribution	The normally distributed mean of interbank assets, used for modeling financial shocks in the banking network	[0;+∞]	<i>mu</i>	Value of the mean that will be taken into consideration for the normal distribution	Between 0 and +∞. In our study, we had set <i>mu</i> to 0 in order to ensure a Gaussian distribution of the interbank assets
Change-for-governments-intervention	If the agent enters a default-state, the variable describes the maximum chance that the agent has to be saved by the government	(0;100]	<i>Sigma</i>	Value of the standard deviation that will be taken into consideration for the normal distribution	Between 0 and +∞. In our study, we had set <i>sigma</i> to 1.5

(continued)

Table 23.1 (continued)

Parameters and variables	Description	Range of value	Parameters and variables	Description	Range of value
Setup	The simulation's starting point. The purpose of this variable is to assign an initial state to each agent, depending on the aforementioned variables	N/A	Default random bank	In order to achieve a financial contagion, at least one of the banks will be chosen randomly, regardless of their size, to be in a default-state at the beginning of the simulation. For the chosen ones, an exogenous shock will be applied to them	[1; number of banks, regardless of their size]
Go	This variable will execute the simulation by performing a certain number of iterations. Each iteration may result in other agents entering a default-state	N/A	Count links	The total number of connections among all the banks	Between 1 and 32

- (i) We consider randomly selecting the size classification of the bank, either the smallest bank or the biggest bank, from which the default event will originate:
 - a. RB—random bank;
 - b. SB—smallest bank;
 - c. BB—biggest bank.
- (ii) For each of the three cases defined below, we will have the following scenarios:
 - a. 25% connectivity, 25% government intervention;
 - b. 25% connectivity, 50% government intervention;
 - c. 25% connectivity, 75% government intervention;
 - d. 50% connectivity, 25% government intervention;
 - e. 50% connectivity, 50% government intervention;
 - f. 50% connectivity, 75% government intervention;
 - g. 75% connectivity, 25% government intervention;
 - h. 75% connectivity, 50% government intervention;
 - i. 75% connectivity, 75% government intervention.

This approach allows for the exploration of various combinations of scenarios to assess the impact on the banking system in the context of financial contagion. Through multiple simulations, a deeper understanding can be gained of how different levels of connectivity and government intervention can influence the spread of contagion effects within the banking network. These simulations provide a valuable perspective for formulating policies and risk management strategies in the financial domain.

Figure 23.1 depicts the interface of the model, showcasing the parameters and variables outlined earlier. Additionally, this representation illustrates only the scenario where the largest bank defaults (red agent), being interconnected with 18 other banks in the banking system.

In Fig. 23.2, we observe in blue the banks that were not affected by the default event of the largest bank, the green banks represent those injected with liquidity by

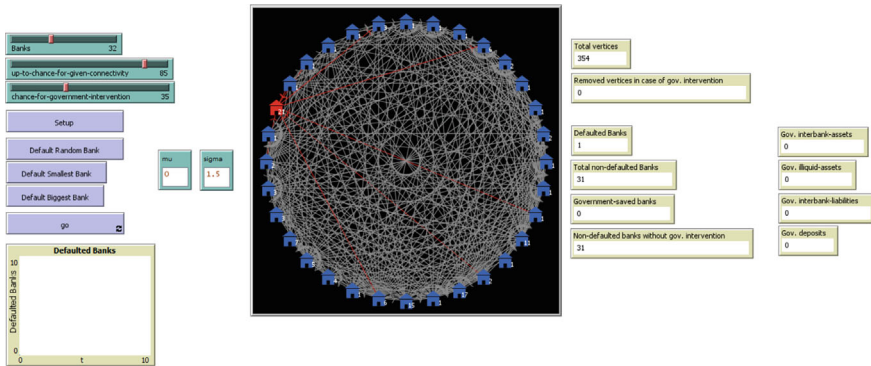


Fig. 23.1 The visualization of the agent-based model is based on the scenario of default for the biggest bank

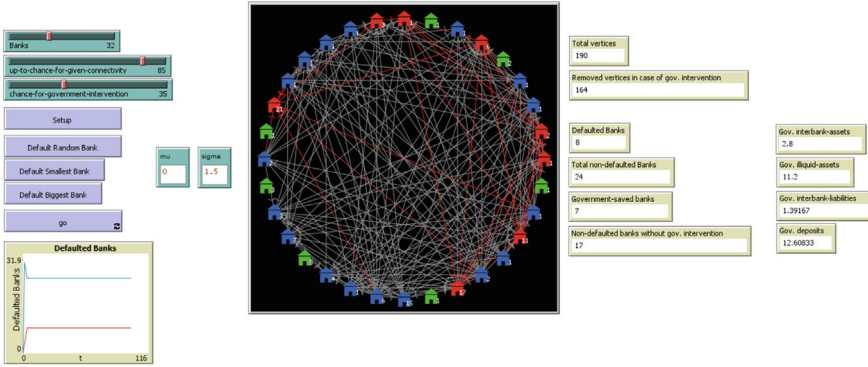


Fig. 23.2 The final state when the government intervenes with liquidity injection into the banking system

the government, thus being saved from the contagion event caused by the largest bank, and the red banks represent those still facing issues.

Following the 1000 simulations conducted for each predefined scenario, according to the output results that can be seen by accessing the following GitHub link: <https://github.com/qfancee/IE2024>, the below observations were made:

- The number of banks defaulting varies depending on the level of connectivity and government intervention. For all types of banks (RB, SB, and BB), we calculated the minimum, median, and maximum values of banks entering default for each scenario. We observed that a higher level of government intervention reduces the number of banks entering default.
- It was observed that for banks not entering default, government intervention represents an effective measure in maintaining the stability of the banking system.
- Regarding banks saved by government intervention, a positive impact on maintaining stability and confidence in the banking system was highlighted.
- Banks marked in blue in the output logs are non-defaulted banks without government intervention. Maintaining these banks in a non-default state demonstrates their ability to survive under financial stress conditions without external assistance.

In conclusion, the simulations provide a detailed insight into how different levels of connectivity and government intervention can influence the stability and resilience of the banking system against financial stress and contagion. This information can be valuable for formulating policies and risk management strategies in the financial domain.

23.5 Conclusions and Recommendations

Our study explores how the effect of financial contagion propagates in the banking system of Romania, consisting of 32 banks, and the effects of government intervention in case of default risk. Our simulations demonstrate that the level of interconnectedness among banks plays a crucial role in determining the extent of systemic risk and the propagation of financial contagion within the banking network. Higher levels of interconnectedness increase the risk of contagion, underscoring the importance of understanding and monitoring interbank linkages. Additionally, our results highlight the importance of government intervention in reducing the effects of financial contagion. We observed that higher levels of government intervention led to a reduction in the number of banks that default, indicating the effectiveness of policy measures in stabilizing the banking system during crises. Banks that do not default, especially those without government intervention, demonstrate resilience in the face of financial stress. Their ability to maintain stability without external assistance underscores the importance of sound risk management practices and capital buffers in protecting against systemic shocks. The information obtained from our simulations has significant implications for policymakers and regulators. Understanding the dynamics of financial contagion and the factors influencing systemic risk can inform the development of more robust regulatory frameworks and risk management strategies aimed at enhancing the stability and resilience of the financial system.

As with any study, it is important to acknowledge its limitations. In our case, the model we developed using agent-based modeling approach may not fully capture the complexity of the banking system or the interactions between banks and other financial institutions. Additionally, some aspects of the banking sector and external factors that could influence systemic risk may not have been included. Furthermore, the way we modeled government intervention may not always reflect the reality of government policies and actions during a financial crisis.

Regarding future research directions, our model could be further developed and expanded to include a wide range of both banking and non-banking financial institutions and their interactions.

Acknowledgements The work is supported by a grant of the Romanian Ministry of Research, Innovation and Digitalization, project CF 178/31.07.2023—‘JobKG—A Knowledge Graph of the Romanian Job Market based on Natural Language Processing’. This study was co-financed by The Bucharest University of Economic Studies during the PhD program.

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References





1. Franch, F., Nocciola, L., Vouldis, A.: Temporal networks and financial contagion. *J. Financ. Stab.* **71**, 101224 (2024). <https://doi.org/10.1016/j.jfs.2024.101224>
2. Camera, G., Gioffré, A.: Financial contagion and financial lockdowns. *J. Econ. Behav. Organ.* **218**, 613–631 (2024). <https://doi.org/10.1016/j.jebo.2024.01.002>
3. Allen, F., Gale, D.: Financial contagion. *J. Polit. Econ.* **108**, 1–33 (2000). <https://doi.org/10.1086/262109>
4. Bae, K.-H., Karolyi, G.A., Stulz, R.M.: A new approach to measuring financial contagion. *Rev. Financ. Stud.* **16**, 717–763 (2003). <https://doi.org/10.1093/rfs/hhg012>
5. Lai, J., McNelis, P.D.: Financial contagion among the GSIBs and regulatory interventions. *J. Financ. Stab.* **72**, 101252 (2024). <https://doi.org/10.1016/j.jfs.2024.101252>
6. Nica, I., Delcea, C., Chiriță, N., Ionescu, Ș.: Quantifying impact, uncovering trends: a comprehensive bibliometric analysis of shadow banking and financial contagion dynamics. *Int. J. Financ. Stud.* **12** (2024). <https://doi.org/10.3390/ijfs12010025>
7. Akhtaruzzaman, M., Boubaker, S., Sensoy, A.: Financial contagion during COVID–19 crisis. *Financ. Res. Lett.* **38**, 101604 (2021). <https://doi.org/10.1016/j.frl.2020.101604>
8. Pallante, G., Guerini, M., Napoletano, M., Roventini, A.: Robust-less-fragile: tackling systemic risk and financial contagion in a macro agent-based model (2024)
9. Gai, P., Kapadia, S.: Contagion in financial networks. *Proc. R. Soc. A.* **466**, 2401–2423 (2010). <https://doi.org/10.1098/rspa.2009.0410>
10. Helbing, D.: Agent-based modeling. In: *Social Self-Organization*, Helbing, D., (ed.) *Understanding Complex Systems*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 25–70 (2012). ISBN 978-3-642-24003-4
11. Ionescu, Ș., Delcea, C., Chiriță, N., Nica, I.: Exploring the use of artificial intelligence in agent-based modeling applications: a bibliometric study. *Algorithms* **17**, 21 (2024). <https://doi.org/10.3390/a17010021>
12. Ionescu, Ș., Chiriță, N., Nica, I., Delcea, C.: An analysis of residual financial contagion in Romania’s banking market for mortgage loans. *Sustainability* **15**, 12037 (2023). <https://doi.org/10.3390/su151512037>
13. Ionescu, Ș., Nica, I., Chiriță, N.: Cybernetics approach using agent-based modeling in the process of evacuating educational institutions in case of disasters. *Sustainability* **13**, 10277 (2021). <https://doi.org/10.3390/su131810277>
14. Chiriță, N., Delcea, C., Nica, I., Craciunescu, S.; Ionescu, Ș.: Financial contagion and identifying speculative frenzies: unraveling price bubbles in cryptocurrency markets. *Theor. Appl. Econ. Asociația Generală a Economistilor din România–AGER*, **0**(3(636), A), 21–40 (2023)
15. Gai, P., Kapadia, S.: Contagion in financial networks. *SSRN J.* (2010). <https://doi.org/10.2139/ssrn.1577043>
16. National Bank of Romania Banks in Romania. <https://www.bnro.ro/Banks-in-Romania-1724.aspx>. Accessed on 4 Feb 2024
17. Rand, W., Rust, R.T.: Agent-based modeling in marketing: guidelines for rigor. *Int. J. Res. Mark.* **28**, 181–193 (2011). <https://doi.org/10.1016/j.ijresmar.2011.04.002>
18. Delcea, C., Cotfas, L.-A., Salari, M., Milne, R.: Investigating the random seat boarding method without seat assignments with common boarding practices using an agent-based modeling. *Sustainability* **10**, 4623 (2018). <https://doi.org/10.3390/su10124623>
19. Chan-Lau, J.: ABBA: an agent-based model for the banking system. *IMF Working Papers*, WP/17/136 (2017)
20. Steinbacher, M., Raddant, M., Karimi, F., Camacho Cuenca, E., Alfaro, S., Iori, G., Lux, T.: Advances in the agent-based modeling of economic and social behavior. *SN Bus. Econ.* **1**, 99 (2021). <https://doi.org/10.1007/s43546-021-00103-3>
21. Chakraborti, A., Toke, I.M., Patriarca, M., Abergel, F.: Econophysics review: II agent-based models. *Quant. Financ.* **11**, 1013–1041 (2011). <https://doi.org/10.1080/14697688.2010.539249>
22. Halaj, G.: Agent-based model of system-wide implications of funding risk. *European Central Bank, Working Paper Series* (2018)

23. Delcea, C., Cotfas, L.-A., Chiriță, N., Nica, I.: A two-door airplane boarding approach when using apron buses. *Sustainability* **10**, 3619 (2018). <https://doi.org/10.3390/su10103619>
24. Wilensky, U., Rand, W.: An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with Netlogo. MIT Press: Cambridge, USA (2015). ISBN 978-0-262-73189-8

Chapter 24

Modeling Socio-economic Phenomena: Some Critical Thoughts and a Positive Example



Jan W. Owsinski , Jarosław Stańczak , Przemysław Śleszyński ,
and Rafał Wiśniewski 

Abstract The paper presents some general thoughts on the subject of modeling of socio-economic phenomena and processes, and the problems, related to such modeling. In particular, the subject of reliability and verifiability of models is taken up. The aspects of data and methodology, as well as theory are touched upon. Against this background a straightforward modeling exercise is shown of migrations in Poland at the municipality level. For the data on some 2500 Polish municipalities over 20 years in the twenty-first century basic models are built, founded on the fundamental hypothesis of relation between unemployment rate and migration flows. It is shown that not only the hypothesis gets confirmed, even if with a great variety of divergences, but quite plausible conclusions can be drawn from the results obtained, leading also to deeper analyses.

24.1 Introduction

This paper is meant, first, to highlight some of the issues, pertaining to the vast area of modeling of socio-economic phenomena and processes. This is done in the perspective of the experience of the authors and the current capacities of the broadly conceived ICT. We do not, certainly, attempt to provide a holistic view of the domain, but rather to pinpoint some of the essential issues in its effectiveness. This is done through presentation of a sequence of model structures and applications, moving from the simplest (and most reliable) towards those perceived as more refined, more complex, and of greater scope in terms of economic, social and geographic reach. The accompanying discussion refers to the essential issues of data, their availability

J. W. Owsinski (✉) · J. Stańczak
Systems Research Institute, Polish Academy of Sciences, Warszawa, Poland
e-mail: owsinski@ibspan.waw.pl.com

P. Śleszyński · R. Wiśniewski
Institute of Geography and Spatial Organization, Polish Academy of Sciences, Warszawa, Poland

and reliability, methodologies applied and theoretical explanations. These, of course, have an important, if not decisive, bearing on the applicability of respective models.

Then, we proceed to present an example of a model that turns out “sufficiently” positive, and that in spite of its straightforwardness as opposed to the apparent complexity of the phenomenon modelled¹. We model, namely, migrations at the level of municipalities in Poland, the lowest administrative units, totaling roughly 2500. We speak only of internal migrations, with data spanning over 20 years, from 2003 to 2022. In fact, we dispose of full inter-municipality migration matrices for each of these years (in principle: approximately 2500×2500 matrices, but, actually, not very dense ones). Yet, for the modeling exercise here considered and illustrated we use for each municipality the net result of the summary immigration and outmigration numbers, i.e. total net migration for each municipality, meaning that we deal with a vector of some 2500 data items.

The basic hypothesis verified was the dependence of migration flows upon the difference in unemployment rates, here, given that we do not consider the inter-municipality flows explicitly, reduced to just levels (rates) of unemployment.

The results obtained, even though statistically not very spectacular, first, fully confirm the hypothesis, then, provide a very interesting spatial image in terms of conformity with the model and the magnitudes and signs of its errors, and thus—offer a highly valuable instrument not only for further, more in-depth cognitive analysis, but also for policy making.

This means that even now, in the age of AI and cloudy computations, quite basic tools, if smartly applied, can bring highly valued results, while at least some of the very sophisticated methodologies and devices not always fulfil their promise and have to be used with great care.

24.2 Socio-economic Modeling: Some Fundamentals

24.2.1 *The Data and the Representation of Reality*

A model is a representation of a certain fragment of reality, such that it is, at the same time, sufficiently accurate for a given purpose, and allows for performing analyses and manipulations, which would be difficult, too costly, or outright impossible to be done directly to the respective fragment of reality. This quasi-definition is fulfilled also by physical models, like those used, for instance, in aerodynamics. Yet, we refer here to computer-based models, mostly (although not necessarily) founded on some kind of mathematical structure, even if sometimes quite modest.

It can be said that a computer-based model is made of three components: (1) the structure, which often takes a mathematical form; (2) the data, which fill the

¹ This part of research was financed by the National Science Centre in Poland, grant no. 2021/41/B/HS4/02035, “Influence of internal migrations on socio-economic development of Poland after 1989”.

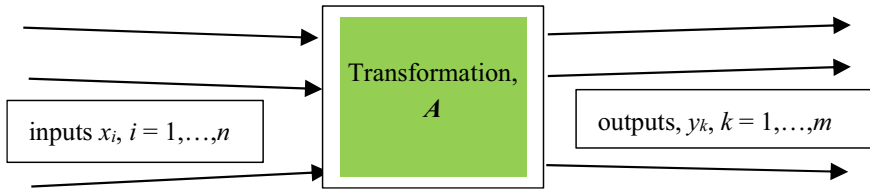


Fig. 24.1 A schematic view of a model for a very wide class of processes and phenomena

structure, and (3) the algorithmic and software implementation (although sometimes it is worthwhile to yet distinguish the latter two).

For a very wide class of processes and phenomena we can see almost any corresponding socio-economic model as sketched in Fig. 24.1.

Possibly one of the simplest structures of this kind is that of a person’s purse or of private money collection for some purposes. So, for instance, it is typical for Polish educational system that parents of pupils in a class collect money for some extra-curricular purposes. The respective form of the “model” is

$$\sum_i x_i = \sum_k y_k + r \tag{24.1}$$

where index i denotes i -th source (donor) of means, x_i is the corresponding amount, k denotes k -th purpose, and y_k its cost, while r is the rest (theoretically could even be negative if there are some outstanding expenditures to be made), which might also be treated as yet another y_k . Here, the transformation of inputs into outputs is totally “inside” the decision making body and may take on a very wide variety of forms, indeed.

Trivial as it may appear, it is, actually, not. First, (24.1) is a model of any accounting activity, however complex it may be. Suffice to add an assumption, concerning time. Second, it is only a slight simplification of a model, in which inputs x_i get transformed into outputs y_k in an economic activity, and that for the objective of economic profit, efficiency, or whatever else one may assume. If we treat the set of x_i as a vector, \mathbf{x} , and of y_k as a vector, \mathbf{y} , then the transformation, at least in approximation (sufficiently close to a definite known and somehow measurable “point of work”), can be expressed as

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{24.2}$$

where \mathbf{A} is the matrix of “technical coefficients”, linking individual inputs with individual outputs. Let us denote these as a_{ik} , meaning, roughly, that input x_i is used up at the rate of a_{ik} in the process of producing the output y_k . Hence, if we wish to get a picture for each output y_k separately, the expression is

$$y_k = \sum_i a_{ik}x_i \tag{24.3}$$

Thus, if we knew the values of a_{ik} , and the assumption of constancy of these coefficients in time and space, at least within certain limits, can be adopted, then we can “predict” the values of y for changing x , and even optimize over x with respect to some quality function $Q(x,y)$, supposedly expressing costs and benefits, linked to x and y .

24.2.2 Identification and Other Problems

There is an enormous volume of literature on both such models and their nonlinear, time-wise and probabilistic extensions, but we wish to take up only a very simple aspect while still sticking to deterministic linearity. Namely: how do we identify a_{ik} ? The answer is simple, and, in fact, leads to an effective procedure, in only two cases:

- I. the processes, described by (24.3), are separable in the sense that for each of $y_k, k = 1, \dots, m$, a separate subset of x_i is used, that is – we deal with subsets of $\{1, \dots, n\}$, denoted I_k , such that $a_{ik} > 0$ for $i \in I_k$, and $a_{ik} = 0$ otherwise; if so, we can actually measure the use of x_i when producing y_k separately for each y_k and obtain the values of a_{ik} ;
- II. we are capable of manipulating each input x_i separately, while keeping all other inputs constant (“ceteris paribus”), changing it, say, by Δx_i , and obtaining the corresponding changes in outputs, Δy_k , thereby obtaining the values $a_{ik} = \Delta y_k / \Delta x_i$.

In the socio-economic reality none of these, except for completely trivial situations, is feasible. Suffice to point out such input to any company’s activity as electric power. Coming back to bookkeeping this is exactly the problem of assigning (or splitting) the overhead or indirect cost or general cost among the particular activities or products (lack of separability).

An extreme, but highly popular and very seriously treated variation of the model here discussed is based on the so-called Input–Output (I/O) tables, or Leontief type model (see [5] and, e.g., [2]). Most countries elaborate through their official statistical systems the I/O tables for the national economies, based, essentially, on inter-branch flows and consumptions of the economy. In this model, inputs and outputs are in principle the same, but “moved in time”, in a certain simplification:

$$x_{t+1} = Ax_t \tag{24.4}$$

meaning that we deal with a kind of a “closed model” (provided demand is appropriately integrated into x and we appropriately account within it for foreign trade, final consumption, prices etc.). The coefficients used are based on inter-branch flows, elasticities (reactions to change) and prices. The recently popular Computable General Equilibrium models (CGE) are also derived from this basic paradigm (see, for instance, [1] or [3]). These models are frequently very large, involving thousands

of coefficients, which have to be established (if we have, say, only 50 branches / products, we have 2500 entries if they are single, but usually they are multiple...).

Yet, even if we make a much more detailed division into the respective “branches”, for which data can be collected, say, 250 by 250 (meaning already some 60000 coefficients), we still do not deal with individual, relatively homogeneous “products”, and the coefficients are aggregates of rather little known structure, to not to say—reliability and prognostic value. Worse even, as we “go down” with disaggregation, we become dependent upon marginal changes such as a single appearance or disappearance of a plant.

In general, though, one does not use a model in order to check what will happen if we move from a given “working point” by a tiny fraction (say, prices shall change by 1%). What is really not only interesting, but truly important in life—and fascinating—is: *what if a qualitative change occurs?*

This is like asking a question: what if inflation moves from 3 to 18%? what if we have to pay for energy three times as much as today? what if we exit the European Union? what if a war breaks up in the nearest neighborhood? and the like.

The models we speak of do NOT provide any reliable answer in such circumstances. If so, if even a well done bookkeeping is not capable of providing an image of functioning of a relatively modest enterprise, in terms of a_{ik} , and the large models are not only expensive, but also often similarly unreliable, and have to be used with utmost care and caution, is there a sense to socio-economic modelling?

24.2.3 What Can Be Done?

The essential issue—also in the light of the above considerations—is *what do we expect from a model?* Actually: from a definite model, not a model in general. In the socio-economic domain we are not having the luxury of, say, engineering, where experiments may be designed and carried out, with numerous repetitions and variations. We always face high uncertainty, this being well illustrated by the values of the correlation coefficient or the factor loadings. While in engineering correlation of + 0.82 would not suffice to state that two variables are really linked to each other, in socio-economic domain we would agree fully with such hypothesis.

Yet, in more general terms, the situation is similar. First, we must know *what question we are going to ask*. The question itself has to be formulated in a rigorous manner, with *full understanding of its components*, as well as *purposes* of asking this particular question.

In particular, we have to specify the *conditions*, under which we wish to have the answer, and these conditions are often highly hypothetical by themselves. That is why people often refer to the *scenario* modelling, in which a number of conditions are established and for these conditions an answer is provided to a given question. Yet, as we know, none of the scenarios may get fulfilled and our answer will be either wrong, or only highly hypothetical (unless we are justified in making an interpolation between the scenarios, which is rarely true).

The most appropriate use of models in the socio-economic domain is, therefore, to test the hypotheses (and this can also be done with the I/O based models). Do the data and the well-founded knowledge that we have support a definite hypothesis?

We shall now present an instance of exactly such a study.

24.3 Internal Migrations and Their Analysis

24.3.1 The Issue

Migrations constitute one of the most important dynamic factors in modern world. This applies, especially, to the international, or even intercontinental migrations. Still, migrations within a given country also pose very serious problems in terms of economy, society, demography, regional and spatial planning, etc. It is of very high value to know the actual course, the patterns, and the factors behind peoples' movements in space. Yet, there exist definite difficulties, facing respective analyses.

The very first of these is, of course, the problem of data. This problem is closely associated with the working of national administration, the regulations, and the methodologies of the public statistics. One can, of course, imagine a highly regulated country, in which true migration is well tracked by means of residence registration. Even in such a country, though, many people would avoid registering, for various reasons, and not only this, but a number of other aspects, associated with migration, may get unnoticed (e.g. where the migrant works? where the migrant's children go to school? where does s/he do the shopping? etc.).

Then, there is the question of the mechanism that propels migration. What is this mechanism? How strong it is? Can we really measure its impact? Can it be controlled? And so on. The set of answers to these questions amounts to a model of migration, to be proposed and identified (Fig. 24.2).

It is generally accepted that migration is propelled by the *perceived difference of the conditions of living between the origin and the destination* (examples of such models are provided, in particular, in the already classical [8], but also, quite recently, for instance in [6]). In some, extreme, cases this difference amounts to survival. But how is it expressed and measured?

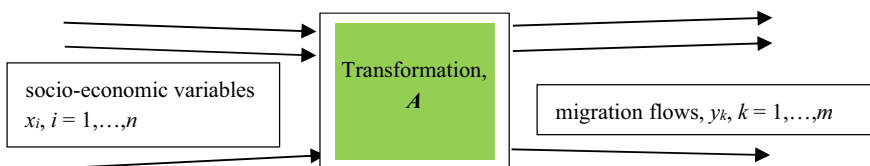


Fig. 24.2 A schematic view of a model of migration flows

24.3.2 *The Model and the Data*

We shall present here a small study for quite a fundamental, basic, and straightforward model. Namely, we assume its general following form:

$$y_{ij} = a_0 + a_1x_{ij} + \varepsilon_{ij} \quad (24.5)$$

where y_{ij} is the flow of migrants from geographical (administrative) unit i to unit j , x_{ij} is a measure of difference in “perceived conditions of living”, a_0 and a_1 are coefficients, and ε_{ij} is the error of the model for a given pair of units.

For pragmatic reasons, we shall yet simplify this basic model. Namely, we wish to model migrations inside Poland at the level of municipalities. There are some 2500 municipalities in Poland, and hence there are roughly $6.25 \cdot 10^6$ pairs (i,j) of municipalities. It is, of course, possible to identify a model for such data, but it is not that simple, first, and it is not at all certain that it is worthwhile, second.

The data we use concern residence registrations (in- and out-). They are recorded at municipality level, and we dispose of such data for each year. Our attention is focused on the period 2003–2022 (two decades), and during this period some 13 million residence changes were registered inside Poland (see, e.g., [7]). The matrices of migration flow data for each year t , i.e. the values of $\{y_{ij}\}_t$, are, first, relatively sparse (high proportion of pairs, between which no flow occurs), and, second, the values are highly diversified, from single occurrences to thousands. On the top of this, the data are insofar inaccurate as many people do not register their residence changes, or, vice versa, if they do, this does not necessarily correspond to reality, for a variety of reasons [4]. It is considered, though, that even if not accurate, the data reflect at least correctly the diversity and the proportions of flow magnitudes.

So, the simplified model, which was actually subject to analysis, had the following form:

$$y_i = a_0 + a_1x_i + \varepsilon_i \quad (24.6)$$

where y_i is the net migration to/from a given municipality (immigration minus outmigration), in relative terms, i.e. divided by the population number, and x_i is the level of unemployment in a municipality (number of registered unemployed divided by the number of persons in working age). So, we deal with some 2500 data items per year, and try to verify a very simple hypothesis: *migration depends (linearly) upon the level of unemployment*.

Here, the magnitude of unemployment stands for the “perceived difference...” etc. It can be said that unemployment is used as a proxy for the entirety of socio-economic conditions, representing the overall socio-economic situation of a given area (in relative terms).

It was, of course, assumed from the very beginning that this model may have to be somehow complemented or modified, in order to obtain the possibly good results. Yet, there is, of course, an almost innumerable variety of ways to complicate a model.

First—one has to check the basic, possibly the simplest version and see its adequacy and capacity.

24.3.3 The Results

Table 24.1 shows an excerpt from the results obtained for the model analysed, i.e. first of all the values of coefficients for the consecutive years, and also a characterization of errors.

The table is followed by Fig. 24.3, illustrating graphically a single case (i.e. for a single year, here the last one considered, 2022).

Altogether, on the basis of these results, it can be concluded that the basic hypothesis was positively verified—the higher the unemployment, the lower the net migration of a unit, starting from some positive value (for the putative zero unemployment) downwards, crossing zero at some point and becoming negative for higher unemployment rates.

Table 24.1 Model data for consecutive years

Year	a_0	a_1	Model error		
			Max	Mean	Sum/sum of absolute values
2003	0.004977	− 0.031785	0.072422	0.004537	$1.0930 \cdot 10^{-14} / 11.246$
2004	0.005427	− 0.033655	0.071667	0.004762	$1.5166 \cdot 10^{-15} / 11.805$
2005	0.005539	− 0.040604	0.060012	0.004496	$− 4.0279 \cdot 10^{-14} / 11.145$
2006	0.006174	− 0.052696	0.074140	0.004773	$− 9.4792 \cdot 10^{-15} / 11.831$
2007	0.006665	− 0.069522	0.076484	0.005331	$− 1.0776 \cdot 10^{-14} / 13.216$
2008	0.005200	− 0.063325	0.052913	0.004193	$1.3080 \cdot 10^{-14} / 10.395$
2009	0.005798	− 0.057930	0.053129	0.004343	$3.3701 \cdot 10^{-14} / 10.765$
2010	0.005185	− 0.061323	0.071134	0.004378	$− 7.9546 \cdot 10^{-15} / 10.852$
2011	0.005820	− 0.061135	0.049553	0.004350	$− 1.4339 \cdot 10^{-14} / 10.784$
2012	0.005343	− 0.054087	0.039587	0.004008	$1.7992 \cdot 10^{-14} / 9.9357$
2013	0.005919	− 0.059938	0.045612	0.004211	$1.7404 \cdot 10^{-14} / 10.439$
2014	0.004630	− 0.058011	0.038719	0.003872	$2.4911 \cdot 10^{-15} / 9.7990$
2015	0.003982	− 0.061810	0.033928	0.003795	$− 5.3896 \cdot 10^{-15} / 9.4079$
2016	0.003399	− 0.061002	0.046455	0.003809	$− 3.0556 \cdot 10^{-15} / 9.4433$
2017	0.003089	− 0.071200	0.044713	0.003998	$− 1.9382 \cdot 10^{-14} / 9.9099$
2018	0.003324	− 0.084093	0.071719	0.004325	$− 7.7707 \cdot 10^{-15} / 10.722$
2019	0.003045	− 0.093832	0.066853	0.004593	$− 8.7957 \cdot 10^{-15} / 11.386$
2020	0.003705	− 0.075329	0.039780	0.004123	$− 6.9174 \cdot 10^{-15} / 10.222$
2021	0.003753	− 0.086029	0.043339	0.004593	$3.2387 \cdot 10^{-15} / 11.387$
2022	0.003861	− 0.085469	0.038152	0.004506	$− 1.1605 \cdot 10^{-15} / 11.171$

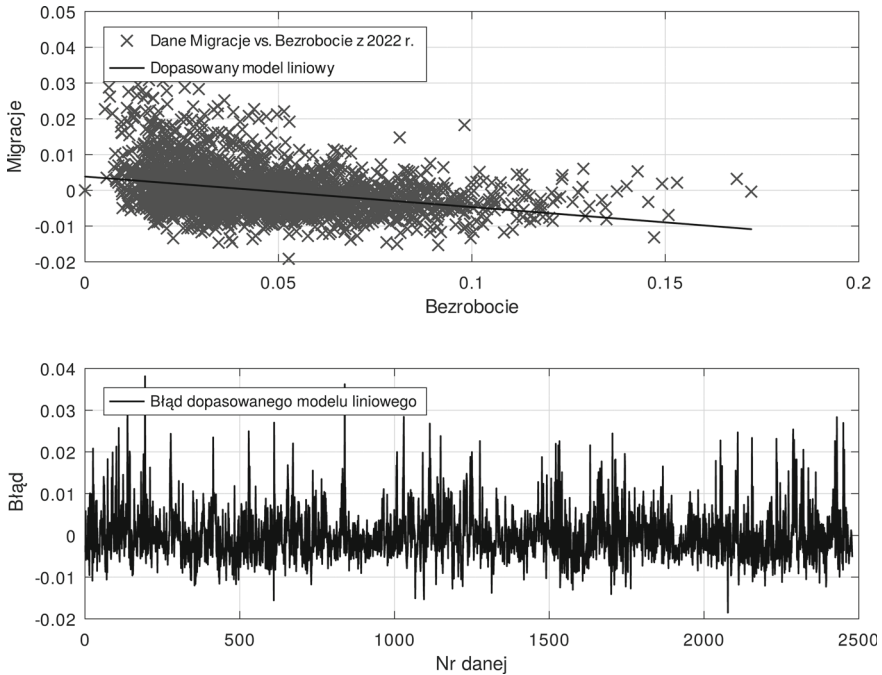


Fig. 24.3 An example of model results for year 2022: the cloud of points, corresponding to municipalities along the identified straight line of the model (upper diagram) and the errors for particular municipalities (lower diagram)

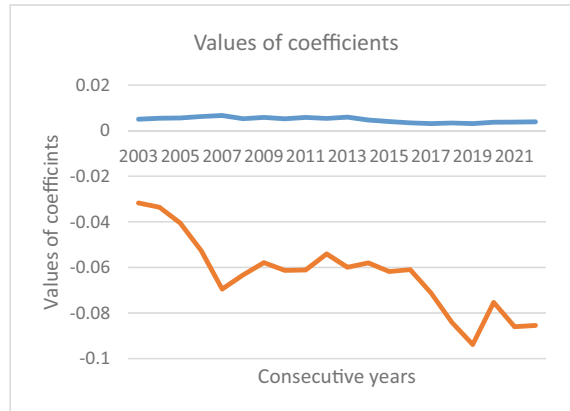
The statistical properties of the models obtained are not very imposing, the correlation coefficient between the model and the data being very roughly at + 0.4, and this could certainly be expected from both the character of the processes, the nature of data, and the simplicity of the model assumed. Yet, irrespective of this, the sequence of values of model coefficients along time, as provided in Table 24.1, shows the values that are by no means haphazard, but display quite consistent trend or trends, which can be seen in Fig. 24.4.

Thus, not only the basic, quite simple model turned out to quite adequately represent a truly complex socio-economic reality, but, moreover, and very importantly, provided material for further studies. We shall illustrate these with just an example in the next section.

24.3.4 *Spatial Analysis*

As already mentioned before, the data on migrations are only an approximation, in almost all cases an underestimation of the true flows. This fact, and not only the quite basic nature of the model, may be the source of errors, which are well visible

Fig. 24.4 The values of model coefficients along time, as shown in Table 24.1



in Fig. 24.3. Before—and perhaps even partly instead of—engaging in formulation and identification of more intricate models, an analysis of these errors is in place.

Of special interest is the distribution of the errors over space, at exactly the level of municipalities. Such an analysis was performed for the majority of the identified models, and the result is shown in here in Fig. 24.5.

The meaning of Fig. 24.5 is obvious: the errors form truly strikingly regular spatial shapes, indicating primarily the areas surrounding the larger cities, included, actually, in respective metropolitan regions. The reason is largely obvious: migrants settle in the suburban municipalities of larger centers, although they may work or take education in the respective centers. The gradient of the error has a clear spatial expression, corresponding to the degree of peripherality, on the one hand, and to the form of the model, on the other.

This motivated further analyses, in which, in particular, various definitions of distance between the municipalities and the neighboring centers were added as a second independent variable. The quality of the models, in terms of correlation coefficient values, increased, but in qualitative terms they were somewhat less clear than for the basic model.

24.4 Conclusions and Further Work

Modeling of socio-economic processes and phenomena is very difficult, much more so than in the case of sciences, where experiments can be conducted under controlled conditions and where measurements can be made (almost) without (gross and unknown) errors. Hence, the models built in the socio-economic realm have to be treated in a special manner. First of all, this concerns the purpose of the model building and the treatment of their results.

We show here, however, how a very straightforward model of a complex socio-economic phenomenon can provide valuable information and the material for further

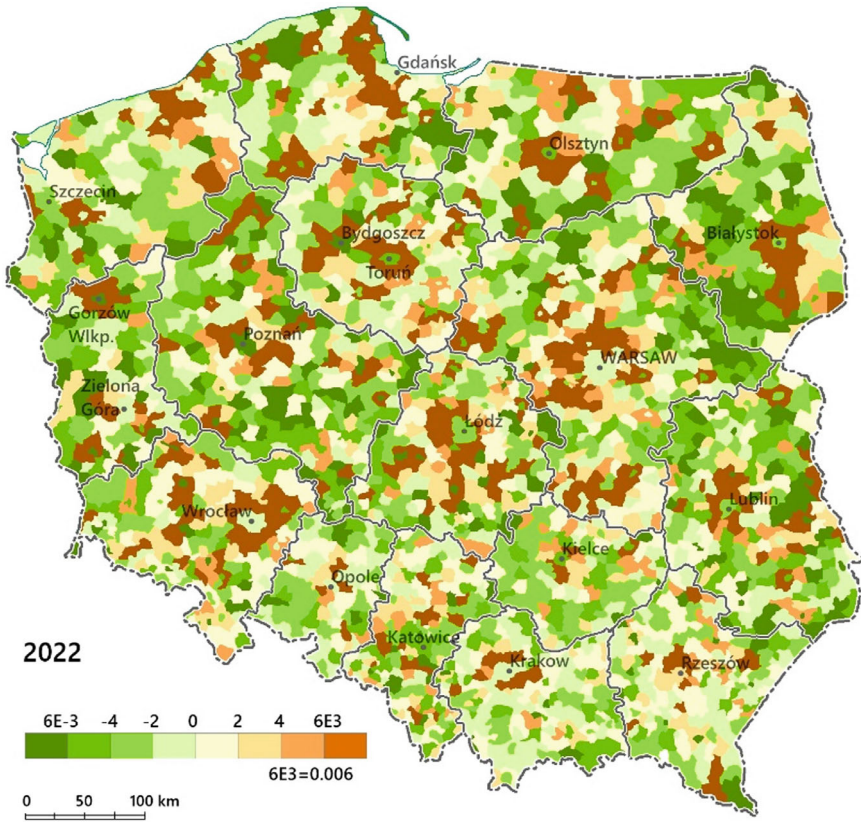


Fig. 24.5 Spatial distribution of model errors across polish municipalities (2022)

study. Not only is the hypothesis behind this basic model confirmed, but also vested with corresponding telling quantitative characteristics. The spatial analyses, associated with the results of this model, give a strikingly clear and interpretable image, even if the statistical quality of the model is still not very high.

The directions of further analyses are very obviously indicated, even at the level of the shorthand here presented: (a) with reference to Fig. 24.4: the association of the changes in the coefficients of the model with the macroeconomic conditions over time; (b) with reference to Fig. 24.5: the association of the model error distribution with the potential errors in the original data; or (c) the temporal evolution of the error shapes in space (hypothetically corresponding to the dynamics of the functional urban regions).

Note that in the above proposals we do not refer to the obvious scientific endeavor of the development of the potential more intricate models, possibly better representing the reality of the migration flows.



References

1. Burfisher, M.: Introduction to computable general equilibrium models. Cambridge University Press, Cambridge (2011)
2. Dietzenbacher, E., Lahr, M.L. (eds): Wassily leontief and input–output economics. Cambridge University Press (2004)
3. Ginsburgh, V., Keyzer, M.: The Structure of applied general equilibrium models. MIT Press (1997)
4. Jończy, R.: Problem nierejestrowanej emigracji definitywnej (emigracji zawieszonyj) w badaniu procesów społeczno-gospodarczych na obszarach wiejskich (Problem of unregistered definitive emigration (suspended emigration) in the study of the socio-economic processes in rural areas in Polish). *Prace Naukowe UE, Wrocław* **360**, 11–18 (2014)
5. Leontief, W.W.: Input-output economics, 2nd edn. Oxford University Press, New York (1986)
6. Mbachu, H.I.: Modelling the effect of gross domestic product (GDP) per capita on population growth, youth unemployment and net migration. *Int. J. Math. Stat. Inven.* **10**(1), 12–21 (2022)
7. Śleszyński, P., Heffner, K., Solga, B., Wiśniewski, R.: Geographical research on migration in Poland after 1989. *Przegląd Geogr.* **93**(4) (2021)
8. Todaro, M.P.: A model of labor migration and urban unemployment in less developed countries. *Am. Econ. Rev.* **59**(1), 138–148 (1969)

Chapter 25

Enhancing Social Acceptance in Microgrid Implementation: A Share-of-Choice Optimization Approach



Emmanuel Fragniere , Francesco Moresino , Sarah Sandoz ,
and Nabil Abdennadher 

Abstract The role of social acceptance is crucial in securing the commitment and cooperation of prosumers (i.e. contraction of producer and consumer), which are essential for the optimal operation and regulation of micro smartgrids. Although designing for social acceptance in smart grid diffusion within energy communities is increasingly recognized in both research and policy domains, its widespread implementation and empirical validation remain limited. This study aims to promote the widespread adoption of the micro smart grid paradigm to advance energy transition initiatives. To achieve this goal, it is necessary to actively engage with prosumers, local communities, and regulatory bodies, demonstrate sustained commitment from stakeholders, and ensure optimal microgrid orchestration through social acceptance. To tackle these challenges, we developed a Share-of-Choice optimization model that aims to determine the most cost-effective setup for smart grid energy communities. This method starts by identifying social characteristics (i.e. salient attributes), such as the frequency of community meetings, to encourage efficient co-participation among prosumer roles and related demand-side management practices. Next, using conjoint analysis, we will calculate part-worth utilities to inform the Share-of-Choice

E. Fragniere (✉) · S. Sandoz
University of Applied Sciences Western Switzerland, HES-SO Valais-Wallis, Sion, Switzerland
e-mail: emmanuel.fragniere@hevs.ch

S. Sandoz
e-mail: sarah.sandoz@hevs.ch

F. Moresino
University of Applied Sciences Western Switzerland, HES-SO HEG Geneva, Carouge,
Switzerland
e-mail: francesco.moresino@hesge.ch

N. Abdennadher
University of Applied Sciences Western Switzerland, HES-SO HEPIA Geneva, Geneva,
Switzerland
e-mail: nabil.abdennadher@hesge.ch

optimization models. These models integrate cost structures and prosumer profiles to design roles that maximize social acceptance in achieving CO₂ reduction goals. The model is currently theoretical, but we offer an illustrative example based on fictional data to showcase its potential outcomes. By using the Share-of-Choice approach, we aim to offer policymakers and stakeholders practical insights to guide the development and implementation of socially sustainable microgrid initiatives.

25.1 Introduction

The deployment of smart grid technologies and the transition towards more sustainable and autonomous energy consumption require careful attention to social acceptance and community engagement. Existing literature emphasizes the importance of moral values, socio-economic factors, learning processes, and consumer participation in shaping the acceptance and effectiveness of smart grid technologies. Milchram et al. [13] emphasize the impact of moral values on the acceptance of smart grid technologies, specifically highlighting concerns related to privacy, justice, and trust as key drivers or barriers to acceptance. It is crucial to address these moral considerations in order to foster acceptance. In a socioeconomic context, assert that smart grid projects have a significant influence on electricity consumers. Although private costs, such as monetary expenditures, are important, it is also crucial to consider social costs, which include consumer perceptions, privacy, cybersecurity, and regulation, as they can hinder technology adoption. Therefore, research should focus on comprehensively elucidating the implications of these social costs. The transition towards smart grid technologies requires social and cultural transformations. Alviyal-Palavicino et al. [1] emphasize the importance of learning processes and reflexivity among stakeholders. The implementation of smart grids depends on community-specific characteristics and stakeholders' adaptability to the evolving context. Community participation is crucial for the sustainability of mini-grid systems, as highlighted by [9]. Social sustainability is often positively correlated with community involvement in the operation and maintenance of smart grid technology. Rathnayaka et al. [15] emphasize the critical roles of consumer participation and management schemes in sustaining smart grid energy-sharing processes. Encouraging and supporting consumer participation is imperative for effective energy-sharing approaches. Klaassen et al. [11] argue that demand-side management is another crucial facet associated with smart grids. They contend that residential smart grid community participation necessitates demand management integration. The importance of engaging consumers as demand management actors is underscored by their roles in overall energy consumption. Achieving a sustainable energy mix requires the integration of renewable energy resources, electrical energy storage, demand response, and electric vehicles. Smart grid success is dependent on civil society's social acceptance, as highlighted by these authors. Finally, [2] suggest addressing community governance, communication, and management issues to fully realize the benefits of smart grids. Smart grid

design should prioritize maximizing grid resource utilization, ensuring reliability, and providing security.

Therefore, it is essential to integrate aspects of social acceptance into any smart grid project, along with technical and technological considerations, to ensure long-term manageability of these smart grid system options in order to have a real impact on the energy transition. In this research, we have developed a Share-of-Choice model to quantify social acceptance regarding certain CO₂ targets. We explain thus now the notion of conjoint analysis and how its resulting input will be used to feed the theoretical model that has been developed.

The use of conjoint analysis in service design optimization was a significant development in the 1970s. Zufryden et al. [16] was a pioneer in this field, introducing a service design optimizer based on conjoint analysis principles. His model aimed to maximize the Share-of-Choice and was formulated as a mixed integer linear program. Easton [5] expanded upon the Share-of-Choice model by integrating production considerations, which laid the groundwork for further advancements [10, 12]. Debély [3, 6] developed a meta-model that combined the Share-of-Choice model with a production model, merging conjoint analysis with operations research techniques. This method has been shown to be effective in various domains [7]. However, its complexity is a significant limitation, especially for non-specialists, as it requires a solid understanding of operations research principles for model construction. Recent efforts have aimed to mitigate this challenge. Moresino [14] introduced the profit-of-choice synthetic model as a user-friendly alternative to intricate modeling. This development democratizes the method, making it accessible even to individuals lacking specialized expertise in operations research. The streamlined approach allows for broader audiences, including policymakers and stakeholders, to engage with and comprehend the methodology, facilitating its application in diverse contexts.

25.2 The Case Study

In this section, we present the case study that will be used to illustrate our method.

This case study is kept small and easy for didactical reasons. We suppose that a community has to design a new smart grid. Table 25.1 gives the salient attributes and their levels (see for the qualitative study where they come from [8]). In our example, we assume that the smart grid has to be self-financed. This means that the price is fully determined by the choice of the levels of the other attributes. That is why this important attribute is not included in the list given in Table 25.1. However, when conducting the survey to determine the utility functions (named part-worth in conjoint analysis), the price must be included in the card system as it is a primordial attribute for the consumer. Note also that the attribute “meeting” has an impact on social acceptance. However, before being in a real situation, the respondent can barely imagine how it is important. This impact of this attribute can probably be

Table 25.1 Salient attributes and their levels

Attribute	Attribute level	Description
Electricity	Photovoltaic both external	Not connected to the grid, only photovoltaic panel connected to the grid with photovoltaic panel connected to the grid without photovoltaic panel
Battery	With without	With battery (mandatory in case of electricity = photovoltaic) without battery
Pricing	Flat dynamic	Flat rate prices are adapted to supply and demand over the current period
Quota	Yes no	Each household has a quota no quota
Meeting	Frequent rare	Community meetings on electricity-related issues are held once a month community meetings on electricity-related issues are held once a year

Table 25.2 Design for the alternative

Attribute	Attribute level
Electricity	External
Battery	Without
Pricing	Flat
Quota	No
Meeting	Rare

better estimated by conducting a separate study rather than including this attribute in the card system.

We assume that the alternative choice for the community is to connect to the electrical grid and not make any efforts with respect to emissions reduction. The design of this alternative is given in Table 25.2. The social acceptance is the proportion of users that prefer the design of the new smart grid than the alternative design. We assume the part-worths have been estimated with a survey based on conjoint analysis techniques. We aim to design the best grid considering costs, CO₂ emission reductions, and social acceptance.

25.3 The Model

In this section, we introduce a meta-model that combines a Share-of-Choice model with a synthetic production model. In this meta-model, the synthetic production model is characterized by the costs associated with each attribute’s level of service. For the algebraic formulation of the model, we will use the following sets:

- Respondent $k = 1 \dots K$,
- Attribute $i = 1 \dots I$,
- Attribute level $j = 1 \dots J_i$.

For the part-worth, let $u(k,i,j)$ be the utility of respondent k for attribute i and attribute level j . Let $u_c(k)$ be the utility perceived by respondent k for the alternative design given in Table 25.2. Let $u_m(k)$ be the minimum increase of utility to convince respondent k to accept another design than the alternative design.

For the decision variables, let $X(i,j)$ be the service configuration. $X(i,j) = 1$ if attribute i is set to level j and $X(i,j) = 0$ otherwise. For the dependent variables, let $Y(k)$ be the preference for respondent k . $Y(k) = 1$ if respondent k prefers the new design and $Y(k) = 0$ otherwise. The Share-of-Choice model can thus be written as follows. For each respondent $k = 1 \dots K$, the following two inequalities must hold

$$\sum_{i=1}^I \sum_{j=1}^{j_i} u(k, i, j) \cdot X(i, j) \geq u_c(k) + u_m(k) + (Y(k) - 1) \cdot M, \tag{25.1}$$

$$\sum_{i=1}^I \sum_{j=1}^{j_i} u(k, i, j) \cdot X(i, j) \leq u_c(k) + u_m(k) + Y(k) \cdot M, \tag{25.2}$$

where M a sufficiently large constant These two equations insure that a respondent is counted as having accepted the new design if, and only if, the perceived utility of the new design is greater enough compared to the alternative design. For each attribute $i = 1 \dots I$, we must have the normalization

$$\sum_{j=1}^{j_i} X(i, j) = 1. \tag{25.3}$$

This equation insures that each attribute is set to one, and only one, level. Let $w(k)$ be the weight of respondent k , with $\sum_{k=1}^K w(k) = 1$. The social acceptance, which is defined as the weighted proportion of the preferences, is then

$$A = \sum_{k=1}^K w(k) \cdot Y(k). \tag{25.4}$$

Let A_{\min} be the minimal social acceptance desired, this requirement is written with the constraint

$$A \geq A_{\min}. \tag{25.5}$$

To compute the CO₂ reduction, we take as reference scenario the alternative design (see Table 25.2). Let $r(i,j)$ be the CO₂ reduction if attribute i is set to level j . The total CO₂ reduction is then

$$R = \sum_{i=1}^I \sum_{j=1}^{j_i} r(i, j) \cdot X(i, j). \tag{25.6}$$

Let R_{\min} be the minimal CO₂ reduction desired, this requirement is written with the constraint

$$R \geq R_{\min}. \quad (25.7)$$

For our model, we have the following constraint: if the community is not linked to the grid, it must be equipped with a battery and households must have a consumption quota. The linear programming formalism is very convenient to model such constraints:

$$X(\text{battery,with}) \geq X(\text{electricity, photovoltaic}) \quad (25.8)$$

$$X(\text{quota, yes}) \geq X(\text{electricity, photovoltaic}) \quad (25.9)$$

Finally, the objective is to minimize costs,

$$\min_X \sum_{i=1}^I \sum_{j=1}^{J_i} c(i, j) \cdot X(i, j), \quad (25.10)$$

where $c(i, j)$ represents the costs if attribute i is set to the level j . The system (25.1, 25.10) is a 0–1 programming problem that can be solved with standard techniques.

25.4 Numerical Results

To illustrate our method, we present a numerical experiment with fictive input data detailed in Tables 25.3 and 25.4. The data in Table 25.3 are not part-worth utilities but represent prosumer profiles ranging from 0, 10, and 20, indicating a spectrum from individuals highly inclined to adopt a prosumer posture to those who are not. The numerical values in Table 25.3 are straightforward, just to illustrate the case. We are currently developing a card-based questionnaire to be administered in Switzerland and Sweden. This questionnaire, planned for completion end of 2024, will employ conjoint analysis to derive part-worth utilities for input into the present theoretical model. Anyway, this fictitious setup allows us to observe the model's behavior under various scenarios. Although the data in Table 25.4a and b are also fictitious, their magnitudes are representative of real cases.

We suppose that a community of 20 households must design a new smart grid.

Data for the case study is given in Tables 25.3 and 25.4. The size of the linear program is 34 constraints and 24 variables, of which 21 are 0–1 variables. We conducted a parametric analysis letting varying R_{\min} , the minimal CO₂ reduction desired, without constraint on A_{\min} , the minimal social acceptance desired. We report here the different levels of CO₂ reduction to produce Figs. 25.1 and 25.2: 0, 1.1, 7.7,

Table 25.3 Utility part-worths for each respondent

Attribute level	Respondent									
	1	2	3	4	5	6	7	8	9	10
Photovoltaic	20	20	10	10	0	0	0	0	0	0
Both	10	10	20	20	20	20	20	10	10	10
External	0	0	0	0	10	10	10	20	20	20
With	0	10	0	0	0	0	0	0	0	0
Without	10	0	10	10	10	10	10	10	10	10
Dynamic	10	10	10	10	10	10	0	10	10	0
Flat	0	0	0	0	0	0	10	0	0	10
Yes	10	10	10	0	10	0	0	10	0	0
No	0	0	0	10	0	10	10	0	10	10
Frequent	10	10	10	10	10	10	0	0	0	0
Rare	0	0	0	0	0	0	10	10	10	10

Table 25.4 Data for the model

	(a) Annual costs in CHF	(b) Annual CO ₂ emissions reduction in tons
	Attribute level $c(i, j)$	Attribute level $r(i, j)$
Photovoltaic	22,500	8.58
Both	19,800	7.7
External	18,000	0
With	10,000	0.22
Without	0	0
Dynamic	750	0.022
Flat	0	0
Yes	500	1.1
No	0	0
Frequent	2000	0.022
Rare	0	0

8.8, 9.02, 9.9, 9.922, which correspond to changes in the optimal solutions is 34 constraints and 24 variables, of which 21 are 0–1 variables. We conducted a parametric analysis letting varying R_{\min} , the minimal CO₂ reduction desired, without constraint on A_{\min} , the minimal social acceptance desired. We report here the different levels of CO₂ reduction to produce Figs. 25.1 and 25.2: 0, 1.1, 7.7, 8.8, 9.02, 9.9, 9.922, which correspond to changes in the optimal solutions.

The software used is GLPK (GNU Linear Programming Kit) [4]. Figure 25.1 shows the social acceptance with respect to the annual reduction of CO₂ emissions in tons.

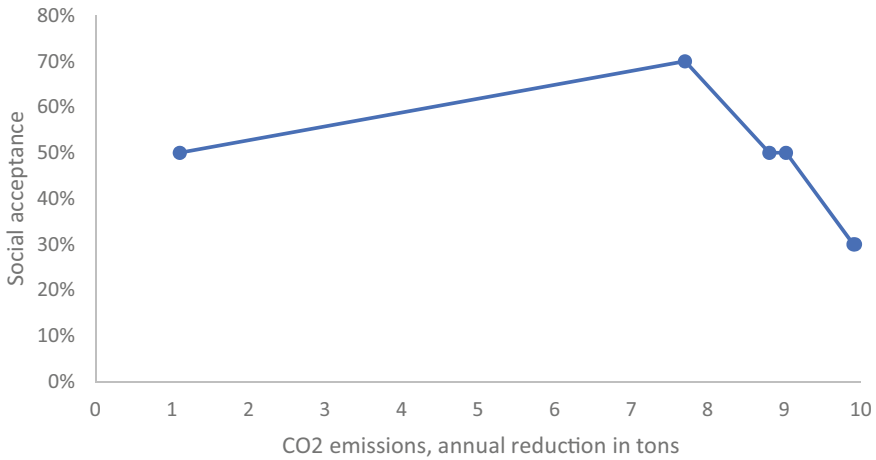


Fig. 25.1 Social acceptance with respect to the annual reduction of CO₂ emissions in tons

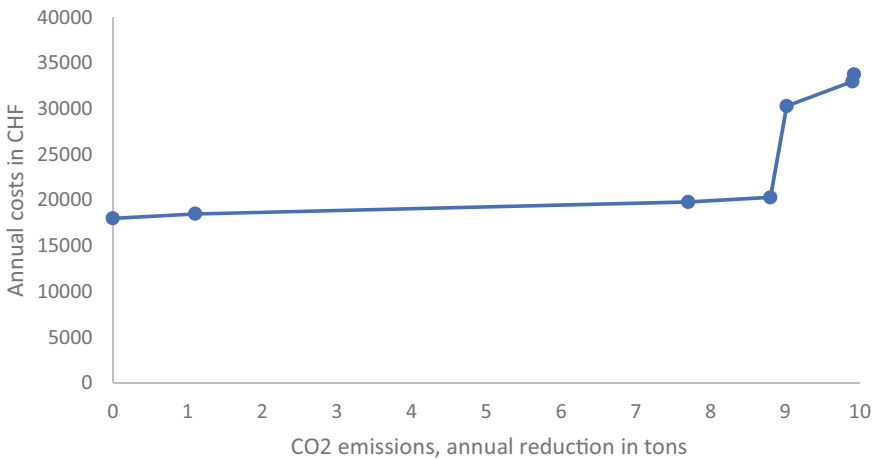


Fig. 25.2 Annual costs in CHF with respect to the annual reduction of CO₂ emissions in tons

Remember that, for this parametric analysis, annual costs are minimized given the constraint on R_{\min} . Figure 25.2 shows the annual costs in CHF with respect to the annual reduction of CO₂ emissions in tons.

In our analysis, it is important to note that the discontinuities in the graphics are a result of utilizing a discrete choice model, where decisions are made among a finite set of options. Upon examining the results, it becomes apparent that one particular design stands out among the others when considering annual costs, social acceptance, and CO₂ emissions. This design exhibits remarkable performance in terms of social acceptance, achieving a high acceptance rate ($A = 70\%$). Although it

Table 25.5 Best design in terms of annual costs, social acceptance and CO₂ emissions

Attribute	Attribute level
Electricity	Both
Battery	Without
Pricing	Flat
Quota	No
Meeting	Rare

ranks lower in terms of CO₂ reduction ($R = 7.7$ tons) compared to the best outcome ($R = 9.9$ tons), it still has a commendable environmental impact. Furthermore, this design is very competitive in terms of annual price, costing 19,800 CHF, only slightly higher than the best possibility, 18,000 CHF. Table 25.5 presents a comprehensive overview of the characteristics of this optimal design, enabling further examination and comparison with other designs. The design’s overall balance between social acceptance, environmental impact, and cost-effectiveness makes it a strong contender for implementation within the microgrid community. Overall, these results emphasize the significance of evaluating and selecting “prosumer role” designs for microgrid systems based on multiple factors, such as social acceptance, environmental sustainability, and economic feasibility. The identified design serves as an illustrative solution that addresses various aspects of microgrid operation and community engagement, providing a foundation for future research and implementation efforts in sustainable energy systems.

25.5 Conclusion

In conclusion, our Share-of-Choice optimization model represents a significant advancement in evaluating the economic concept of social acceptance within microgrid contexts. The ongoing energy transition towards renewable energy necessitates the formation of more energy communities, transforming consumers into prosumers. As governments worldwide advocate for renewable energy mixes, the shift from consumer to prosumer roles becomes imperative. However, the adoption of micro smart grid technology requires more than just technological integration; it demands “social acceptance by design.”

The theoretical Share-of-Choice model that we have developed serves as a valuable tool for assessing and integrating social acceptance factors into microgrid development. In the Summer of 2024, a conjoint analysis survey will be launched in Switzerland and Sweden to gather insights from stakeholders.

The originality of our Share-of-Choice model lies in its application of optimization techniques to the configuration of energy communities. This is achieved through the design of a prosumer role that enables the dispersion of micro smart grids via social acceptance. In contrast to traditional electricity supply models, where logistical and managerial elements are typically managed by a single utility, our project

addresses the complexities arising from the emergence of prosumers. It emphasizes the importance of considering both technical constraints and social acceptability characteristics to ensure the success of microgrid initiatives.

In the subsequent phase of this research, the results of the questionnaire based card survey, obtained through conjoint analysis, will be integrated into the Share-of-Choice model. This integration will provide actionable information to policymakers, facilitating the fostering of socially sustainable autonomous electric systems based on renewable energy, run by energy communities.

References

1. Alvial-Palavicino, C., Garrido-Echeverría, N., Jimenez-Estévez, G., Reyes, L., Palma-Behnke, R.: A methodology for community engagement in the introduction of renewable based smart microgrid. *Energy Sustain. Dev.* **15**(3), 314–323 (2011)
2. Bari, A., Jiang, J., Saad, W., Jaekel, A.: Challenges in the smart grid applications: an overview. *Int. J. Distrib. Sens. Netw.* **10**(2), 974682 (2014)
3. Debély, J., Dubosson, M., Fragnière, E.: The pricing of knowledge-based services: Insights from the environmental sciences. *J. Serv. Res. Spec. Issue* 167–181 (2008)
4. Donoghue, J.R.: Comparison of integer programming (ip) solvers for automated test assembly (ata). *ETS Res. Rep. Ser.* **2015**(1), 1–12 (2015)
5. Easton, F.F., Pullman, M.E.: Optimizing service attributes: the seller's utility problem. *Decis. Sci.* **32**(2), 251–276 (2001)
6. Fragnière, E., Heitz, C., Moresino, F.: The concept of shadow price to monetize the intangible value of expertise. In: Proceedings of the IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE, pp. 1736–1741 (2008)
7. Fragnière, E., Kanala, R., Moresino, F., Reveiu, A., Smeureanu, I.: Coupling techno-economic energy models with behavioral approaches. *Oper. Res.* pp. 1–15 (2016)
8. Fragniere, E., Sandoz, S., Abdenadher, N., Moussa, M., Di Marzo Serugendo, G., Glass, P.: Improving the social acceptability of microgrids. In: 2023 IEEE Renewable Energy and Sustainable E-Mobility Conference (RESEM), pp. 1–5 (2023). <https://doi.org/10.1109/RESEM57584.2023.10236324>
9. Gill-Wiehl, A., Miles, S., Wu, J., Kammen, D.: Beyond customer acquisition: a comprehensive review of community participation in mini grid projects. *Renew. Sustain. Energy Rev.* **153**, 111778 (2022)
10. Gunneç, D., Raghavan, S.: Integrating social network effects in the share-of-choice problem. *Decis. Sci.* **48**(6), 1098–1131 (2017)
11. Klaassen, E., Veldman, E., Slootweg, J., Kling, W.L.: Energy efficient residential areas through smart grids. In: 2013 IEEE Power and Energy society general meeting, IEEE, pp. 1–5 (2013)
12. Liu, M., Zheng, L., Liu, C., Zhang, Z.H.: From share of choice to buyers' welfare maximization: bridging the gap through distributionally robust optimization. *Prod. Oper. Manag.* **32**(4), 1205–1222 (2023)
13. Milchram, C., Van de Kaa, G., Doorn, N., Ku'neke, R.: Moral values as factors for social acceptance of smart grid technologies. *Sustainability* **10**(8), 2703 (2018)

14. Moresino, F.: Optimal design and pricing of e-services. In: Proceedings of the IE2017 International Conference, pp. 396–401 (2017)
15. Rathnayaka, A.D., Potdar, V., Ou, M.H.: Prosumer management in socio-technical smart grid. In: Proceedings of the CUBE international information technology conference, pp. 483–489 (2012)
16. Zufryden, F.S.: A conjoint measurement-based approach for optimal new product design and market segmentation. In: Shocker, A. (ed.) Analytic Approaches to Product and Marketing Planning: Proceedings of Workshop, pp. 100–114, Marketing Science Institute (1977)

Chapter 26

Greening the Workplace for Happiness: Exploring CSR's and Sustainable Practice's Effects on Employees Well-being



Alexandra-Nicoleta Ciucu Durnoi , Adina-Iuliana Jigani ,
Bianca Cibu , Andra Sandu , and Camelia Delcea 

Abstract In contemporary times, there has been a notable surge in companies' attention toward the welfare of their employees. Employee well-being is no longer solely contingent upon financial incentives. Factors such as the work environment, corporate social responsibility (CSR) initiatives, green practices, and organizational culture have emerged as pivotal considerations among employees when evaluating job satisfaction. Through a structured questionnaire administered to employees across various industries, data were collected on perceptions of organizational culture, CSR practices, green initiatives, and self-reported levels of happiness. Utilizing Smart PLS 4 analysis, a robust statistical model was constructed to examine the interplay between these variables. Our research reveals a significant positive correlation between CSR endeavors, encompassing green practices, and employee happiness. Through elucidating the pathways via which CSR and green initiatives impact employee well-being, this study provides valuable insights for fostering a comprehensive approach to corporate sustainability and employee contentment. These results emphasize the strategic necessity for organizations not only to prioritize CSR and green practices but also to integrate them within their organizational culture, thereby enhancing both employee well-being and overall organizational effectiveness.

26.1 Introduction

Starting from a common problem nowadays, we have chosen to investigate through this work, the interest of the employer on the working environment in which the employee works. If until a few years ago, employees were only interested in the

A.-N. Ciucu Durnoi · A.-I. Jigani (✉) · B. Cibu · A. Sandu · C. Delcea
Department of Economic Informatics and Cybernetics, Bucharest University of Economic
Studies, Bucharest, Romania
e-mail: adina.jigani@csie.ase.ro

financial side, lately their attention has been directed to other benefits they receive. Starting from this hypothesis, we chose to investigate employees' views on corporate social responsibility (CSR) practices, organizational culture, self-reported happiness levels, and environmental initiatives.

In the first part of our paper, we describe the literature, to observe the perspectives of other researchers on the studied field, and then we explain the data collection and analysis techniques used. We validated the hypothesis proposed by the methodology with Smart PLS4 because, according to Hensler [1], it is considered to be the most comprehensive software to perform PLS-SEM analyses. Next, we interpreted the results obtained and drew some conclusions about what was achieved with the research conducted.

26.2 Literature Review

CSR is the idea that organizations should be aware of their impact on employees and other stakeholders. CSR actions are not limited to an organization's economic objectives but also include concerns for society, the public good, and the environment [2]. CSR and sustainability are closely linked, as they shift the focus of organizations from purely economic to sustainable environmental and social aspects. This shift can influence competitiveness, reputation, staff, and customer attraction and retention, as well as employee morale, engagement, and productivity [3, 4].

According to the World Bank, CSR is seen as the involvement of businesses in sustainable economic development, working with employees, their families, the community, and society to improve the quality of life and generate both business and development benefits [5].

According to Mu et al. [6] study on happiness management factors, both primary and secondary stakeholder-oriented CSR had positive and significant effects on happiness management. Job satisfaction was also found to be positively and significantly associated with happiness management.

Researchers' interest in employee happiness and well-being is evident, with numerous studies on concepts such as subjective well-being, eudaimonia, and job satisfaction. Eudaimonia is seen as the equivalent of happiness, especially individual happiness. This idea has a distant origin, starting with Aristotle, who argued that eudaimonia is a rational process by which people can pursue what they consider to be a worthwhile life [7].

Tziner et al. [8] found that CSR has a positive impact on both job satisfaction and organizational justice. Furthermore, Wisse et al. [9] found that CSR has a significantly stronger influence on the satisfaction of older employees compared to younger employees.

We also need to talk about the impact of CSR on workplace performance, as corporate success depends on the output of employees, who have been shown to be more effective when the level of burnout is lower [10]. Also, job burnout was felt significantly during the COVID-19 pandemic in various fields of work [11, 12].

In addition to the workplace, the role of CSR or organizational culture on well-being is also identified in other areas such as health [13, 14], technology [15, 16], hospitality [17, 18], or education [19–21].

It is also worth mentioning the role of green practices in finding a path to happiness at work, with numerous studies focusing on the implementation of green well-being in areas such as health [22] hospitality [23, 24], automotive industry [25], etc.

26.3 Methodology

For the current research, a questionnaire with 52 questions was developed and distributed online through the Google form. The 123 answers provided were imported and processed in Excel to have the format desired by the software used. Out of the 52 questions asked, 10 of them correspond to obtaining information about respondents (here can be found data on job, age, educational status, demographic characteristics, etc.). All other questions are in the form of the Likert scale, with answers ranging from 1 to 5, with one stating total disagreement and five agreeing totally. Although the questions were generated randomly, without giving respondents the opportunity to identify latent variables, 16 questions were asked to discover information about organizational culture, 11 for internal CSR, eight about green practices, and seven to identify their degree of happiness.

The answers provided by the respondents were loaded into the Smart PLS4 program [26], which is one of the most widely used PLS-SEM software, having applications in various areas [27, 28]. This study implies two methods for model validation. The initial step involves outer model validation as outlined by [29], succeeded by inner model assessment. The evaluation of the outer model necessitates distinct approaches contingent upon the type of construct under consideration: formative or reflective. Formative indicators are not interchangeable, omitting an indicator is akin to omitting a component of the construct [30]. In contrast, reflective indicators are interchangeable, where the removal of an item does not fundamentally alter the construct. In the evaluation of formative constructs, the process entails verifying indicator multicollinearity, estimating convergent validity, testing the magnitude and significance of indicator weights, scrutinizing the magnitude and significance of indicator loadings, and assessing predictive validity [31]. The evaluation of reflective constructs encompasses five steps: estimating loadings and their associated p-values, assessing indicator reliability, evaluating internal consistency reliability, computing the average variance extracted (AVE), and ensuring discriminant validity through the Heterotrait-Monotrait index (HTMT) [32].

Figure 26.1 presents the suppositions that represent the foundation for the current model according to which Organizational culture impacts Internal CSR (which in turn is formed by the variables Work-life balance and equal opportunities, Adaptability to change, and Occupational health and safety).

In this regard, the first four hypotheses formulated refer to the impact that Clan, Adhocracy, Market, and Hierarchy Culture (H1–H4) have on CSR. The following

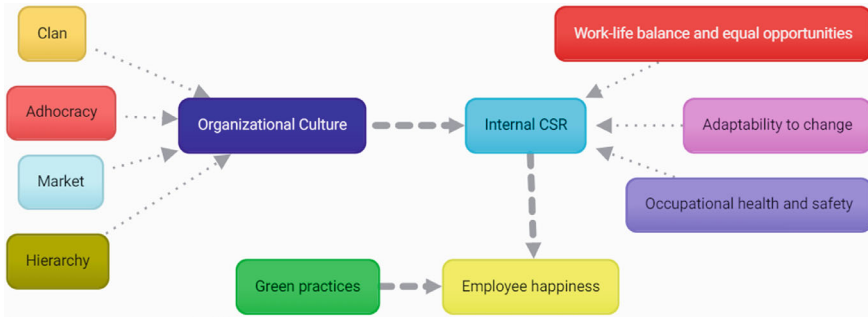


Fig. 26.1 Research model

two hypotheses develop the effect of Internal CSR (H5) and Green practices (H6) on Employee happiness.

26.4 Results

The formulated model is inspired by articles written by [33], which discusses the role of internal CSR on employee happiness, and [34] who talks about the connection between green practices and employee well-being.

Figure 26.2 illustrates the initial model comprising Clam Adhocracy, Market, Hierarchy, Internal CSR, and Green practices as formative constructs, while the Happiness variable is reflective; each of these variables consists of several items.

The first step in developing the model consisted in eliminating those items that compose formative constructs whose VIF exceeds the value five, being identified HE1, GR4, AND GR1. The next step was to run the Bootstrapping algorithm for 5000 subsamples, with a significance level of 0.05. Thus, items of formative variables whose p-value associated with outer weights exceeded the threshold of 0.05 were deleted.

After applying the first two steps detailed above, the model obtained is the one shown in Fig. 26.3. The model performances and testing of the assumptions provided in the methodology will be detailed using Tables 26.1, 26.2 and 26.3.

Table 26.1 presents that for all queries constituting formative variables, the values fall below the threshold of 5% ($p\text{-value} < 0.05$), accompanied by a T-value higher than 1.96. Additionally, for all these, the VIF is below 5, with most values below three, indicating the absence of collinearity. Concerning the reflective variable, both AVE and CR values surpass 0.7, affirming the model’s validity. Moreover, for the reflective construct, the loading values exceed 0.8.

Building upon the hypotheses formulated in the research methodology, it was determined that not all the assumptions posited in the present study have been substantiated, as evidenced by the model indicating a lack of association between adhocracy

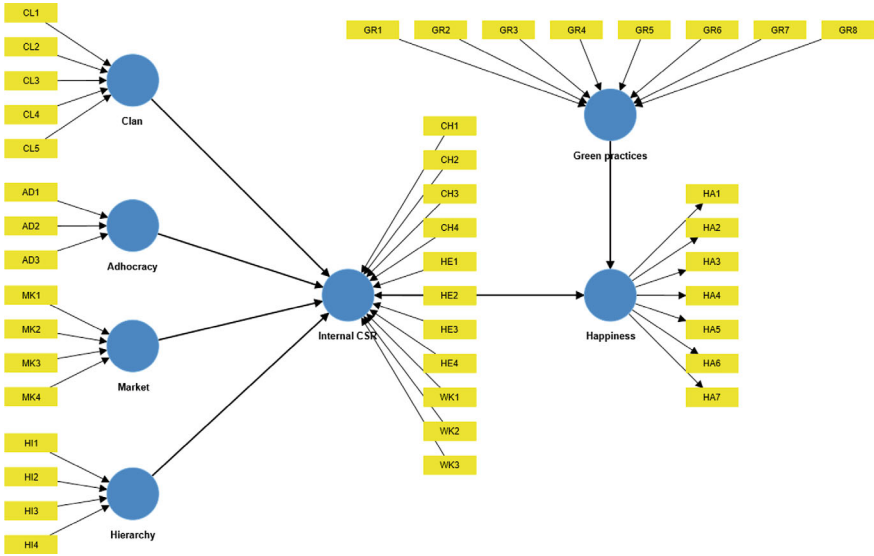


Fig. 26.2 Initial model

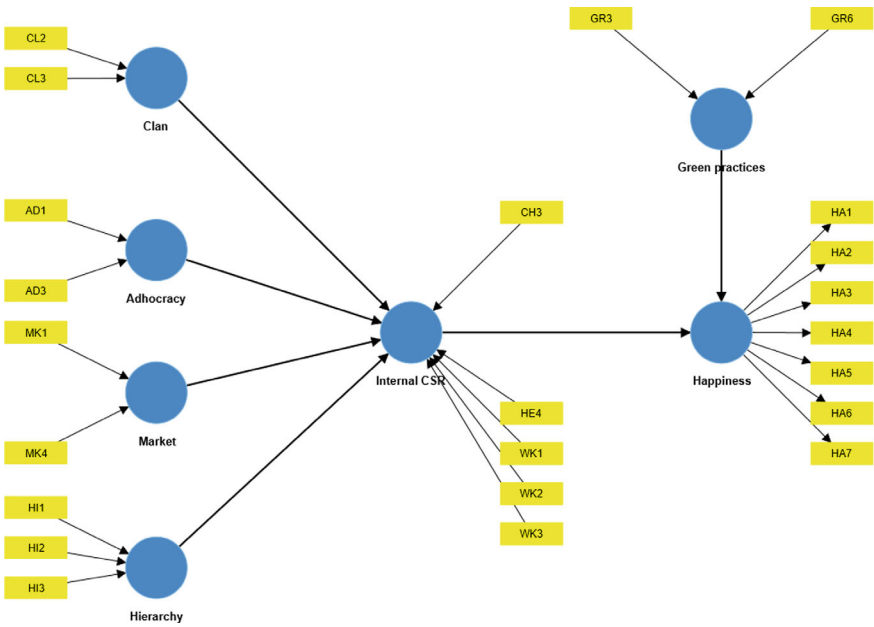


Fig. 26.3 Final model

Table 26.1 Individual components indicators

Construct	Item	Scale	Loadings/weights	AVE/T-values	CR/VIF
Clan	CL2	Formative	0.51	5.64***	2.69
	CL3		0.55	5.89***	2.69
Adhocracy	AD1	Formative	0.66	6.67***	2.12
	AD3		0.41	3.92***	2.12
Market	MK1	Formative	0.83	11.28***	1.50
	MK4		0.26	2.71**	1.50
Hierarchy	HI1	Formative	0.34	3.22**	3.17
	HI2		0.30	2.39**	3.78
	HI3		0.46	3.99***	2.48
Internal CSR	CH3	Formative	0.27	3.43**	2.16
	HE4		0.30	3.25**	2.23
	WK1		0.19	2.23**	3.42
	WK2		0.17	2.23*	2.64
	WK3		0.23	2.36*	3.13
Green practices	GR3	Formative	0.74	6.17***	2.35
	GR6		0.32	2.43*	2.35
Employee happiness	HA1	Reflective	0.86	0.82	0.96
	HA2		0.90		
	HA3		0.89		
	HA4		0.92		
	HA5		0.94		
	HA6		0.93		
	HA7		0.91		

Note * p < 0.05, ** p < 0.01, *** p < 0.001

Source Authors' own research

and internal CSR, as well as between marketing and internal CSR (refer to Table 26.2). Nonetheless, the pivotal hypothesis asserting that internal CSR and Green practices influence employee happiness has been effectively confirmed (p-value < 0.05).

Concerning the degree to which exogenous variables elucidate variations in endogenous ones, Table 26.3 indicates that internal CSR variables account for 78% of this component, whereas the variable of employee happiness explains 78%, with the remaining portion attributed to other unexamined variables.

According to what was mentioned before, most of the formulated hypotheses were demonstrated through the current model, thus, organizational culture, internal CSR, and green practices determining a significant proportion of employee happiness.

Table 26.2 Hypothesis testing

Hypothesis	Relationship	Decision
H1	Clan -> Internal CSR	Supported***
H2	Adhocracy -> Internal CSR	Not supported
H3	Market -> Internal CSR	Not supported
H4	Hierarchy -> Internal CSR	Supported***
H5	Internal CSR -> Employee happiness	Supported***
H6	Green practices -> Employee happiness	Supported***

Note * p < 0.05, ** p < 0.01, *** p < 0.001

Source Authors' own research

Table 26.3 R² values

Construct	R-square
Employee happiness	0.78
Internal CSR	0.78

Source Authors' own research

26.5 Conclusion

Sustainable practices are becoming more widespread in society, with businesses introducing green practices to reduce pollution and more. Social responsibility is another component of the introduction of sustainable work through which corporations undertake to take into account the well-being of employees. The research aimed to identify the impact brought by internal CSR and green practice on the well-being of employees, noting that 78% of the happiness component is determined based on the two variables presented previously.

As with any study, there are certain limitations, in the case of our research, the results are strictly dependent on the answers provided by the respondents. Although the number of respondents who participated in this study was not exactly small, it is possible that the results would have been more concrete if we had received a larger number of responses.

Acknowledgements The work is supported by a grant of the Romanian Ministry of Research, Innovation, and Digitalization, project CF 178/31.07.2023—‘JobKG—A Knowledge Graph of the Romanian Job Market based on Natural Language Processing’. This study was co-financed by The Bucharest University of Economic Studies during the PhD program.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Henseler, J., Leeflang, P.S.H., Wieringa, J.E., Bijmolt, T.H.A., Pauwels, K.H., (eds.): *Partial Least Squares Path Modelling*, pp. 361–381. Springer International Publishing, Cham, (2017)
2. Du, S., Bhattacharya, C.B., Sen, S.: Corporate social responsibility and competitive advantage: overcoming the trust barrier. *Manage. Sci.* **57**, 1528–1545 (2011). <https://doi.org/10.1287/mnsc.1110.1403>
3. Brammer, S., Millington, A., Rayton, B.: The contribution of corporate social responsibility to organizational commitment. *Int. J. Hum. Resour. Manag.* **18**, 1701–1719 (2007). <https://doi.org/10.1080/09585190701570866>
4. Chaudhary, R.: Corporate social responsibility and employee performance: a study among Indian business executives. *Int. J. Hum. Resour. Manag.* **31**, 2761–2784 (2020). <https://doi.org/10.1080/09585192.2018.1469159>
5. Ward, H.: Public sector roles in strengthening corporate social responsibility: taking stock. <https://documents1.worldbank.org/curated/en/548301468313740636/pdf/346560CSR1Taking1Stock.pdf> (2004). Accessed 14 April 2024
6. Mu, H.-L., Xu, J., Chen, S.: The impact of corporate social responsibility types on happiness management: a stakeholder theory perspective. *Manag. Decis.* **62**, 591–613 (2023). <https://doi.org/10.1108/MD-02-2023-0267>
7. Moore, C.: What Is eudaimonia? aristotle and eudaimonic wellbeing. <https://positivepsychology.com/eudaimonia/> Accessed on 14 April 2024
8. Tziner, A., Oren, L., Bar, Y., Kadosh, G.: Corporate social responsibility, organizational justice and job satisfaction: how do they interrelate, if at all? *J. Work Organ. Psychol.* **27**, 67–72 (2011). <https://doi.org/10.5093/tr2011v27n1a7>
9. Wisse, B., van Eijbergen, R., Rietzschel, E.F., Scheibe, S.: Catering to the needs of an aging workforce: the role of employee age in the relationship between corporate social responsibility and employee satisfaction. *J. Bus. Ethics* **147**, 875–888 (2018). <https://doi.org/10.1007/s10551-015-2983-8>
10. Ramdhan, R.M., Kisahwan, D., Winarno, A., Hermana, D.: Internal corporate social responsibility as a microfoundation of employee well-being and job performance. *Sustainability* **14**, 9065 (2022). <https://doi.org/10.3390/su14159065>
11. Carroll, N., Sadowski, A., Laila, A., Hruska, V., Nixon, M., Ma, D.W.L., Haines, J.: On behalf of the Guelph family health study the impact of COVID-19 on health behavior, stress, financial and food security among middle to high income canadian families with young children. *Nutrients* **12**, 2352 (2020). <https://doi.org/10.3390/nu12082352>
12. Zhou, T., Xu, C., Wang, C., Sha, S., Wang, Z., Zhou, Y., Zhang, X., Hu, D., Liu, Y., Tian, T., et al.: Burnout and well-being of healthcare workers in the post-pandemic period of COVID-19: a perspective from the job demands-resources model. *BMC Health Serv. Res.* **22**, 284 (2022). <https://doi.org/10.1186/s12913-022-07608-z>
13. Droppert, H., Bennett, S.: Corporate social responsibility in global health: an exploratory study of multinational pharmaceutical firms. *Glob. Health* **11**, 15 (2015). <https://doi.org/10.1186/s12992-015-0100-5>
14. Tzeng, S.-T., Su, B.-Y., Chen, H.-M.: Correlation among workplace burnout, resilience, and well-being in nursing staff: a cross-sectional study in Taiwan. *J. Nurs. Res.* **31**, e294 (2023). <https://doi.org/10.1097/jnr.0000000000000577>
15. Jaiswal, A., Arun, C.J.: Impact of happiness-enhancing activities and positive practices on employee well-being. *J. Asia Bus. Stud.* **16**, 988–1005 (2021). <https://doi.org/10.1108/JABS-06-2021-0251>
16. Yang, H., Shi, X., Bhutto, M.Y., Ertz, M.: Do corporate social responsibility and technological innovation get along? A systematic review and future research agenda. *J. Innov. Knowl.* **9**, 100462 (2024). <https://doi.org/10.1016/j.jik.2024.100462>
17. Hayat, A., Afshari, L.: CSR and employee well-being in hospitality industry: a mediation model of job satisfaction and affective commitment. *J. Hosp. Tour. Manag.* **51**, 387–396 (2022). <https://doi.org/10.1016/j.jhtm.2022.04.008>

18. Rhou, Y., Singal, M.: A review of the business case for CSR in the hospitality industry. *Int. J. Hosp. Manag.* **84**, 102330 (2020). <https://doi.org/10.1016/j.ijhm.2019.102330>
19. Gigauri, I., Panait, M., Palazzo, M.: Teaching corporate social responsibility and business ethics at economic programs. *LUMEN Proc.* **15**, 24–37 (2021). <https://doi.org/10.18662/lumproc/gekos2021/3>
20. Liu, H., Peng, F., Zeng, X., Zhao, J., Zhang, X.: Authoritarian personality and subjective well-being in chinese college students: the moderation effect of the organizational culture context. *Pers. Individ. Differ.* **138**, 79–83 (2019). <https://doi.org/10.1016/j.paid.2018.09.030>
21. Nițulescu, L.-M., Hosu, E.-A.: Corporate social responsibility for education. *Educ.* **21** 42–47 (2020). <https://doi.org/10.24193/ed21.2020.19.05>
22. Borges De Oliveira, K., De Oliveira, O.J.: Making hospitals sustainable: towards greener, fairer and more prosperous services. *Sustainability* **14**, 9730 (2022). <https://doi.org/10.3390/su14159730>
23. Khairy, H.A., Elzek, Y., Aliane, N., Agina, M.F.: Perceived environmental corporate social responsibility effect on green perceived value and green attitude in hospitality and tourism industry: the mediating role of environmental well-being. *Sustainability* **15**, 4746 (2023). <https://doi.org/10.3390/su15064746>
24. Su, L., Swanson, S.R.: Perceived corporate social responsibility's impact on the well-being and supportive green behaviors of hotel employees: the mediating role of the employee-corporate relationship. *Tour. Manage.* **72**, 437–450 (2019). <https://doi.org/10.1016/j.tourman.2019.01.009>
25. Manoj, G., Jenefa, L., Sarah, S., Velmurugan: Impact of green HRM on work-life balance of employees in automobile industry: an empirical investigation. *QAS* **23** (2022). <https://doi.org/10.47750/QAS/23.191.15>
26. Ringle, C.M., Wende, S., Becker, J.-M.: *SmartPLS 4*. Oststeinbek: smartPLS GmbH. <http://www.smartpls.com>
27. Cepeda-Carrion, G., Cegarra-Navarro, J.-G., Cillo, V.: Tips to use partial least squares structural equation modelling (PLS-SEM) in knowledge management. *JKM* **23**, 67–89 (2019). <https://doi.org/10.1108/JKM-05-2018-0322>
28. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Thiele, K.O.: Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. *J. Acad. Mark. Sci.* **45**, 616–632 (2017). <https://doi.org/10.1007/s11747-017-0517-x>
29. Hair, J., Alamer, A.: Partial least squares structural equation modeling (PLS-SEM) in second language and education research: guidelines using an applied example. *Res. Methods Appl. Linguist.* **1**, 100027 (2022). <https://doi.org/10.1016/j.rmal.2022.100027>
30. Diamantopoulos, A., Winklhofer, H.M.: Index construction with formative indicators: an alternative to scale development. *J. Mark. Res.* **38**, 269–277 (2001). <https://doi.org/10.1509/jmkr.38.2.269.18845>
31. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Danks, N.P., Ray, S.: *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. In: *classroom companion: business*, Springer International Publishing: Cham, (2021). ISBN 978–3–030–80518–0
32. Hair, J.F., Sarstedt, M.: Factors versus composites: guidelines for choosing the right structural equation modeling method. *Proj. Manag. J.* **50**, 619–624 (2019). <https://doi.org/10.1177/8756972819882132>
33. Espasandín-Bustelo, F., Ganaza-Vargas, J., Diaz-Carrion, R.: Employee happiness and corporate social responsibility: the role of organizational culture. *ER* **43**, 609–629 (2021). <https://doi.org/10.1108/ER-07-2020-0343>
34. Gyensare, M.A., Adomako, S., Amankwah-Amoah, J.: Green HRM practices, employee well-being, and sustainable work behavior: examining the moderating role of resource commitment. *Bus Strat. Env.* bse.3642 (2023). <https://doi.org/10.1002/bse.3642>

Chapter 27

Digital Skills and Labor Market Prospects: A Conjoint Experiment on Students' Preferences



Maer Matei Monica Mihaela , Mocanu Cristina , Zamfir Ana Maria ,
and Năstasă Anamaria 

Abstract This study aims to investigate students' preferences when selecting a master's program. These preferences are expressed considering the labor market prospects. The findings of this study reflect students' perceptions of the utility of relevant digital competencies. The students who participated in the experiment are studying a field that combines mathematics, economics, and information technology to understand economic systems. Upon graduation, these students should have the necessary skills to model economic behavior, utilize computational methods to solve economic problems, and make informed decisions using various quantitative methods. To investigate the students' preferences, the study uses a method called conjoint analysis. This method provides a framework to understand how people make decisions when faced with multiple options. In our study, the students are the decision-makers who are selecting their future master's program. Each student received two profiles and had to select one. This task simulates a decision-making process where respondents have to choose a preferred alternative from two options, where options vary across two or more attributes. In our experiment, the attributes are represented by a set of skills that are relevant to ICT professionals. The findings of the study emphasize the skills that enhance job prospects in this field. These skills include the ability to access and analyze digital data and to use digital tools for collaboration and productivity.

M. M. M. Mihaela (✉)
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: monica.matei@csie.ase.ro

M. Cristina · Z. A. Maria · N. Anamaria
National Scientific Research Institute for Labour and Social Protection Bucharest, Research
Institute, Bucharest, Romania

27.1 Introduction

The research undertaken within this paper subscribes to the interdisciplinary topic of understanding educational choices. The selection of education paths has implications for students’ career and personal growth. Moreover, the findings of this research could offer insights for policymakers or educational institutions in structuring more effective and relevant educational programs. According to the most used author keywords in the literature indexed in the Web of Science database (see Fig. 27.1), various factors can influence educational choices. Researchers identified socio-economic background characteristics as important influences when choosing an educational path or program. For instance, studies showed that social class and cultural factors related to parents’ occupation, income, or education impact the educational choice of young individuals [1–4]. Other studies have highlighted the effect of ascriptive factors such as gender, migratory background, ethnicity, and race [4–6]. The social network of an individual is also relevant to educational choice. Consequently, parents, siblings, and friends play key roles in students’ educational decisions [7–9].

Labor market outcomes are also relevant when choosing an educational program; studies show the increasing importance of career/occupational aspirations, skill acquisition, risk perceptions, and expected returns in terms of wages [10–13]. Available information sources also play an important role in choosing educational paths [14, 15]. Other authors also revealed that current skills and qualifications, including the current academic track and academic standings (such as standardized test scores and interests), are relevant when choosing the educational path [12, 16, 17]. At the individual level, student’s personal values, motivations, and attitudes also matter when choosing academic degrees [18, 19].



Fig. 27.1 Top author keywords by the number of occurrences in the publications regarding educational choice. Source Web of Science database

Other relevant factors when choosing an educational path or program are related to the institution and can be related to the quality of the program/institution, facilities offered, and characteristics (course program and content, rankings, type of institution, entrance criteria, etc.) [17, 20]. The Curriculum and the extent to which a specific program offers the skills students need for their chosen career [21] are also among the relevant criteria when selecting future study programs.

In our study, we investigate students' preferences when selecting a master's program. These preferences are expressed considering the labor market prospects. The findings of this study reflect students' perceptions of the utility of relevant digital competencies. The factors influencing study path selection are dependent on the field of study [3, 22, 23]. In our investigation, the students who participated in the experiment are studying a field that combines mathematics, economics, and information technology to understand economic systems. Upon graduation, these students should have the necessary skills to model economic behavior, utilize computational methods to solve economic problems, and make informed decisions using various quantitative methods.

The research methods employed to understand decision-making in terms of education paths are usually based on quantitative surveys. To investigate the students' preferences, this study uses a method called conjoint analysis. This method provides a framework to understand how people make decisions when faced with multiple options. In our study, the students are the decision-makers who are selecting their future master's program. Each student received two profiles and had to select one. This task simulates a decision-making process where respondents have to choose a preferred alternative from two options, where options vary across two or more attributes. In our experiment, the attributes are represented by a set of skills that are relevant to ICT professionals.

27.2 Methodology and Data

Understanding human behavior in the decision-making process represents a constant challenge in socio-economic research. The complexity arises from the simultaneous criteria that come into play. In the traditional approach, decision-makers are directly asked to name the most important criteria influencing their decisions. However, this method has a major drawback in that it fails to account for the potential trade-offs occurring when decisions have to be made on multiple criteria. In 1964, the psychologist Luce and statistician Tukey developed conjoint analysis. This technique considers the factors influencing a decision together and has since become an analysis tool used for preference estimation in a variety of fields.

In the investigation of complex decision processes, conjoint experiments findings are superior to those based on rank-ordering tasks or qualitative surveys because they simulate the real-world decision-making process by jointly considering the relevant factors. The main concepts conjoint analysis operates with are attribute, profile, or alternative. The attributes refer to the distinguishing features of the entity that is

being considered while making a decision, each with a set of possible states called levels. A certain combination of levels set for each attribute will reveal one profile. In conjoint analysis, a set of profiles is generated, and respondents evaluate pairs of different profiles through several tasks. The estimation is undertaken on a number of observations given by the sample size multiplied by the number of tasks per respondent. To estimate the value associated with each attribute level, the conjoint approach uses the respondents' reactions toward the evaluated profiles.

In a conjoint experiment, the decision-making behavior is observed by two means: sort or choice. In the first approach, the respondent's task is to sort a set of profiles, while the second one involves choosing the most preferred profile. It is considered that the second approach reflects a more accurate real-life decision process. Moreover, the task is less demanding given that the evaluation is performed on a reduced number of profiles at a time. In a choice-based conjoint, the respondent receives two or more profiles from which to choose. This task is repeated for multiple combinations of profiles. Each task simulates a decision-making process where respondents have to choose a preferred alternative from different options, where options vary across two or more attributes.

The specific data collection process increases the number of profiles included in the analysis compared to standard experiments since respondents in conjoint experiments evaluate multiple profiles with randomized attribute values. Additionally, this feature allows impact estimation at the individual level, which is relevant for heterogeneity investigation.

Although conjoint analysis was developed over 50 years ago, the methods used for analyzing data from a conjoint experiment have evolved over time. Initially, traditional parametric models of McFadden [24] based on logistic regression were used to estimate the importance assigned to decision criteria. However, nowadays, nonparametric approaches are frequently used, which infer the average marginal component effect (AMCE) of each factor by averaging over the distribution of other factors (Hainmueller et al. 2014). In addition, machine learning solutions are often being used [25, 26] to analyze data from conjoint experiments.

In our experiment, the profiles being evaluated are master programs. The profiles in question are defined by five skills that are relevant to the field of study being considered. These attributes were chosen using a tool called Skills OVATE. This is a dashboard developed by CEDEFOP which presents the skills that employers are looking for in Europe, extracted from online job advertisements. We have selected the most commonly requested skills (level 3 ESCO skill) for ICT professionals that are relevant to the jobs that the respondents are targeting:

- Accessing and analyzing digital data (A1);
- Managing, gathering, and storing digital data (A2);
- Using digital tools for collaboration and productivity (A3);
- Using digital tools for processing sound and images (A4);
- Protecting ICT devices (A5).

Each skill is categorized as basic or advanced, and a profile is determined by its level of development in each skill. When designing the experiment, we set the

restriction that two of the skills must be provided at an advanced level. Hence, with five attributes each having two levels, there are 2^5 possible profiles. However, due to the previously mentioned restriction, only 10 profiles are used to create the choice sets.

The R library used to build the experiment design and to define the data collection instrument is “idefix” [27]. The solutions implemented within this library allowed us to base our investigation on an adaptive efficient design. This means that we took into consideration the individual preference heterogeneity in this stage of the experiment. Individually adaptive sequential Bayesian (IASB) design reduces the vulnerability of the Bayesian approach to misspecification of the prior by adapting this distribution during the survey.

The design generated for a random respondent is shown in Table 27.1. This is represented by a matrix in which each row is a profile. The two alternatives are two master programs. In this representation, it can be seen that each respondent was given six tasks to choose from two profiles. Their answer was then recorded in the “Selected” column. For instance, for the respondent with the ID = 2, the first task asked him to choose between a program that develops advanced skills in accessing and analyzing digital data and using digital tools for collaboration and productivity, and another program that develops advanced skills in managing, gathering, and storing digital data as well as protecting ICT devices. As indicated in the “Selected” column, this respondent preferred the second option.

Finally, a shiny application was generated to gather the empirical data. In our study, the questionnaire simulates the decision-making process by giving the respondents the task of selecting from two master program profiles the one they would choose considering the labor market prospects. An illustrative choice set is presented in the Fig. 27.2.

Table 27.1 A design generated for a specific respondent

Choice sets	A1	A2	A3	A4	A5	Selected	ID
set1.alt1	1	0	1	0	0	0	2
set1.alt2	0	1	0	0	1	1	2
set2.alt1	1	0	1	0	0	1	2
set2.alt2	0	0	0	1	1	0	2
set3.alt1	1	0	1	0	0	0	2
set3.alt2	0	1	0	0	1	1	2
set4.alt1	1	0	1	0	0	1	2
set4.alt2	0	0	0	1	1	0	2
set5.alt1	1	0	1	0	0	1	2
set5.alt2	0	0	0	1	1	0	2
set6.alt1	1	0	1	0	0	0	2
set6.alt2	0	1	0	0	1	1	2

(Source Author’s own elaboration)

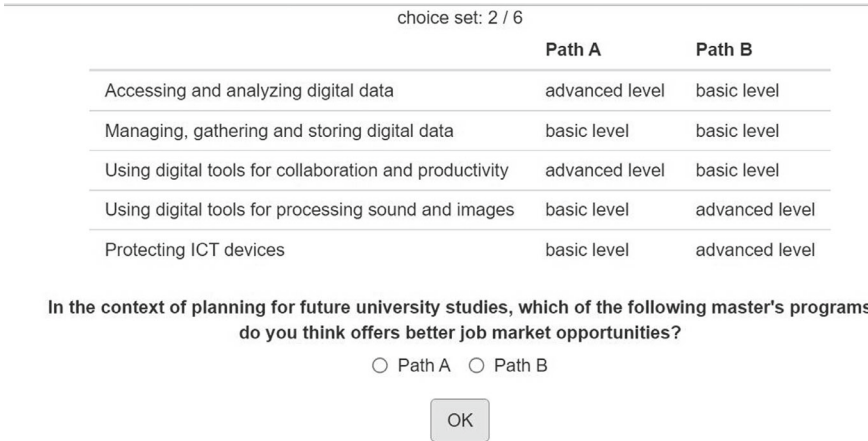


Fig. 27.2 An illustrative task from the conjoint experiment. (Source Author’s own elaboration)

The study collected preferences for educational paths from 90 s-year students specializing in economic cybernetics. The data was collected in April 2024. Due to the specificity of the conjoint design, the estimation results presented in the subsequent section were based on 1080 observations (90×12).

27.3 Results

Our findings are based on the AMCE—Average Marginal Component Effect. This estimate is calculated by observing the randomized features and respondents’ choices repeatedly. It is defined as the average effect of changing one attribute of a candidate profile on the probability that the profile will be chosen. This calculation implies performing a least squares regression and the expectation is taken over the distribution of the other attributes and respondents. The estimation and visualization are performed with library “cregg” in R [28]. The average causal effect is represented in Fig. 27.3 and it shows the effect in percentage points of changing from basic to advanced level for each skill developed through the curriculum of a certain study program on the probability of selection.

The estimation results show that for the field of economic cybernetics, student choices are positively influenced by two attributes: A1-accessing and analyzing digital data and A3-Using digital tools for collaboration and productivity. Therefore, changing a master program’s curriculum from basic to advanced data analysis skills increases enrollment probability by 2 percentage points, holding all else equal.

These results are in accordance with the specifics of this study program since the general objectives are to train graduates who have the capacity to use quantitative tools to model economic systems. However, our study’s limitations result from using

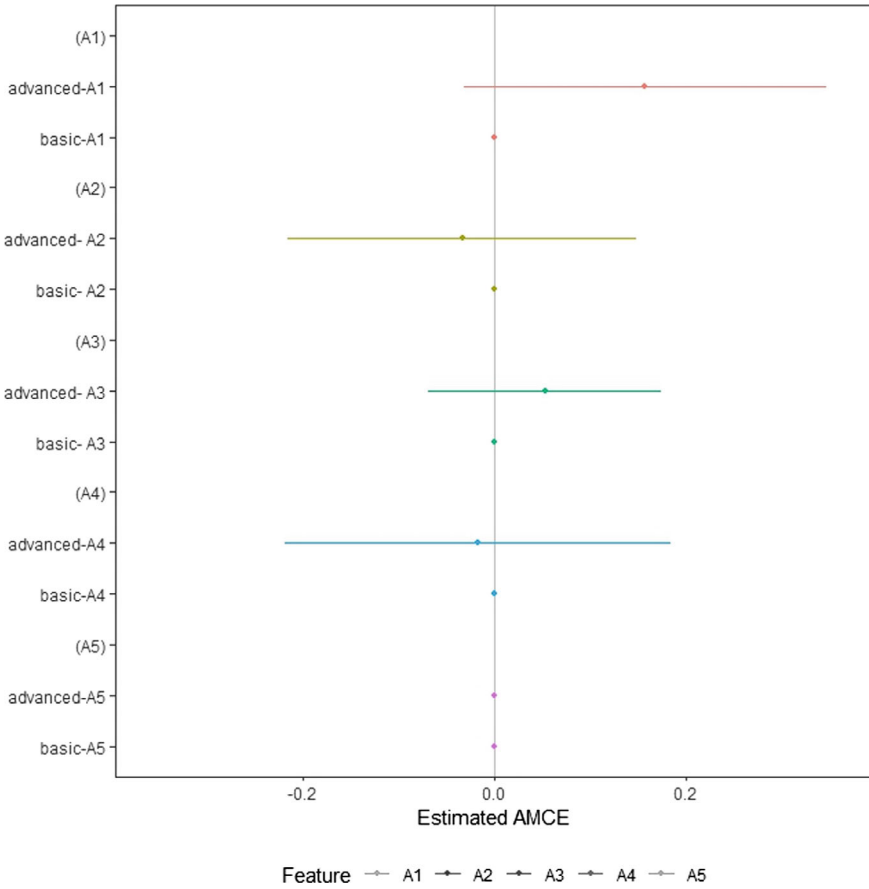


Fig. 27.3 Estimation results. (Source Author’s own estimation)

a small sample of respondents, as only about 50% of the students in the target group have expressed their preferences. Consequently, this study is a pilot experiment that provided us with valuable feedback. The results indicate the general competencies that our graduates find useful for their careers. In the future, we plan to conduct a more detailed study, which will include specific skills and a larger sample size. We will also develop the questionnaire to investigate the heterogeneity of the preferences of the respondents’ characteristics.

27.4 Conclusions

This study builds a conjoint experiment to understand the effect of different educational paths characteristics on the probability that a student would select that path. The experiment was designed using an Individually adaptive sequential Bayesian (IASB) approach. The educational path characteristics included in the design are five of the most requested level 3 ESCO skills for ICT professionals. The study extracted data from a causal estimand indicating how changing a program's specific skill level from basic to advanced affects the selection probability, while keeping other factors constant. According to our estimations, the students attending the field of economic cybernetics give a higher weight to skills like accessing and analyzing digital data or using digital tools for collaboration and productivity when selecting their master's program.

The experiment designed within our study provides a robust framework for understanding student preferences in choosing educational programs. Therefore, the findings drawn from it guide educational institutions to make informed decisions aligning with students' needs and market demands. The value added by using this approach within this context resides in the fact it simulates real-life decision-making scenarios where students must weigh various factors against each other. Hence the estimation results illustrate the trade-offs and how respondents prioritize different aspects.

References

1. Vergolini, L., Vlach, E.: Family background and educational path of Italian graduates. *Higher Edu* **73**, (2017)
2. Attanasio, O., Kaufmann, K.: Educational choices, subjective expectations, and credit constraints. *Natl Bur Econ Res Work Pap No.* 15087, (2009)
3. Sianou-Kyrgiou, E., Tsiplakides, I.: Choice and social class of medical school students in Greece. *Br. J. Sociol. Educ.* **30**(6), 727–740 (2009)
4. Giancola, O., Colarusso, S.: Gender, social origins, and educational choices: How it really works. In: *The Education of gender. The gender of education. sociological research in Italy*, pp. 95–109. Associazione "Per Scuola Democratica", (2020)
5. Haas, C., Hadjar, A: Students' trajectories through higher education: A review of quantitative research. *Higher. Edu.* **79**(6), (2020)
6. Lucas, S.R., Beresford, L.: Naming and classifying: theory, evidence, and equity in education. *Rev. Res. Educ.* **34**(1), 25–84 (2010)
7. Smith, E.: Peer preferences and educational decisions: Heterogeneous associations across student socioeconomic status. *Br. J. Sociol. Educ.* **44**(2), 374–393 (2023)
8. van der Vleuten, M., Weesie, J., Maas, L.: Sibling influence in field of study choices. *Res. Soc. Stratif. Mobil.* **68**, 100525 (2020)
9. Andersen, S.C., Hjortskov, M.: The unnoticed influence of peers on educational preferences. *Behav. Public Policy* **6**(4), 530–553 (2022)
10. Koksharov, V.A., Agarkov, G.A.: Analysis of economic motivation when individuals choose an educational path. *Ekon. Regiona= Econ. Reg.* **1**, 245 (2015)
11. Pinxten, M., De Fraine, B., Van Den Noortgate, W., Van Damme, J., Anumendem, D.: Educational choice in secondary school in Flanders: The relative impact of occupational interests on option choice. *Educ. Res. Eval.* **18**(6), 541–569 (2012)

12. Rapoport, B., Thibout, C.: Why do boys and girls make different educational choices? the influence of expected earnings and test scores. *Econ. Educ. Rev.* **62**, 205–229 (2018)
13. Brodaty, T., Gary-Bobo, R.J., Prieto, A.: Do risk aversion and wages explain educational choices? *J. Public Econ.* **117**, 125–148 (2014)
14. Dirin, A., Nieminen, M., Alamäki, A.: Social media and social bonding in students' decision-making regarding their study path. *Int. J. Inf. Commun. Technol. Educ. (IJICTE)* **17**(1), (2021)
15. Bonnema, J., Van der Waldt, D.L.R.: Information and Source Preferences of a Student Market in Higher Education. *Int. J. Educ. Manag.* **22**(4), 314–327 (2008)
16. Strømme, T.B.: Vocational and academic decisions in 'classed' school environments. *J. Educ. Work.* **33**(3), 197–211 (2020)
17. Hemsley-Brown, J., Oplotka, I.: University choice: What do we know, what don't we know and what do we still need to find out? *Int. J. Educ. Manag.* **29**(3), 254–274 (2015)
18. Kopanidis, F.Z., Shaw, M.J.: Courses and careers: Measuring how students' personal values matter. *Educ. Train.* **56**(5), 397–413 (2014)
19. Jensen, F., Henriksen, E.K.: Short stories of educational choice: In the words of science and technology students. In: *Understanding student participation and choice in science and technology education*, pp. 135–151. Springer, Dordrecht (2014)
20. Bernardi, F., Cebolla, H.: Social class and school performance as predictors of educational paths in Spain. *Rev. Esp. Investig. Sociol.* **146**(1), (2014)
21. Määttä, K., Uusiautti, S.: How to enhance the smoothness of university students' study paths. *Int. J. Res. Stud. Educ.* **1**(1), (2012)
22. Spigarska, E., Majerowska, E.: Economic and social factors of studying finance and accounting: The case of the master degree students at The University of Gdansk. In: *EDULEARN17 proceedings*, pp. 553–559. IATED, (2017)
23. Sholihah, I.A., Prasetya, H., Widyaningsih, V.: Path analysis on factors affecting the decision to choose school of acupuncture. *J. Health Policy Manag.* **5**(2), (2020)
24. McFadden, D.: Conditional logit analysis of qualitative choice behavior. In: *Frontiers in econometrics*, pp. 105–142. Academic Press, (1974)
25. Ham, D.W., Imai, K., Janson, L.: Using machine learning to test causal hypotheses in conjoint analysis. *Political Anal.* (2022)
26. Robinson, T.S., Duch, R.M.: How to detect heterogeneity in conjoint experiments. *J. Polit.* **86**(2), (2024).
27. Traets, F., Sanchez, D.G., Vandebroek, M.: Generating optimal designs for discrete choice experiments in R: The idifix package. *J. Stat. Softw.*, **96**(3), 1–41, (2020). <https://doi.org/10.18637/jss.v096.i03>
28. Leeper, T.J.: Cregg: Simple conjoint analyses and visualization. R package version 0.3.6, (2020)

Chapter 28

Estimating Probability of Default. A New Qualitative Perspective



Florin George Dragu

Abstract In today's economic landscape, accurate credit risk assessment is paramount for prudent lending decisions. This paper proposes a qualitative framework for modeling default probability (PD) in credit risk assessment. Unlike traditional quantitative models, this approach integrates a broader range of factors including market position, ownership structure, and sector risk. By complementing quantitative techniques with qualitative insights, this model enhances PD assessments and can be utilized in stress testing and macroprudential analyses. The methodology involves systematically evaluating qualitative dimensions contributing to credit risk and employs a logit model with backward selection. Drawing from a comprehensive dataset, including approximately 30 qualitative characteristics from financial statements, Credit Registry loan-level information, and Business Register data, this approach offers a robust framework for assessing default probability.

28.1 Introduction and Literature Review

The model outlined in this paper is constructed as a complementary tool to enhance the quantitative estimation of credit risk associated with corporate exposures. While traditional assessments rely mostly on financial indicators such as profitability, liquidity, and indebtedness positions, this model adopts a broader perspective. It places greater emphasis on qualitative variables, including market position, shareholder support, activity sector, and sensitivity to economic cycles.

By incorporating these qualitative dimensions, the model aims to provide a more comprehensive understanding of a firm's creditworthiness. For instance, assessing market positioning allows for insights into competitive advantages and market share dynamics, while considering shareholder support offers valuable information about governance structures and stakeholder confidence. Additionally, analyzing the firm's

F. G. Dragu (✉)
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: florin.dragu@gmail.com

activity sector and its dependency on economic cycles enables a deeper assessment of industry-specific risks and macroeconomic influences.

Credit risk modeling is crucial for financial institutions to assess the probability of borrower default. While quantitative models have traditionally dominated credit risk management, qualitative factors need to be taken as well into account to enhance the accuracy of the estimation. Even though, generally, the qualitative assessment is performed by an analyst, this study proposes an automated model to streamline and enhance the qualitative evaluation process.

This paper builds on the existing literature on credit risk modeling. The current approaches for qualitative assessments can be, generally, grouped into two main categories: (i) expert-based models and (ii) qualitative scoring systems.

In situations when historical data is limited or unreliable, expert judgment models can play a significant role in credit risk assessment. Soares et al. [9] conducted a survey of experts from major Portuguese banks, revealing the significance of qualitative criteria, notably management experience and reliability. Clemen and Winkler [4] show that experts' judgments are crucial due to limited "hard data" availability, making their input valuable in probabilistic risk analysis. García et al. [6] introduce an approach utilizing goal programming to integrate expert judgment into a solvency model. Furthermore, expert judgment models offer flexibility in addressing evolving risk factors and market dynamics. In rapidly changing environments, where historical data may not reflect current conditions, expert opinions provide valuable insights into emerging risks and vulnerabilities.

Qualitative scoring systems are another important component in credit risk assessment, offering a systematic approach to assess borrowers' creditworthiness. Altman's Z-score model [1] is a landmark in quantitative assessment, using financial ratios to categorize firms into risk groups. Building upon Altman's work, subsequent research has expanded the model to integrate additional qualitative elements. For instance, Chava and Jarrow [3] extended the Z-score model to account for industry-specific dynamics, acknowledging the significance of contextual factors in credit risk evaluation.

Moreover, qualitative scoring systems enable lenders to integrate non-financial information into the credit risk assessment process. Factors such as management quality, corporate governance practices, and market reputation are increasingly recognized as critical determinants of creditworthiness [7]. In addition, Bozanic et al. [2] find that firms' qualitative disclosure significantly influences credit analysts' soft adjustments, particularly regarding debt-related topics.

The clarity and simplicity of qualitative scoring systems make them attractive to lenders and investors, providing straightforward criteria for assessing credit risk. These models streamline decision-making and risk management processes. Moreover, they can be tailored to meet the unique requirements and preferences of various stakeholders, enhancing flexibility in risk assessment methods.

Despite their strengths, qualitative scoring systems are not without limitations. One challenge lies in the subjective nature of qualitative assessments, which may introduce biases or inconsistencies in credit risk evaluations. The evolving nature

of industries and markets necessitates regular updates and refinements to scoring systems to maintain their relevance and effectiveness.

The qualitative module proposed in this paper can be strategically integrated to automatically offer nuanced insights into a firm's capacity to service its bank debt obligations. It seeks to capture qualitative influences that may not be fully reflected by quantitative metrics alone, thereby enhancing the overall accuracy and reliability of the credit risk assessment process. This approach aims to provide a more robust framework for evaluating credit risk and supporting informed decision-making in credit risk management.

28.2 Data and Methodology

The dataset used to quantify the probability of default model is constructed based on all Romanian firms' financial statements and loan-level information from the Credit Registry, covering the period 2015–2022. The frequency of the data is annual.

The default flag is formulated based on a subset of the Capital Requirements Regulation definition for defaulted exposures, namely, “a default shall be considered to have occurred with regard to a particular obligor when either or both of the following have taken place:

- the institution considers that the obligor is unlikely to pay its credit obligations to the institution, the parent undertaking or any of its subsidiaries in full, without recourse by the institution to actions such as realizing security;
- the obligor is more than 90 days past due on any material credit obligation to the institution, the parent undertaking or any of its subsidiaries.”

The dependent variable is a binary indicator that takes the value 1 if a firm meets the second criterion outlined above for default and 0, otherwise (Eq. 28.1).

$$y_i = \begin{cases} 1, & \text{default} \\ 0, & \text{otherwise} \end{cases} \quad (28.1)$$

The model is estimated using a logistic regression, which employs as explanatory variables two types of indicators: (i) grades-based variables and (ii) dummy variables. These provide qualitative information on firms' non-financial characteristics, such as their track record in the non-financial companies' sector, their ability to cope within the market they operate in or their dependency on the business cycle.

$$P(y_i = 1) = F(x_i\beta) = \frac{1}{1 + e^{-x_i\beta}} \quad (28.2)$$

Data filtering and validation are performed to ensure high-quality data for the probability of the default model. The explanatory variables in the sample are winsorized to exclude extreme values. To develop the final default model, additional filters, such

as the Kolmogorov–Smirnov (KS) test and discriminatory power tests, are applied to the pool of candidate explanatory variables and intermediate default models.

The grades-based variable aims to categorize non-financial companies into distinct groups, taking into account various qualitative factors commonly evaluated in credit risk models. Additionally, it enables more nuanced decision-making by identifying specific risk profiles, thereby enhancing the precision of the model.

A first significant indicator is the firm's industry positioning within a specific 2-digit NACE group, calculated as the proportion of the company's sales relative to the total sales of the NACE group. The grades range from 1 to 5, reflecting a spectrum from weakest to strongest, based on increasing turnover shares. Companies reporting zero sales are assigned a grade of 1, while those surpassing the 75th percentile are rated as grade 5, providing a comprehensive assessment of market penetration and competitive positioning.

Another variable under consideration is competition intensity, measured by the Herfindahl–Hirschman Index (HHI) within 3-digit NACE groups. The HHI categorizes activity sectors into three tiers based on US Department of Justice competition thresholds (U.S. Department of Justice & FTC, Merger Guidelines § 2.1 [10]: (i) unconcentrated markets (HHI below 1000), (ii) moderately concentrated markets (HHI between 1000 and 1800), and (iii) highly concentrated markets (HHI above 1800). Subsequently, companies are graded from 1 to 3 according to their sector classification: 1 indicating high competition, 2 reflecting moderate competition, and 3 representing low competition.

Another important factor influencing credit risk is considered to be the track record of firms in the corporate sector. Companies are segmented into four cohorts based on their age, with a higher age corresponding to a better grade. This is rooted in the belief that mature companies are better positioned to capitalize on business opportunities and maintain superior financial stability. The categorization is defined relative to the median age of Romanian companies (6 years). The first two groups comprise (i) companies aged below 3 years (start-ups) and (ii) firms with a track record of 3 to 6 years. The top tiers consist of firms aged (i) between 6 and 10 years or (ii) over 10 years.

An additional qualitative aspect explored is the level of support from companies' shareholders. Initially, firms with debt owed to shareholders or affiliated entities are identified. Subsequently, each firm is graded on a scale of 1 to 4, based on the proportion of debt owed to shareholders in relation to total debt. A higher grade is assigned to companies with a larger share of debt to shareholders, with grades ranging from 1 for those in the first quartile to 4 for those in the fourth quartile.

Utilizing both historical non-performing loan (NPL) data and sector risk assessments from the National Bank of Romania's (NBR) Bank Lending Survey, companies operating within specific sectors are graded accordingly. This process involves assigning grades based on sector riskiness, with lower-risk sectors receiving higher grades for companies operating within them.

A final variable within the first category pertains to companies' proportion of arrears. Firms with minimal past-due debts to suppliers and other creditors are allocated the highest grade and vice-versa. This reflects a strong financial position and

timely payment practices, indicative of lower credit risk. This emphasis on arrears underscores the significance of prompt payment behavior in mitigating credit risk.

The second category of indicators incorporated into the model development comprises a series of binary variables, denoted by a value of 1 when a specific condition is met and 0 otherwise. These dummy variables include:

Companies exhibiting higher growth rates compared to the entire non-financial corporation (NFC) sector: A positive link exists between credit risk and companies demonstrating higher growth rates compared to the entire non-financial corporation (NFC) sector. Rapid growth often signals strong market demand, innovation, and competitive advantages, making these companies attractive investment prospects. Effective execution of strategic initiatives and adaptability to industry trends contribute to long-term viability and creditworthiness.

Firms displaying greater sensitivity to business cycle fluctuations, characterized by increased growth during economic expansions and sharper declines during contractions: Higher sensitivity to business cycle fluctuations can increase firms' vulnerabilities during economic downturns. While these companies may thrive during expansions, fueled by increased demand and favorable conditions, they often struggle during contractions. Reduced demand, tighter credit, and market volatility can impede their ability to sustain growth and meet financial obligations, heightening the risk of default and financial distress. Consequently, despite short-term gains, their susceptibility to economic cycles exposes them to heightened credit risk.

Degree of sector development, identifying companies engaged in technology- and knowledge-intensive activities: Sectors characterized by advanced technology and knowledge-intensive activities often exhibit greater resilience and growth potential, reducing credit risk. Conversely, sectors with lower levels of development may face challenges such as market saturation, technological obsolescence, and heightened competition, elevating credit risk for companies operating within them.

Companies exposed to foreign exchange (FX) risk, such as those with foreign-currency-denominated loans or external debt. The negative correlation between credit risk and firms exposed to foreign exchange (FX) risk, particularly those holding foreign-currency-denominated loans or external debt, stems from the susceptibility to market volatility, which in turn can amplify the debt burden and escalate debt servicing expenses, hedging costs, macroeconomic influences, and the ensuing downward pressure on credit ratings attributed to exchange rate fluctuations. These factors collectively strain cash flows, diminish profitability, and prompt creditor apprehension regarding a firm's stability and ability to fulfill debt obligations.

Identification of atypical situations, such as negative equity, absence of employees, or lack of sales. In atypical situations like negative equity, absence of employees, or lack of sales, the connection between bank credit risk becomes evident, highlighting the need for careful risk assessment. Negative equity signifies financial distress and limited collateral availability, elevating the risk of default. Likewise, the absence of employees or lack of sales may signal operational challenges or an unsustainable business model, raising doubts about the borrower's capacity to generate revenue and repay debt.

Companies with records of payment incidents, reflecting instances of delayed or defaulted payments to their business partners. Firms that register payment incidents pose significant credit risk for banks, as past delinquencies signal financial instability and increase the likelihood of default.

These dummy variables provide additional granularity to the model, capturing various dimensions of firms' operational characteristics and risk exposures. In total, the qualitative module relies on approximately 30 indicators.

The modeling process consists of several stages, constructed to ensure the validity of the results.

The first stage consists of a logit model estimation on the entire timespan 2015–2022, having a through-the-cycle approach. The initial list of variables is determined using a stepwise procedure with a 99% significance level. Finally, several logit models are being run to check the robustness and stability of the overall results in various configurations. This through-the-cycle model relies on approximately 580 thousand observations, of which 14.7 thousand are defaults.

Secondly, a point-in-time estimation is conducted to cross-check the robustness of the results. The estimation follows an approach similar to the one proposed by Costeiu and Neagu [5]. The estimation is carried out for each year in the sample, using one year of data for the development sample, on which the model is being trained. Then, the model is re-estimated on the validation sample. The logit methodology employed uses a backward selection method and is conducted in several steps. Initial models (100) are estimated on random samples of 50:50 defaulting to non-defaulting companies on the development sample (indicators at year $t-1$). Then, variables are filtered out based on minimum significance and likelihood ratio tests. Secondly, the final model is selected based on the frequency of explanatory variables and model specification. Thirdly, the model is re-estimated with selected explanatory variables on random samples of 50:50 defaulting to non-defaulting companies with indicators at year t (out-of-sample estimation). The model parameters are derived as the mean of coefficients' distribution. Finally, to account for the true default rate, the intercept within the logit formula is adjusted using the methodology proposed by King and Zeng [8] that involves shifting the distribution of the PDs.

Rating scales provide a systematic method for evaluating the creditworthiness of individuals or entities by assigning default probabilities to various risk categories. The rating scale used in this paper to evaluate the riskiness of non-financial companies is based on the one developed by Standard and Poor's, with several adjustments, proposing a higher granularity in the lower part of the scale. These adjustments account for the tendency of Romanian companies to cluster in the lower part of the rating scale and aim to ensure adequate granularity in the rating distribution by subdividing several of the lower rating classes.

28.3 Results

The qualitative Probability of Default (PD) model customized for Romania's corporate sector, as depicted in Table 28.1, is developed following the methodological framework outlined in Sect. 28.2. Data utilized for constructing the PD model is procured from diverse sources. The dataset employed in model development comprises around 580,000 observations spanning the timeframe from 2015 to 2022. Companies exhibiting anomalies in their financial reports, such as negative turnover or total assets, are meticulously excluded from the analysis to ensure data integrity and model robustness.

The through-the-cycle Probability of Default (PD) logit model estimation unveils significant insights into various factors influencing firms' probability of default within the Romanian non-financial companies' sector.

The statistical significance evident in the overall results presented in Table 28.1 underscores the reliability of the model estimation. Across the various variables analyzed, the coefficient values consistently demonstrate statistical significance. Furthermore, all variables have the expected sign and are in line with the economic theory.

In more details, the coefficient value of approximately -0.36 shows a negative link between a company's market position and its probability of default. This means that firms with a stronger foothold in their markets tend to demonstrate a lower likelihood of default. This trend persists consistently across all estimations and maintains statistical significance. This is due to the fact that, in general, companies with a robust market presence often benefit from more diversified revenue streams, expansive customer bases, and heightened brand recognition. These factors collectively bolster their resilience during economic downturns and mitigate the risk of default.

Conversely, the presence of payment incidents is associated with a substantial increase in the probability of default, with a positive coefficient value exceeding 2.5. This indicates that companies facing payment issues are at a significantly higher risk of default.

Similarly, companies encountering atypical financial situations exhibit a markedly higher probability of default, with positive coefficient values around 0.85. This variable consistently retains statistical significance across all estimations. Furthermore, companies with higher arrears tend to demonstrate a higher probability of default.

The presence of foreign exchange (FX) loans also contributes to an increased probability of default, with positive coefficient values ranging from approximately 0.25 to 0.28, a finding that remains statistically significant across all estimations.

Moreover, companies whose Gross Value Added (GVA) cycles synchronize with the broader economic cycle demonstrate a lower probability of default, as indicated by coefficient values ranging between approximately -0.36 and -0.51 . This variable maintains as well statistical significance across all estimations.

Table 28.1 Through-the-cycle logit model estimations

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Company's market position	-0.3584*** (0.00)	-0.3757*** (0.00)	-0.3662*** (0.00)	-0.3627*** (0.00)	-0.3670*** (0.00)	-0.3684*** (0.00)
Payment incidents (dummy)	2.5605*** (0.00)	2.5640*** (0.00)	2.5601*** (0.00)	2.5632*** (0.00)	2.5614*** (0.00)	2.5554*** (0.00)
GVA cycle synchronization (dummy)	-0.3610*** (0.00)	-0.5144*** (0.00)	-0.5009*** (0.00)	-0.4960*** (0.00)	-0.4805*** (0.00)	-0.5020*** (0.00)
Atypical financial situations (dummy)	0.8540*** (0.00)	0.8637*** (0.00)	0.8631*** (0.00)	0.8534*** (0.00)	0.8549*** (0.00)	0.9130*** (0.00)
FX loans (dummy)	0.2589*** (0.00)	0.2506*** (0.00)	0.2832*** (0.00)	0.2827*** (0.00)	0.2665*** (0.00)	0.2516*** (0.00)
Arrears (grades)	-0.3715*** (0.00)	-0.3685*** (0.00)	-0.3724*** (0.00)	-0.3729*** (0.00)	-0.3758*** (0.00)	-0.3591*** (0.00)
Sales cycle synchronization (dummy)	-0.2510*** (0.00)					
Employee cycle dependency (dummy)		0.2641*** (0.00)				
Private-owned companies (dummy)			-0.1932*** (0.00)			
Export activity (dummy)				-0.2390*** (0.00)		
Company's track record (grades)					-0.0356*** (0.00)	

(continued)

Table 28.1 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Shareholder support (grades)						-0.1174***
						(0.00)
Constant	-1.2153***	-1.2155***	-1.2068***	-1.2127***	-1.0827***	-1.0327***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	580,484	580,484	580,484	580,484	580,484	580,484
Pseudo R2	15.53%	15.53%	15.48%	15.49%	15.48%	15.58%
ROC	81.16%	81.02%	81.04%	81.07%	81.11%	81.10%

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$; p -values in parentheses

Source Author's estimations

Additionally, the model is estimated in several different specifications, which include other variables such as sales cycle synchronization, employee cycle dependency, private-owned companies, export activity, company's track record, and shareholder support, in order to check the robustness of the estimations. In essence, the models' performance remains consistent regardless of the model specification.

The model's performance is assessed through pseudo-R-squared and Receiver Operating Characteristic (ROC) metrics. The pseudo-R-squared values are ranging from 15.48 to 15.58%. Moreover, the ROC remains relatively stable at 81% across the models, indicating a sufficient level of predictive accuracy in distinguishing between default and non-default cases.

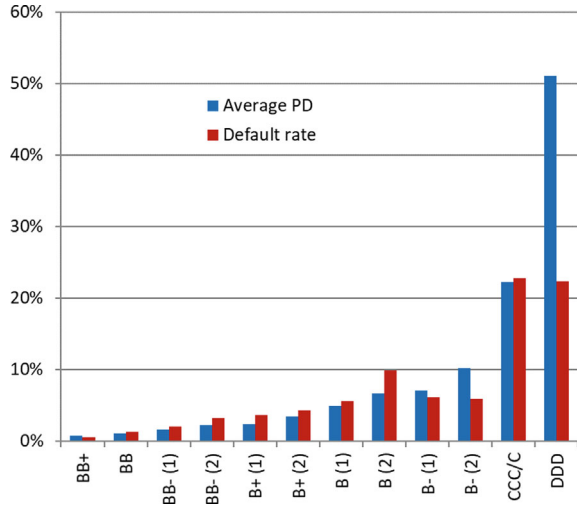
The through-the-cycle PD models are cross-referenced with a point-in-time version to validate their reliability across different temporal economic contexts. The results exhibit general consistency between the outcomes generated by both models, affirming the model's stability and accuracy across diverse timeframes.

28.4 Conclusion

The current study utilizes a range of logit estimations with varied specifications to evaluate the likelihood of default among non-financial companies in Romania, providing additional valuable insights into the determinants of default probability.

The analysis reveals several key determinants of default within the Romanian corporate sector. Notably, an inverse relationship between a company's market position and its likelihood of default, highlighting the resilience of firms with stronger market footholds. Conversely, the presence of payment incidents and atypical financial situations significantly heightens the risk of default. The robustness of the model's estimations is confirmed through various specifications, with consistent performance assessed by pseudo-R-squared and ROC metrics. Moreover, cross-referencing the through-the-cycle PD models with a point-in-time version affirms their reliability

Fig. 28.1 (Average PD versus realized default rate)



across different economic contexts, further enhancing confidence in their predictive capabilities.

In conclusion, the qualitative module presented in this paper should be combined with a model based on financial indicators to offer nuanced insights into a firm’s ability to meet its bank debt obligations. Capturing qualitative factors often overlooked by quantitative metrics alone enhances the accuracy and reliability of credit risk assessment. This approach establishes a more robust framework for evaluating credit risk and facilitates informed decision-making in credit risk management.

This tool has multiple applications. Firstly, it can serve as an input for stress-testing frameworks that assess the impact of various scenarios on the probability of default for non-financial companies. Secondly, it can be integrated into macroprudential analyses. Lastly, and importantly, it can be employed in supervisory activities to benchmark Internal Ratings-Based (IRB) models submitted by banks for approval.

There are several limitations of the qualitative PD model proposed in this paper. Firstly, the default indicator currently relies on a subset of the CRR default definition, which may not capture all relevant default scenarios. Additionally, the model’s reliance on the pandemic crisis as the sole instance of economic contraction might introduce some bias, as this period’s unique circumstances might lead to an overestimation of the probability of default, as we can observe from the results during that period.

28.5 Annex

See Figs. 28.1, 28.2, 28.3, 28.4 and 28.5.

Fig. 28.2 Calibration quality (Binomial test)

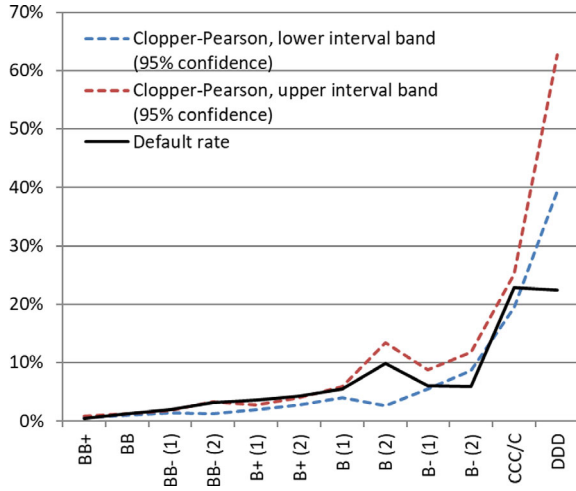


Fig. 28.3 Discriminatory power

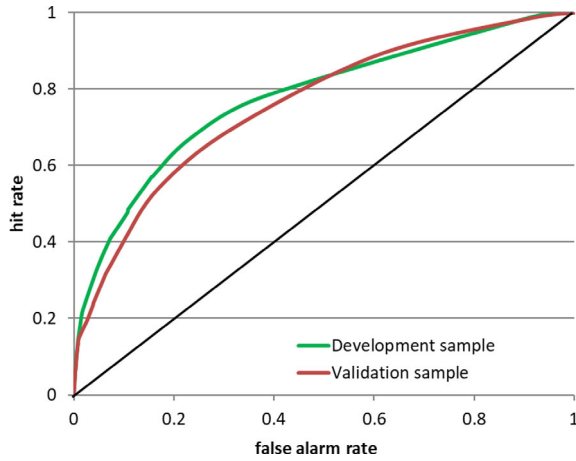
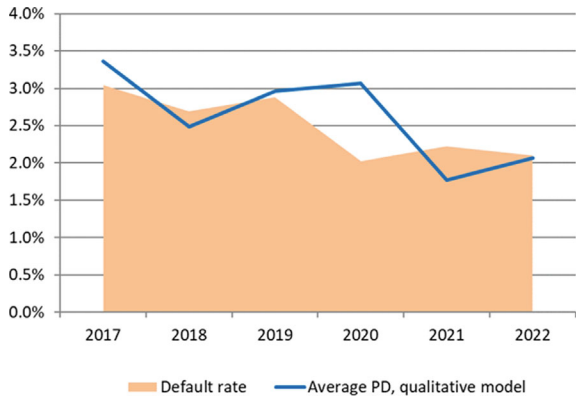


Fig. 28.4 Estimated PD versus realized default rate



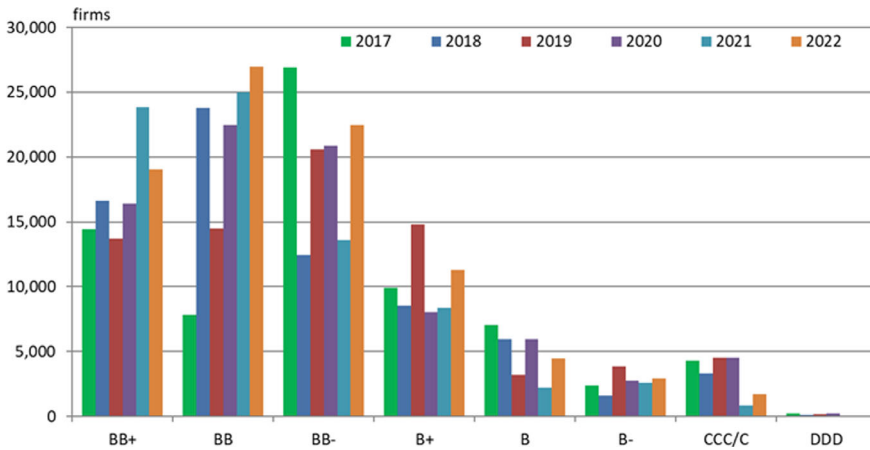


Fig. 28.5 Rating distribution 2017–2022

References

1. Altman, E.I.: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Financ.* **23**(4), 589–609 (1968)
2. Bozanic, Z., Li, X., Pincus, M.: Soft information in hard adjustments: evidence from credit rating agencies' reports. *J. Account. Res.* **60**(4), 1219–1259 (2022)
3. Chava, S., Jarrow, R.: Bankruptcy prediction with industry effects. *Rev. Financ.* **8**(4), 537–569 (2004)
4. Clemen, R.T., Winkler, R.L.: Combining probability distributions from experts in risk analysis. *Risk Anal: Int J* **19**(2), 187–203 (1999)
5. Costeiu, A., Neagu, F.: Bridging the banking sector with the real economy. A financial stability perspective. ECB Working Paper Series no. 1592, (2013)
6. García, J., Fernández, E., Salinas, A.: Integrating expert judgment into the solvency assessment of insurance companies. *Insur.: Math. Econ.*, **53**(3), 654–665 (2013)
7. Kaur, M., Vij, A.: The role of qualitative factors in assessing credit risk. *J. Bank. Finance* **131**, 106469 (2023)
8. King, G., Zeng, L.: Logistic regression in rare events data. Harvard University, Center for Basic Research in the Social Sciences (2001)
9. Soares, N., Nunes, L.C., Serrasqueiro, Z.: Credit risk assessment of small firms: Evidence from a survey of Portuguese banks. *Small Bus. Econ.* **37**(4), 477–492 (2011)
10. U.S. Department of Justice & FTC, Merger Guidelines § 2.1 (2023). <https://www.justice.gov/atr/herfindahl-hirschman-index>

Chapter 29

Brand Associations Bibliometric Analysis



Alexandra Raluca Jelea , Adriana Manolică , Teodora Roman ,
and Patricea Elena Bertea 

Abstract Connectivity in modern society has revolutionized communication and interaction, transcending geographical barriers and fostering unprecedented exchange of information and ideas. Enabled by rapid technological advancement, connectivity has democratized access to knowledge, transformed various sectors including education, healthcare, and business, and reshaped social interactions. Moreover, connectivity has played a crucial role in reshaping the global economy, facilitating seamless trade, and creating new opportunities for businesses and entrepreneurs. In the context of brand associations, a bibliometric analysis was conducted to provide insights into the existing literature. Results revealed a growing interest in the topic, with a significant number of publications primarily focused on quantitative studies within the business domain. However, there remains a notable gap in qualitative studies exploring brand associations, indicating opportunities for further research and innovation. Overall, the analysis underscores the importance of connectivity in driving growth, innovation, and deeper connections with consumers in the sphere of brand associations.

29.1 From Connectivity to Brand Associations

In modern society, connectivity has become a pillar that revolutionized the way we communicate, collaborate, and interact with the world around us [1]. In this era of rapid technological advancement, connectivity serves as the bridge that transcends geographical barriers, connecting individuals, communities, and nations in ways never thought possible.

The advent of the digital revolution has unleashed unprecedented connectivity, facilitating the exchange of information and ideas across vast distances. The development of smartphones, high-speed Internet, and social media platforms has democratized access to knowledge, enabling individuals from diverse backgrounds to connect,

A. R. Jelea (✉) · A. Manolică · T. Roman · P. E. Bertea
Alexandru Ioan Cuza University of Iasi, Iasi, Romania
e-mail: alexandra.jelea@uaic.ro

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
C. Ciurea et al. (eds.), *Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)*, Smart Innovation, Systems and Technologies 426,
https://doi.org/10.1007/978-981-96-0161-5_29

333

learn, and share experiences [2]. Connectivity has transformed education, healthcare, business, and entertainment, creating opportunities for economic growth, innovation, and social progress.

Connectivity has revolutionized social interactions, fostering a sense of belonging and unity among individuals irrespective of their physical location. Social media platforms have become virtual communities where people connect, share their thoughts, and engage in meaningful conversations [3]. These online communities ignore geographical boundaries, bringing together individuals with shared interests and passions, leading to the formation of global networks and collaborations. Moreover, connectivity has enabled marginalized communities to find their voices.

The global economy was reshaped by facilitating seamless trade and creating new opportunities for businesses and entrepreneurs. E-commerce platforms have opened markets beyond borders, allowing small-scale enterprises to reach a global customer base [4]. Through online platforms, individuals can now engage in freelance work, connecting with clients and employers worldwide. Connectivity has fostered economic integration, breaking down barriers to trade and investment, and fuelling economic growth and development.

Brand associations represent the connections or associations that consumers make between brands and various attributes like quality, value for money, or reputation [5]. The impact of these associations is significant since strong and positive associations can lead to fast purchasing decisions and help increasing the brand loyalty.

In the past, brand associations were mostly viewed as a way of highlighting the differences between products or services provided by different companies [5]. By focusing on building positive associations with their own brands using advertising, creating a strong visual identity (logos or mascots), or other promoting methods the companies would gain a better reach to the customers and obtain a clear and recognizable brand image.

Recently, the way that companies view and approach brand associations has become more complex. Simply focusing on positive associations is not enough and many companies are looking to establish a network of positive associations including links with other brands, organizations, and causes that match their values and missions [6].

A good example is the concept of “purpose-driven branding,” where companies focus on aligning their brand with a specific social or environmental cause in order to build positive associations and connect with consumers who care about those issues [7]. Another approach is “experience-driven branding” where the focus is on creating a general experience for the customer that aligns with the brand message and values.

Smartly connected brands rely on data and technology to identify opportunities for partnerships and collaborations that align with their values and mission [8]. They leverage these partnerships to reach new audiences, gain access to new resources and expertise, and create new opportunities for growth and innovation. For example, a brand may partner with a non-profit organization to support a social cause or collaborate with another brand to develop a new product or service. These partnerships help to build positive associations with the brand, while also providing valuable resources and expertise that can help the brand to grow and evolve.

Using technology to personalize their communication with the customers, and to create a seamless customer experience across different touchpoints is another approach used by smartly connected brands. This can be achieved by gathering data from customer interactions with the brand and using it to personalize the customer experience and build loyalty.

Overall, smartly connected brands are those that use data and technology to identify and pursue strategic partnerships and collaborations, which ultimately drive growth and innovation upwards, and create deeper connections with consumers.

In the following chapter, the research methodology will be further explained, followed by the bibliometric analysis conducted on brand associations, where also the topic of the bibliometric table is addressed, leading in the end to a conclusion part.

29.2 Research and Methodology

In our study, we employed the Web of Science platform, a widely respected bibliographic database, as our primary tool for conducting a bibliometric analysis. Initially, we conducted searches within the Web of Science database using the term “brand associations,” searching across all available fields, meaning that the search was not restricted to be found only in the title or abstract. Other constructs were not taken into consideration, because Web of Science returned the same results when searching “brand association” as for “brand associations.” This action was made because we wanted to see what was the pool of answers. From the results, it is shown in Fig. 29.1 the domains on which the most articles have been written. From those results, there were excluded the domains which had nothing to do with the economy. Only categories like business, management, hospitality, leisure, environmental sciences, business finance, and behavioral sciences were selected. There was another filtering step, that which only journal articles and conference proceedings were chosen. The search results were then refined once more based on the number of citations, meaning that the dataset was ordered from articles with the highest number of citations to the smallest, and retaining the top 1000 records for further analysis. This sample size was chosen to be only 1000 results, because of the level of difficulty it would have risen for choosing the right articles for the bibliometric table. This approach has been used before by Zübeyir Çelik when creating a bibliometric analysis of Flow Theory [9].

Subsequently, the dataset containing all results after refining the dataset with articles only from the economic domain, was imported into VOSviewer for further examination. Various computations were performed, including the identification of the most co-cited authors and the most frequently used keywords.

To proceed, two experts in the field of Marketing were engaged to select at least one article per year from each dataset, focusing on its relevance to the respective terms. One of the specialists is a Marketing Professor and the other one is a branding specialist at a branding agency. For the construct of “brand associations,” a total of



Fig. 29.1 The domains which got the most publications for the term “brand association.” *Source* Computed from Web of Science

46 studies were selected. The criteria on which they selected the articles was that the articles should have the construct “brand associations” in the title. These chosen studies underwent a detailed bibliometric analysis, exploring seven dimensions.

For the visualization and analysis of co-occurrence maps and keywords, we utilized VOSviewer version 1.6.18, a software developed by Ness Jan van Eck and Ludo Waltman [10].

29.3 Bibliometric Analysis of Brand Associations

To get a broad view of what is the state of existing literature for the term “brand associations,” a bibliometric analysis was made. For this term there was specific information gathered, like the domains in which the term was mostly used, the number of published articles, the number of citations for a specific period, and an analysis using VOSviewer was done to show the co-citation of authors and the mostly used keywords.

29.3.1 Web of Science Info-graphs

When searching the term “brand associations” on Web of Science there were 8,952 results. There were no constraints regarding the search, so the term was found in all fields. This is a term that first appeared in 1979, while other connected terms like smart brands emerged in 1998 and smart objects in 1996. The publications are scattered in different domains and the first 3 places are occupied by the following: business,

management, and public environmental occupational health. A brand association is a mental connection a customer makes between your brand and a concept, image, emotion, experience, person, interest, or activity. This association can be immediately positive or negative and it heavily influences purchase decisions.

Since the field of interest of this research goes toward economy, the following fields were selected Business, Management, Economics, Communication, Neurosciences, Business Finance, Behavioral Sciences, Public Administration, and International Relations. The results after this selection were 3228 publications. The first publication on brand associations from the selected domains was in 1986. The trend is upward with 2020 and 2021 being the year with the highest number of publications and the year 2023 for citations.

The main type of documents which were published were articles, followed by proceeding papers and meeting abstracts. The main affiliations from the published articles are RLUK Research Libraries UK, the State University System of Florida, and the Pennsylvania Commonwealth System of Higher Education PCSHE. The most published authors are Brand Serge, Kowall Neil, and Levey Allan (Fig. 29.2).

From the 3228-result database, the first 1000 publications were chosen and uploaded in VOSviewer using the same computations as for the previous terms and the 3D co-citation visualization of authors resulted. The figure below shows that the author with the highest number of citations is Brand. There are three clusters linked to this co-citation. As seen in the figure below, Hollebeek has been co-cited by Kunkel, Smith, and Sigurdsson. While Kunkel has been cited by Funk, Kim, and Ko (Fig. 29.3).

The second type of analysis made was finding which are the most used keywords in the publications when searching for the term “brand associations.” There were 4286 keywords in total. The most used keywords are impact, satisfaction, behavior, word of mouth, social media, association, memory, etc. There are also some outliers like

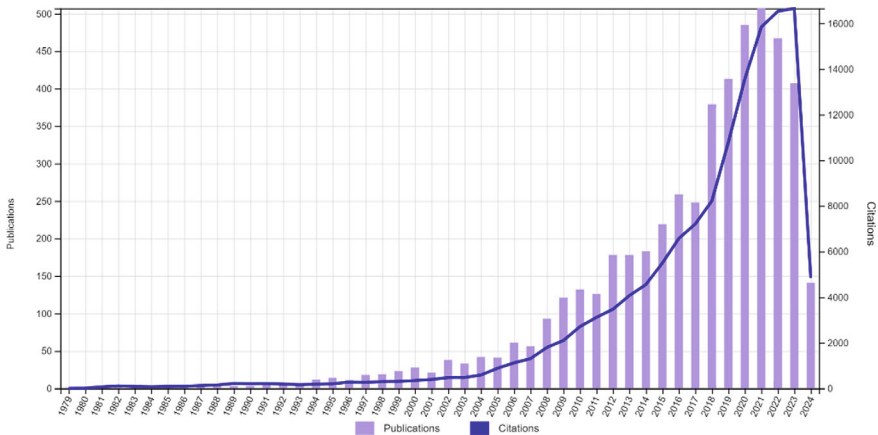


Fig. 29.2 Number of published articles and citations for brand association. *Source* Computed from Web of Science

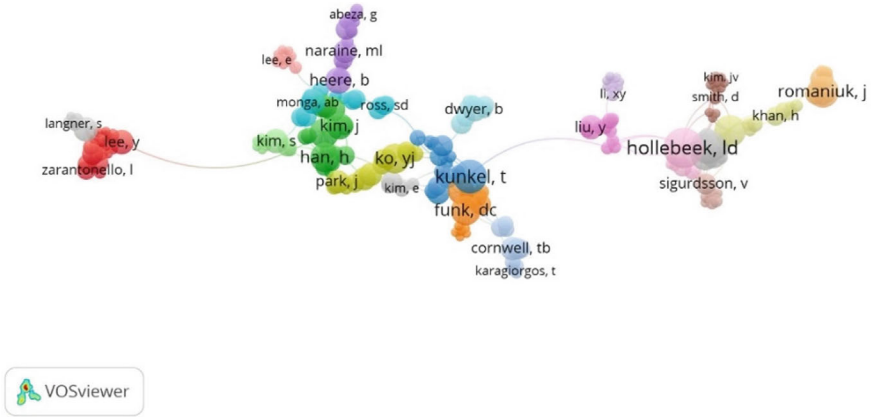


Fig. 29.3 Brand association articles co-citation. Source Own processing in VOSviewer

association patterns, social network analysis, network analysis, user participation, etc. These are very important keywords since association patterns will be used in the practical part of this research. There will be an analysis made on networks conceived from associations of smart consumers.

From Fig. 29.4, there are around four clusters which are formed based on color:

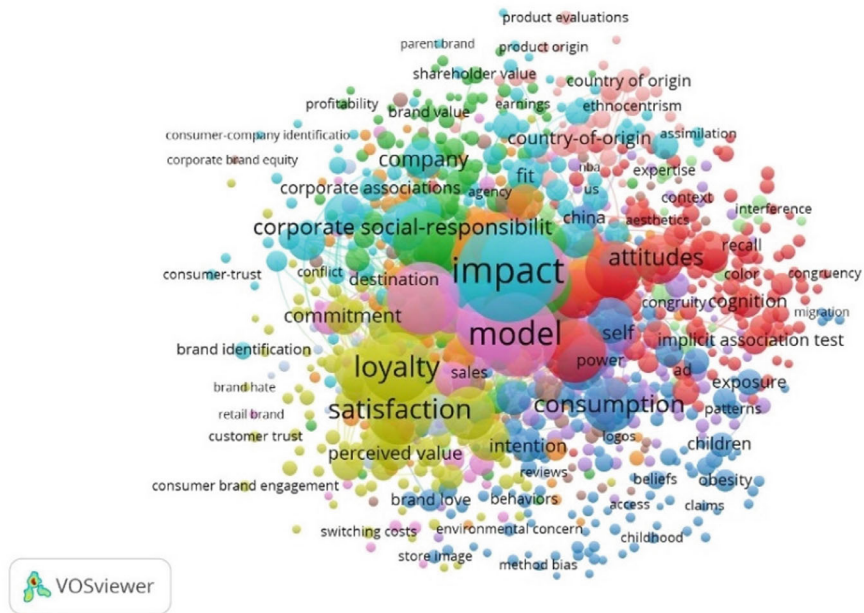


Fig. 29.4. Keywords for the term “brand association.” Source Own processing in VOSviewer

- Red color cluster: reflects the attitudes and cognitive side of consumers, consisting in words like recall, context, and association test.
- Dark blue color cluster: reflects the consumption side of consumers, consisting in words like intention, patterns, and exposure.
- Yellow color cluster: reflects the satisfaction of consumers, consisting in words like loyalty, perceived value, and commitment.
- Light blue color cluster: reflects the impact of corporate social responsibility, consisting in words like company, consumer trust, and brand identification.

29.3.2 Bibliometric Analysis for Brand Association

From the 1000 records database, the publications which are going to be used for the bibliometric analysis were selected. They were selected as following: for the years which have only one publication per year, that publication was considered, and from years which had more publications, one or two articles were chosen by specialists. There were 46 selected studies (Table 29.1).

As can be seen in the table above, a chronological analysis of the findings in the time period from 1989 to the present is given. Among the selected 46 studies, there are more studies with two authors (45,65%) journal articles (97,82%) and quantitative (71,73%) studies on brand associations. At the same time, it is noticeable that among the 46 studies chosen, the other most selected is the analysis unit (71,73%), and the sample size in most of the studies is less than one thousand or strictly one thousand (76,08%). The mixed method (63,04%) was used for data collection and all the studies are 100% open access.

29.4 Conclusions

The bibliometric analysis provided valuable insights into the current state of research on brand associations. By getting into the Web of Science database, we were able to identify key trends and areas of focus within this field: the number of publications on brand associations has steadily increased, particularly in the fields of business, management, and economics. This reflects the growing recognition of brand associations as a crucial factor influencing consumer behavior. The analysis revealed prominent authors in the field and highlighted potential collaborations based on co-citation patterns. This information can be valuable for researchers seeking to build upon existing knowledge. Overall, the bibliometric analysis serves as a valuable starting point for further research on brand associations since it has identified areas of established knowledge and emerging trends. From the bibliometric study, it is observable that there are many studies and articles written specifically on brand associations. However, there is still room for innovation, like for example from all the studies chosen, there were not so many qualitative ones when analyzing the “brand

Table 29.1. Bibliography of studies selected for this study

Bibliographic of selected studies	Time period												
	1979–1999		2000–2008		2009–2016		2017–2022		2020–Present		1989–Present		
	f	%	F	%	f	%	f	%	f	%	f	%	
Number of authors	One author	3	25%	3	21,42%							6	13,04%
	Two authors	7	58,33%	9	64,28%	5	38,46%					21	45,65%
	Three or more authors	2	16,67%	2	14,28%	8	61,53%	7	100%	4	100%	19	41,30%
Study attribute	Journal Article	11	91,67%	14	100%	13	100%	7	100%	4	100%	45	97,82%
	Conference paper/proceedings	1	8,33%									1	2,17%
Study type	Quantitative	7	58,33%	10	71,42%	10	76,92%	6	85,71%	4	100%	33	71,73%
	Qualitative	4	33,33%	3	21,42%	1	7,69%					8	17,39%
	Mixed	1	8,33%	1	7,14%	2	15,38%	1	14,28%			5	10,86%
Analysis unit	Student	3	25%			2	15,38%	3	42,85%	2	50%	8	17,39%
	Other	6	50%	13	92,85%	11	84,61%	3	42,85%	1	25%	33	71,73%
	Source	3	25%	1	7,14%			1	14,28%	1	25%	5	10,86%
Sample size	A thousand and below	8	66,67%	11	78,57%	11	84,61%	5	71,42%	3	75%	35	76,08%
	Over a thousand					1	7,69%	2	28,57%	1	25%	3	6,52%
	Not quantitative	4	33,33%	3	21,42%	1	7,69%					8	17,39%

(continued)

associations” term. Even if quantitative studies prove facts, there are still many ideas and concepts which could be discovered by using the qualitative method. There is a noticeable gap in existing literature regarding the fact that “brand associations” is a term mostly used in marketing, however, from the published articles, the majority is written in the field of business.

This comprehensive understanding of the current research landscape paves the way for a deeper exploration of how brands can build and leverage strong associations to connect with consumers in a meaningful way.

Limitations for this research are residing in the fact that the only chosen database was the Web of Science, meaning that some of the articles which can be found in other online databases have not been taken into consideration.

References

1. Hakvoort, G., Beale, R., Ch'ng, E.: Connect and connectivity: revealing a world of interactions. In: CHI '13 extended abstracts on human factors in computing systems, in CHI EA' 13, pp. 1647–1652. Association for Computing Machinery, New York, NY, USA (2013). <https://doi.org/10.1145/2468356.2468651>
2. Warschauer, M.: Technology and social inclusion: Rethinking the digital divide. MIT Press
3. Preece, J., Maloney-Krichmar, D.: Online communities: Focusing on sociability and usability
4. Savrul, M., Incekara, A., Sener, S.: The potential of E-commerce for SMEs in a globalizing business environment. *Procedia—Social Behav. Sci.* **150**, 35–45 (2014). <https://doi.org/10.1016/j.sbspro.2014.09.005>
5. Thellefsen, T., Sørensen, B.: What brand associations are. *Sign Systems Studies* **43**, 191 (2015). <https://doi.org/10.12697/SSS.2015.43.2-3.03>
6. Priluck Grossman R (1997) Co-branding in advertising: developing effective associations. *J. Product Brand Manag.* **6**(3), 191–201 (1997). <https://doi.org/10.1108/10610429710175709>
7. Berglind, M., Nakata, C.: Cause-related marketing: More buck than bang? *Bus. Horiz.* **48**(5), 443–453 (2005). <https://doi.org/10.1016/j.bushor.2005.04.008>
8. Patil P.: (2018) How brands connect to technology. In: von Gizycki V, Elias CA (eds), Omnichannel branding: Digitalisierung als basis erlebnis- und beziehungsorientierter markenföhrung, pp. 21–44. Springer Fachmedien, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-21450-0_2
9. Çelik Z, Uslu A (2022) Bibliometric analysis of flow theory from past to present with visual mapping technique: A marketing-sided approach. *Öneri Dergisi*, 17(57), Art. no. 57. <https://doi.org/10.14783/maruoneri.990480>
10. “VOSviewer—Visualizing scientific landscapes.” Accessed 31 Mar 2024. [Online]. <https://www.vosviewer.com/>

Chapter 30

Community Detection on an European Airports Network



Mioara Băncescu and Adina Gabriela Iliuță

Abstract Topics related to community detection and complex network indicators are increasing notoriety in the literature. In this research, we focus on network indicators such as average degree, average clustering coefficient, average path length, network density, network diameter, nodes centrality, and on detecting communities with the Louvain algorithm. We build an European airports network and apply and discuss the network indicators, as well as the community detection algorithm. The primary aim of the research is to analyze several network indicators and to check how the split of the network in communities is changing the range of values for the respective network indicators when computed at the community level. The paper addresses a novel approach in the field of air transportation, firstly by representing 50 European airports and the flights operated with the help of a network, and then by illustrating how community detection can be applied to identify relevant groups of airports.

30.1 Introduction

Complex adaptive systems are found all around us, most real systems are complex adaptive ones. Connectivity and interdependence in complex adaptive systems imply that a decision or action made by one element of the system will affect other elements of the system. What makes a system to be a system and not simply a collection of elements are the connections and interactions between its components.

Studying connectivity and interdependence in complex adaptive systems is related in practice to studying complex networks. Many measures and methods have been developed to analyze complex networks, including measures of node importance,

M. Băncescu (✉) · A. G. Iliuță
Bucharest Academy of Economic Studies, Bucharest, Romania
e-mail: mioara.bancescu@csie.ase.ro

A. G. Iliuță
e-mail: iliutaadina19@stud.ase.ro

node clustering, network modularity, the correlation between degrees of neighboring nodes, or extraction of network community structures. Community detection in networks and a diversity of network indicators have attracted great scientific interest.

The network is a heavily overloaded term used to denote relational data in a vast number of applications, and network analysis is carried out in areas such as social networks, transportation systems, electrical circuits, communication networks, epidemiology, bioinformatics, hypertext systems, text analysis, bibliometrics, organization theory, project planning, complex systems, genealogical research or event analysis [2].

The primary novel contributions of this study in the field of air transportation are: firstly to represent 50 European airports and the flights operated with the help of a network, and then to illustrate how community detection can be applied to identify relevant groups of airports.

The paper proceeds as follows. In Sect. 30.2 we present related work, covering relevant studies using network analysis and network clustering. Section 30.3 introduces several network indicators and the Louvain algorithm. In Sects. 30.4 we present entry data, the European airports network designed, computational results after applying the network measures and Louvain algorithm. Conclusions and future research avenues are highlighted in Sect. 30.5.

30.2 Related Work

In the field of air transportation, airport networks have been widely studied. Wang et al. [10] demonstrated that the construction of an airport network from delay data can provide new insights into the fundamental properties of the air transport system and, from a more practical point of view, have identified several influential airports and general patterns in the air transport system, after carrying out the comparison between delay networks in China and USA. Ceria et al. [3] have modeled airport traffic congestion contagion on three airport transportation networks, aiming to identify the airport's vulnerability, i.e., probability of being congested, using nodal infection probabilities.

Network clustering quantifies the propensity of two neighbors of the same vertex to also be neighbors of each other, forming triangle-shaped configurations in the network [9]. Asim et al. [1] have performed extensive experimentations to compare the performance of their proposed network clustering methods with the state-of-the-art algorithms Clauset, Newman's eigenvector, and Louvain algorithm, on five real graph datasets. Makris et al. [7] proposed a novel distributed methodology exploring the value of community detection algorithms for the text annotation problem, introducing community coherence metric.

30.3 Methods

Node degree, isolated network node, average degree, average clustering coefficient, average path length, network density, network diameter, and closeness centrality are introduced at the beginning of this section. More details are available to review in [8].

The degree of a vertex i , d_i , is given by the number of edges incident to that vertex. If no edge is incident to a vertex, we call the respective node an isolated network node. Network average degree \bar{d} , the average number of incident edges to a node, is an indicator that measures how many incident edges a network node has, on average. It is calculated as twice the number of edges, nE , divided by the number of vertices, N :

$$\bar{d} = \frac{2 * nE}{N} \quad (30.1)$$

Network density measures the number of edges of the graph relative to the maximum possible number of edges in the graph:

$$Den = \frac{2 * nE}{N(N - 1)} \quad (30.2)$$

The average path length in the network, l , is given by the average number of edges to be traversed on the shortest path between any two nodes of the network:

$$l = \frac{2}{N(N - 1)} \sum_{i=1}^N \sum_{j=1}^N l_{min}(i, j), i \neq j \quad (30.3)$$

where $l_{min}(i, j)$ is the shortest distance between nodes i and j , i.e., the minimum number of edges to be traversed on the shortest path between the two nodes.

The diameter of the network D represents the longest minimum path in the network:

$$D = \max(l_{min}(i, j)) \quad (30.4)$$

Node clustering coefficient, C_i , is defined as the ratio between the number of existing links E_i and the number of potential links between the neighbors of the analyzed node:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (30.5)$$

where k_i represents the number of neighbors of node i . And the average network clustering coefficient, \bar{C} , signifies how dense the vicinity of network nodes is, on

average, being calculated as an average of all network nodes clustering coefficients:

$$\bar{C} = \frac{\sum_{i=1}^N C_i}{N} \quad (30.6)$$

Node closeness centrality, $C_c(i)$, indicates the importance (centrality) of a node in the network from the perspective of the node's proximity to all other nodes in the network. It is calculated as the ratio of the maximum possible number of connections to the respective node and the sum of the shortest distances from that node to the other nodes of the network:

$$C_c(i) = \frac{N - 1}{\sum_{j=1, j \neq i}^n l_{min}(i, j)} \quad (30.7)$$

Louvain algorithm allows detecting communities (groups, modules, clusters) among the nodes of the network. As a first step, it involves computing the modularity indicator. The network modularity measures the potential for dividing the network into communities based on the density of connections between nodes within each community. The calculation formula is based on the difference between the number of edges existing within a community and the maximum potential number of edges within that community:

$$Q = \frac{1}{2 * nE} \sum_{i,j} (A_{i,j} - \frac{d(i)d(j)}{2 * nE}) \delta(C^{(i)}, C^{(j)}) \quad (30.8)$$

where $A_{i,j}$ denotes adjacency matrix elements, having values of 1 if there is a connection between nodes i and j , or 0 otherwise, $d(i)$ measures the degree of vertex i , $d(j)$ measures the degree of vertex j , $C^{(i)}$ denotes the community for vertex i , $C^{(j)}$ denotes the community for vertex j , and $\delta(C^{(i)}, C^{(j)})$ is *Kronecker delta* function:

$$\delta(C^{(i)}, C^{(j)}) = \begin{cases} 1, & \text{if } C^{(i)} = C^{(j)} \\ 0, & \text{otherwise} \end{cases} \quad (30.9)$$

As the next step of the algorithm, every vertex is a community. Then, according to the change in edge density at the cluster level, the community gradually merges with its neighbors. At every iteration, the network modularity is recalculated: if the indicator value increases with the merge of a node into a community, that node is kept in the respective community, otherwise not. The algorithm stops when it cannot obtain a gain in network modularity by moving any node in the neighborhood community. More on the Louvain algorithm can be reviewed in Kanavos et al. [5] and Gupta [4].

30.4 Results and Discussion

The input data used in the case study was collected from an airline company report (Lufthansa 2022). It contains two data sets: one for the nodes of the networks, consisting in 50 European airports, and another data set with 366 observations, for the edges of the networks, representing the regular flights the company operated at the analysis time, across the airports. A portion of the input data is shown in Tables 30.1 and 30.2.

As shown in Fig. 30.1, we created a graph, with the elements of the adjacency matrix taking value 1, if two airports are connected, and 0, if the respective two nodes are not connected via a direct flight offered. The airports network designed in Gephi provides a visual representation of the entire input dataset.

We apply the concepts of node degree, average degree, average clustering coefficient, average path length, network density, network diameter, and closeness centrality which we described previously in the methodology part. The results in Tables 30.3, 30.4, and 30.5 were computed by using the RStudio program.

Table 30.1 A portion of the network nodes input data

Airport ID	Airport label
IST	Istanbul
LHR	Heathrow London
CDG	Charles de Gaulle Paris
AMS	Schiphol Amsterdam
MAD	Madrid
FRA	Frankfurt
BCN	Barcelona
MUC	Munich
VKO	Vnukovo

Table 30.2 A portion of the network edges input data

Source	Target
FRA	CDG
CDG	FRA
FRA	NCE
NCE	FRA
FRA	LYS
LYS	FRA
FRA	MAD
MAD	FRA
FRA	BCN
BCN	FRA

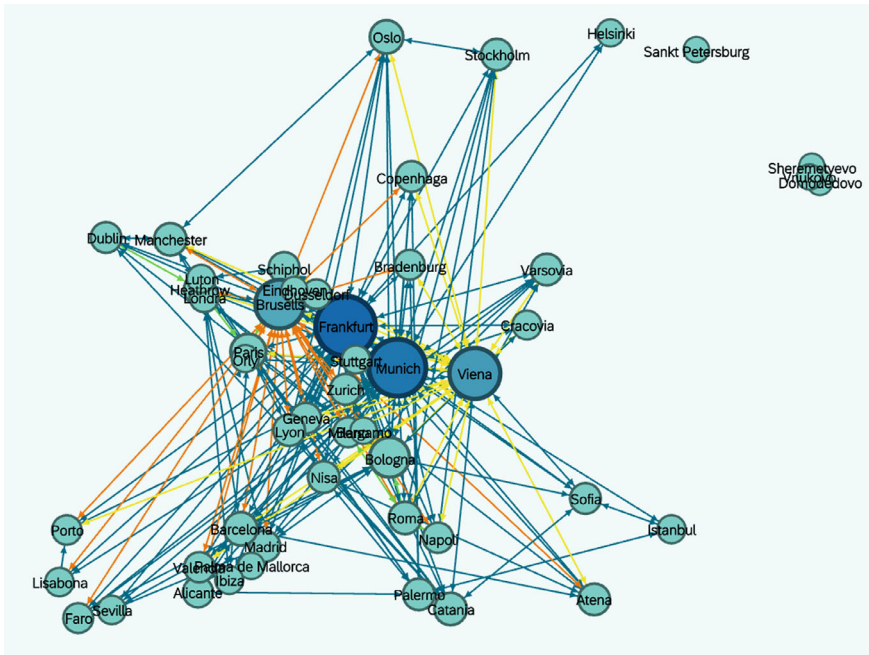


Fig. 30.1 The graphic illustration of the network

Table 30.3 Results of network indicators

Network indicator	Value
Density	0.298
Diameter	5
Average clustering coefficient	0.63
Average path length	2
Average degree	14.93

Table 30.4 Most relevant results of nodes degree and closeness centrality

Node label	Node degree	Node closeness centrality
Frankfurt	78	0.0192
Munich	72	0.0182
Vienna	58	0.0161
Bruxelles	53	0.0156
Bologna	33	0.0133
Paris CDG	23	0.0125
Barcelona	18	0.0122
Athens	14	0.012
Nice	14	0.0117

Table 30.5 Results of community indicators

Community	Community nodes	Community density	Community average path length
Com. 1	FRA, BLQ , IST, MAD, SPC, FAO, WAW, KRK, SVQ, SOF	0.77	1.44
Com. 2	MUC , LHR, LIS, DUB, MAN, HEL, OPO, STR	0.75	1.40
Com. 3	CDG, NCE, ATH , LGW, ORY, IBZ, BGY	0.66	1.67
Com. 4	BRU, BCN , AMS, CPH, BER, ALC, LTN, VLC, DUS, PMO, EIN	0.53	1.66
Com. 5	VIE , ZRH, OSL, ARN, GVA, FCO, NAP, CTA, MXP, LYS	0.82	1.42

The network density of 0.298 shows that the network is relatively sparse, having nearly 30% of the maximum possible number of connections. Airline companies like *Lufthansa* are taking in considerations a mix of economic, social, and political criteria when deciding which routes to operate. The network diameter indicates that the longest shortest path between any two network airports is of 5 connections. The average clustering coefficient having the value of 0.63 shows that the density of connections in the immediate neighborhood of the nodes is medium to high, and this can be interpreted as a signal that the network can be split into several subgraphs, with a high density of connections inside each subgraph. It can be observed that on average, to arrive at a network node from any other network node, it is needed to use two airport connections, as the average path length is of 2. The average network degree is 14.93 , indicating that on average a network node has nearly 15 connections with other network nodes.

For the network nodes are isolated, no airline connection being offered by the company at the analysis time, in the context of the Russian-Ukrainian military conflict.

Looking at node degree, the results show *Frankfurt International Airport* as having the highest number of connections with other European airports, 78 connections, followed by *Munich Airport* (72 connections), and *Vienna Airport* (58 connections). The closeness centrality indicator confirms the same ranking among airports indicated by node degrees: *Frankfurt*, followed by *Munich*, *Vienna*, *Bruxelles*, *Bologna*, *Paris CDG*, *Barcelona*, *Athens*, and *Nice*.

Figure 30.2 illustrates the communities of nodes detected after applying the Louvain algorithm, an algorithm which was detailed in Sect. 30.3. The network nodes are assigned to five communities. The isolated network nodes are not assigned any community by the algorithm. Table 30.5 shows the componence of each community, community density, and community average path length.

It can be observed from Table 30.5 that Communities 1 and 5 have 10 nodes each, community 2 has 8 nodes, community 3 has 7 nodes, and community 4 has 11 nodes. In each community there is at least one airport node with a high degree and

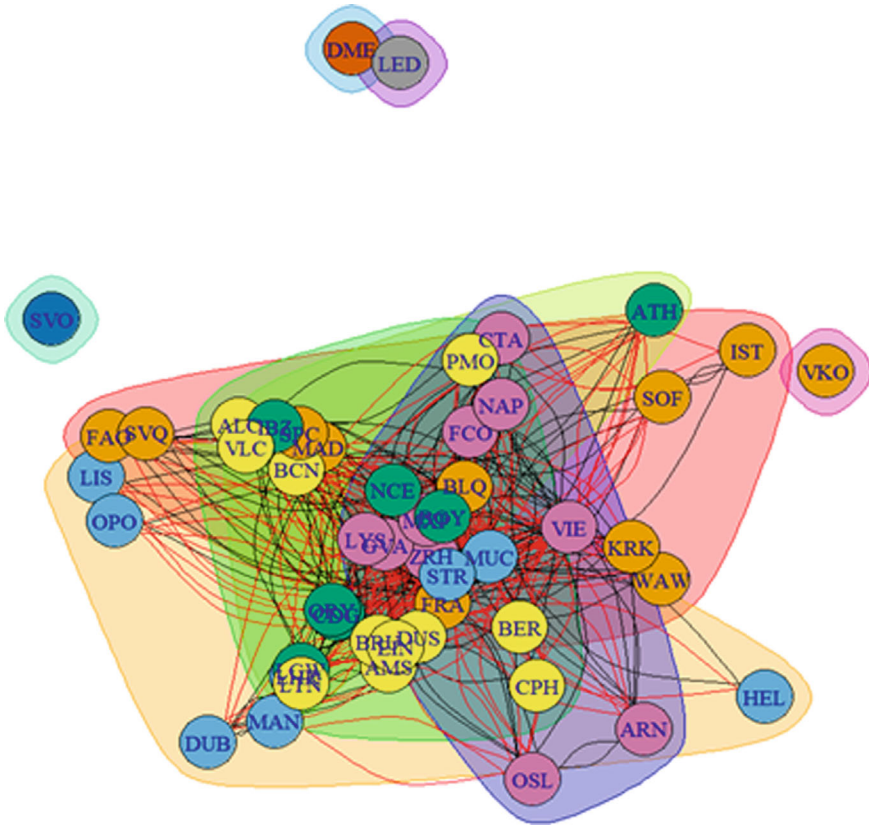


Fig. 30.2 Network communities

closeness centrality: *Frankfurt* and *Bologna* in community 1, *Munich* in community 2, *Paris CDG*, *Nice*, *Athens* in community 3, *Bruxelles*, *Barcelona* in community 4, and *Vienna* in community 5.

The density of connections inside each community ranges from 0.53 (community 4) to 0.82 (community 5). The results show that each community has a higher density of connections than the network density (0.298). Also, the split of the network in communities led to a decrease of the average path length inside each of the 5 groups. If the average path length for the network is 2, inside each community, the average path length ranges from 1.4 to 1.67, meaning that in average, to arrive to a community network node from any other community network node, it is needed to use less than two airport connections.

30.5 Conclusions

In this paper, we addressed the problem of detecting communities in networks with the Louvain algorithm. We illustrated how community detection can be applied to identify relevant groups of European airports, that looked like network nodes. Also, we calculated several network indicators, average degree, average clustering coefficient, average path length, network density, network diameter, and nodes closeness centrality on an European airports network we have designed. We analyzed those network indicators and determined that the split of the network in communities is changing the range of values for the connection density and for the average path length indicators when computed at the community level. We have seen that each community has a higher density of connections than the network density and that the community split of the network led to a decrease in the average path length inside each group.

We described a methodology that can be used to compute network indicators, useful to measure relationships among sets of entities in a network, or to detect communities of nodes among the total set of network nodes. In the case study, we created a network consisting in 50 nodes representing European airports, and 366 network edges, representing regular flights Lufthansa company operates across the airports. We discussed the 5 communities detected based on maximizing network modularity.

Future research avenues are addressing more diverse algorithms for community detection in networks and how the study could be extended to integrate network indicators into various networks: in economics, in education, in social networks, communication networks, or transportation ones.

References

1. Asim, Y., Ghazal, R., Naeem, W., Majeed, A., Raza, B., et al.: Community detection in networks using node attributes and modularity. *Int. J. Adv. Comput. Sci. Appl.* **8**, 1 (2017). <https://doi.org/10.14569/IJACSA.2017.080148>
2. Brandes, U., Erlebach, T.: *Network analysis methodological foundation*. Springer-Verlag, Berlin Heidelberg (2005)
3. Ceria, A., Köstler, K., Gobardhan, R., Wang, H.: Modeling airport congestion contagion by heterogeneous SIS epidemic spreading on airline networks. *PLoS ONE* **16**, 1 (2021). <https://doi.org/10.1371/journal.pone.0245043>
4. Gupta, M.: Community detection in a graph using Louvain algorithm with example. *Data Science in your pocket*. (2022). <https://medium.com/data-science-in-your-pocket/community-detection-in-a-graph-using-louvain-algorithm-with-example-7a77e5e4b079>. Accessed 15 Mar 2024
5. Kanavos, A., Voutos, Y., Grivokostopoulou, F., Mylonas, P.: Evaluating methods for efficient community detection in social networks. *Information* **13**(5), 209 (2022). <https://doi.org/10.3390/info13050209>
6. Lufthansa Group.: Traffic figures. (2022). <https://investor-relations.lufthansagroup.com/en/publications/traffic-figures.html>. Accessed 15 Feb 2024

7. Makris, C., Pispirigos, G., Michael, A.S.: Text semantic annotation: A distributed methodology based on community coherence. *Algorithms* **13**, 7 (2020). <https://doi.org/10.3390/a13070160>
8. Naimzada, A., Stefani, S., Torriero, A.: *Networks, topology and dynamics. Theory and Applications to Economics and Social Systems*. Springer-Verlag Berlin Heidelberg (2009). <https://doi.org/10.1007/978-3-540-68409-1>
9. Newman, M.: *Networks: An introduction*, 1st edn. Oxford University Press, Oxford (2010)
10. Wang, Y., Zheng, H., Wu, F., Chen, J., Hansen, M.: A comparative study on flight delay networks of the USA and China. *J. Adv. Transp.* (2020). <https://doi.org/10.1155/2020/1369591>

Chapter 31

The Impact of Minimum Wage on Employment. The Case of Romania



Georgiana Stănciulescu and Mădălina Ecaterina Popescu

Abstract The minimum wage is probably one of the most debated topics among economists, and its impact on other macroeconomic variables is challenged by numerous studies. Therefore, it is of high importance to understand how the minimum wage influences the economy. The heightened attention that this index receives at the European level is noticeable both through the recently published directives and through the numerous studies dealing with this subject. This article aims to analyze how the minimum wage, together with a set of other relevant economic indicators, influences the employment rate in Romania, both in total and among young people, because the expectation is that the impact is greater among young people, the latter being the ones who can get easier a minimum salary at the beginning of their career. The considered period is 2000–2022, and the examined variables are the following: in-work productivity, minimum wage, and the employment rate. Regression models were estimated for each category separately: total employment, as well as young women and men under 24 years, male and female employed.

31.1 Introduction

The minimum wage does represent one of the most discussed topics both at the national and international levels and European states are still trying to define this concept and determine how it influences the economy or parts of it. The theories are multiple, and the definitions are also different depending on the perspective from which we analyze its influence. One of the most accurate definitions of the

G. Stănciulescu (✉) · M. E. Popescu
The Bucharest University of Economic Studies, Bucharest, Romania
e-mail: georgiana.visan@csie.ase.ro

M. E. Popescu
e-mail: madalina.andreica@csie.ase.ro

M. E. Popescu
The National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania

minimum wage is provided by the International Labor Organization, which describes the minimum wage as “the minimum amount of remuneration that an employer is required to pay wage earners for the work performed during a given period, which cannot be reduced by collective agreement or an individual contract”. Depending on the perspective, definitions may vary; however, the concept of minimum wage is well known by the entire society.

As there are multiple definitions in the literature, each of them refers to a particular aspect of the minimum wage: some are providing information on the nature of this indicator, while others claim its calculation or indexing method. All of them are important and contribute to the overall understanding of this concept. In terms of regulatory limitations, the minimum wage can be set by the state or other competent national authorities, courts, and councils. There are also situations when the minimum wage is set up by collective agreements and also the case when no minimum wage is set on a national/industry level. However, official statistics provided by the International Labor Organization show that in 90% of ILO Member States, Europe is the continent with the highest percentage.

Being such an important topic, economists have been permanently eager to understand and predict the consequences arisen from the evolution of the minimum wage. In particular, for this article, the interest was how it may influence employment. One can quantify the impact of the minimum wage on the employment rate by searching for several regression models which provide relevant information on how this interaction evolves. From a technical perspective, we chose to analyze the relationship between minimum wage and employment rate by considering three categories: total population, females under 24, and males under 24. The reason for choosing the three age groups in the analysis is because the literature indicates a higher impact of minimum wage on young people [4].

31.2 Literature Review

The complexity of the overall minimum wage mechanism is that it can strongly affect the companies’ behavior in relation to employees and authorities. According to most studies regarding minimum wage, firms may intentionally modify employees’ schedule or other aspects in order to respond to minimum wage rises. One of the most recent definitions officially given for the minimum wage belongs to ILO 2014, par. 68.

The worldwide literature is very generous on the subject of minimum wage, with studies showing that among the most important aspects in reducing the national income disparity remains the minimum wage level [6]. Even though at the worldwide level the problematic of minimum wage is well explained and captured, the analytical analysis of wage distribution is partially incomplete for the specific case of Romania, considering the strong disruption of the reforming process on the labor market. Moreover, an important aspect in studying the minimum wage process and behavior is the political field, which in the case of Romania has been quite unstable lately.

Economists have always been interested to investigate the consequences of minimum wages on the occupation level, but also on the overall phenomena generated by the minimum wage policy on the labor market level. In 1946, George Stigler, a remarkable American economist, formulated the first neoclassical hypotheses on the consequences of minimum wage policies, which were at that time contested by Lester economist. The difference between the neoclassical approach and the new economic Lester's approach on the impact of salary is mainly explained by the source of examined data. While the neoclassical approach aims to analyze annual macroeconomic data, the new economic approach uses monthly information procured from population surveys at the microeconomic level. However, the minimum wage remains an extremely controversial topic continuously debated, hence we often find theoretical and empirical contradictory arguments.

Several authors tried to explain or to provide the literature with an exhaustive definition of the minimum wage and to establish this macroeconomic mechanism. According to Levin-Waldman, the most important reason for establishing a minimum wage mechanism is to set up a legal ground for workers to earn at least above the subsistence level. In fact, it is of high importance a minimum wage threshold might provide the person the opportunity to live above the poverty line. This was also one of the earliest arguments for implementing a minimum wage policy in the nineteenth century. Moreover, it is of high importance for a state to assure its population of decent living conditions, one aspect of this network being the minimum wage threshold.

Moreover, the literature tries to provide with a set of answers regarding how the minimum wage threshold would affect productivity or other macroeconomic indicators [9]. One of the first perspectives on this aspect is given by an English economist, Webb [10], who stated that there is a direct relationship between the way the employers are paid, their energy, motivation, and loyalty to the employers. Although the minimum wage is a controversial topic in Romanian politics, the literature often shows that there is a positive effect generated by the minimum wage increase on the Romanian economy and society. In the most recent years, Romania experienced an accelerated increase in the minimum wage, especially between 2014 and 2018: from 900 to 1900 RON, registering an increase of above 200%. Inflation was kept under control in this period, and so was the unemployment rate. Therefore, these resulted in an increase in incomes for the part of the labor force with the lowest bargaining power.

According to the 23rd article of the Universal Declaration of Human Rights, "everyone who works has the right to just and favorable remuneration ensuring for himself and his family an existence worthy of human dignity, and supplemented, if necessary, by other means of social protection". Therefore, considering each development stage of the Member States, the minimum wage threshold may guarantee the fulfillment of the basic needs of employees. In 1928, the International Labor Organization (ILO) has recognized for the first time the concept of minimum wage and provided a set of recommendation throughout the Minimum Wage Fixing Machinery Convention, no 26. The use of minimum wage was however also a controversial subject at that time, both from the economic and the political perspectives. There is a major contradiction between the social-based arguments and the economic-based

theories [12]. While for the first category, the promoters claim that rising minimum wage leads to an increase in the living standards of the low-skilled workers, the latter highlights the possibility of a negative impact on a set of economic indicators like competitiveness, inflation, and employment rates. The entire debate on minimum wage becomes more and more thrilling when it comes to the latest directive issued by the Commission on adequate minimum wages within Member States. The minimum policy therefore raises several questions, such as: what is the optimal level of the minimum wage? What is the impact of the minimum wage increase on the economic growth? Can we find a suitable mechanism for the minimum wage increase? How effective is minimum wage in reducing income inequalities? Is minimum wage a feasible mechanism to reduce the poverty rate among population? What are the biggest challenges for a state when implementing a minimum wage mechanism?

While the first studies were mostly describing how the minimum salary and employment rate are being linked, the literature was gradually enriched during the last few decades, when multiple correlations between minimum wage and other macroeconomic indicators have been seriously studied. One of the first econometric estimations underlying how the minimum wage affects the employment rate was developed in the early 80s. Back then, it was considered to be a common notion that a rise in the minimum wage threshold will decrease the firms' hiring initiatives and thus lower the employment [13]. The theory seemed to be so clear that most economists did not question it until the early 90s. In 1983, Schaafsma and Walsh performed empirical researches covering the 1975–1979 years and pointed out that minimum wage has a negative effect on the employment rate and labor force for all the ages groups included in their study. Generally speaking, teenagers and the lower-skilled workers are the ones mostly affected by the increase in the minimum wage, as most of them are being paid with this salary. Moreover, Brown suggested in several studies from 1982 to 1983 that the employment sensitivity to the minimum wage increase is somewhere between -0.1 and -0.3 . His team also mentioned that the higher values correspond to the employment of young people.

A decade later, Card and Krueger [15] performed several analyses on what the consequences of a minimum wage rise would be in the United States. Their study evaluated the industry of fast-food restaurants because it is commonly characterized by unskilled workers paid at lower rates. The analysis concluded that there is no negative effect of a minimum wage rise from 4 to 5 USD/hour on employment. Moreover, they even showed that for this kind of industry where there are mostly hired lower-skilled workers, there might be even a positive relationship between the two variables. Taking this into account, we might expect that the low-paid workers are positively sensitive to the minimum wage increases.

This perspective was seriously debated even later. In the 2000 years, Neumark and Wascher [16] have conducted several studies and found out that there might be higher elasticity levels for certain categories of employees. It has been even found there are countries having positive levels of elasticity. Studies like the one conducted by Alatas and Cameron [17] analyzed the case of some emerging economies and they proved that an increase in the minimum wage does not negatively affect large firms and corporations, but only the employment in smaller domestically owned firms. As

the time pass, most of the recent studies have shown that the impact of minimum wage on employment is either insignificant or positive, at least on long term.

While all of these studies aim to explain how minimum salary influences certain macroeconomic indicators, the case of Romania nowadays seems to be somehow restricted by the latest European Directive issued by the European Parliament: Directive (EU) 2022/2041 on adequate minimum wages in the European Union. With respect to this, Romania has to implement an adequate minimum wage setting mechanism by the end of 2024. Thus, from the most recent Eurostat data on the monthly minimum wage levels, it seems that Romania (having a minimum wage of around 600 euros) ranks among the last countries in the European Union, only above Hungary and Bulgaria. Although the minimum wage topic does not appear to be such a priority for a couple of countries, it is one of the subjects positioned as a priority on the work agenda of the European institutions.

31.3 Methodology and Data

In this paper, we aim to evaluate the minimum wage impact on employment for the case of Romania considering quarterly data for the 2000–2022 period. The empirical analysis was performed using EViews, which is a very powerful econometric tool, designed to support complex time series analysis. Moreover, it provides with most of the technical tests in order to verify the time series characteristics and also to check the statistical significance of our results (Table 31.1).

We used logarithmic series in order to ensure higher homogeneity in data. Afterward, all the variables were imported into EViews and some short summary statistics were checked. In terms of in-work productivity, the data shows that for each year, the productivity of employees is at its lowest level in the first quarter of each year and at its best value in the fourth quarter.

Table 31.1 Macroeconomic variables selected for this analysis

No	Variable name	Variable description
1	Minimum wage	Minimum wage was retrieved from the Ministry of Labor and Social Protection database and calculated in order to express real values
2	Productivity	Index of real labor productivity per employed person, retrieved from Eurostat database
3	Employment	Total employment rate, retrieved from the National Institute of Statistics
4	Employment rate, females under 24	Total female employment rate, retrieved from the National Institute of Statistics
5	Employment rate, males under 24	Total male employment rate, retrieved from the National Institute of Statistics

Source of data: Eurostat database

One of the first steps when working with time series is checking their stationarity. Stationary series are characterized by their relatively regular oscillations around a reference level. In general terms, a time series is considered stationary if it does not show any systematic change in its mean. One of the most utilized tests for checking stationarity is the Augmented Dickey-Fuller Unit Root Test. Being applied to each variable, it shows us its stationarity. Also, its result may be confronted with the correlogram. When a particular time series was not stationary, the test was applied on the first difference and the stationarity issue was solved. Despite the employment rate, all the other variables were differentiated at the first level in order to obtain stationarity and to be able to continue the investigation on how the minimum wage influences the employment rate. The analysis has been split into three categories:

- The impact of minimum wage on the total employment rate in Romania
- The impact of minimum wage on young female (under24) employment rate
- The impact of minimum wage on young male (under24) employment rate.

The second step of the analysis was estimating the regression model based on the indicators' significance. It was of high importance to include in the regression models the variable itself, with a respective lag. This is because it is predictable that the employment rate is somehow influenced by its own past-period value. The maximum lag considered in the analysis was four, however we have not reached it when estimating the regression equations.

After the regression equation had been estimated, several residual diagnostics were performed in order to verify error serial correlation, heteroscedasticity, and normality. The Breusch-Godfrey test has the null hypothesis that residuals are not autocorrelated. Its statistic involves building an "auxiliary" regression equation by regressing the residuals of the main equation on the initial explanatory variables and the "lags" of the residuals. For all three regression models, the p-value was more than 0.05, meaning that there is no autocorrelation within the data. In terms of heteroscedasticity, the Breusch-Pagan-Godfrey test was performed, suggesting the absence of heteroscedasticity among errors. Finally, the Jarque-Bera Probability for Normality Test was checked and confirmed for two out of the three regression models the presence of normally distributed errors. The summary of these findings is summed up in Table 31.2.

31.4 Main Results

Having in mind our purpose to investigate the influence of minimum wage, along with other relevant economic indicators, on employment in Romania, three econometric models were estimated: (1) the total employment model; (2) the young female employment model, and (3) young male employment model. The results are presented in Table 31.3.

The total employment rate seems to be influenced by its previous period level (first quarter lag), by the in-work productivity, and by the minimum wage. The

Table 31.2 Residual diagnostic tests performed in EViews

Residual test	Total employment model	Young female employment model	Young male employment model
Breusch-Godfrey serial correlation LM test (Prob. Chi-square)	0.195	0.756	0.963
Heteroskedasticity test: Breusch-Pagan-Godfrey	0.102	0.596	0.238
Jarque-Bera probability for normality test	0.367	0.02	0.869

Source Authors' own computation

Table 31.3 Regression models estimated in EViews

Variables	Total employment model	Young female employment model	Young male employment model
Lemploy_rate(-1)	0.685**		
Lfemale_employ24(-1)		0.844**	
Lmale_employ24(-1)			0.816**
Dlproductivity	0.070**	0.0807*	0.082**
Dlminimum_wage(-3)	-0.190**	-0.339**	-0.238**
C	1.2443**	0.5528**	0.713**
<i>R-squared</i>	0.595	0.734	0.733

** Statistically significant at 1% level; * At 5% level

Source Authors' own computation

positive coefficients of the lagged employment rate and the in-work productivity show that if any of these two variables record an increase in the following period, then the employment rate will also increase. The negative coefficient of the minimum wage, however, suggests that the minimum wage threshold from three quarters before influences the evolution of the employment rate in reverse. Considering that we were working with the first difference when it comes to minimum wage and in-work productivity, these measures may be interpreted in terms of differences also.

The findings corresponding to the young female and male employment models are rather similar, also suggesting the negative impact of the minimum wage on youth employment rates. Our results are consistent with previous studies on the topic. For instance, a similar negative three-quarters-delay shock of minimum wage on female employment in Romania was reported by Popescu et al. [1]. These findings suggest a plausible behavior of employers to a minimum wage shock, especially in cases of small and medium-sized enterprises, who have more minimum-wage-paid

employees, are more sensitive to such fiscal policy shocks, and have less financial leverage to adjust but to lay off young employees.

At the international level, other papers support our findings. For instance, Schaafsma and Walsh (1983) also brought empirical evidence in favor of a significant negative minimum wage effect on employment and labor force in almost all age groups in Canada, but with a definitely stronger effect observed on youth employment (15–19 years). Dolado et al. [3] also support such findings of a negative minimum wage shock upon youth employment in Spain. Finally, the results of Majchrowska and Zólkiewski's [2] on the Polish economy indicated that the minimum wage had a negative impact on employment in the period 1999–2010, with the strongest effects felt on young people.

31.5 Conclusion

This article aimed to provide with an overview of the way the minimum wage can influence the evolution of the employment rate, by estimating a set of regression models, for three categories: total, female under 24 and male under 24. This could be a first step in deciding upon an indexing mechanism for the minimum wage, considering the latest Directive issued by the European Commission on the minimum salary over the European Union. The three regression models estimated on total employment, young female and young male employment allowed us to investigate the impact of minimum wage and productivity on employment for the case of Romania. The data was retrieved from public sources like the National Institute of Statistics.

In conclusion, this analysis reiterated the results that the literature had promoted in the past. The results were reconfirmed for the period 2000–2022, showing that among young people the impact of the minimum wage on the employment rate is greater than overall. This shows that the young age group is much more sensitive to changes in the minimum wage level, compared to employees having more experience.

References

1. Popescu, M.E., Stanilă, L., Cristescu, A.: A gender analysis of the minimum wage effects upon employment in Romania. In: Proceedings of the IE 2015 international conference, pp. 444–449, Bucharest (2015)
2. Majchrowska, A., Zólkiewski, Z.: The impact of minimum wage on employment in Poland. *Investig Regnales*, **24**, 211–239, Lotz (2012)
3. Dolado, J., Kramarz, F., Machin, S., Manning, A., Margolis, D., Teulings, C.: The economic impact of minimum wages in Europe. *Econ Policy*, 319–372 (1996)
4. Schaafsma, J., Walsh, W.D.: Employment and labour supply effects of the minimum wage: some pooled time-series estimates from Canadian provincial data. *Can J Econ/Rev Can d'Economique* **16**(1), 86–97 (1983)
5. Andreica, M.E., Stanila, L., Cristescu, A., Vasilescu, M.D., Cataniciu, N.: Rolul salariului minim și tendințe ale performanței pieței muncii, Ed. Universitara, Bucharest (2015)

6. Militaru, E., Popescu, M.E., Cristescu, A., Vasilescu, M.D.: Assessing minimum wage policy implications upon income inequalities. The case of Romania. *Sustain* **11**(9), 2542 (2019)
7. Stigler, G.J.: The economics of minimum wage legislation. *Am. Econ. Rev.* **36**(3), 358–365 (1946)
8. Lester, R.A.: Employment effects of minimum wages. *ILR Rev.* **13**(2), 254–264 (1960)
9. Levin-Waldman, O.M.: *The minimum wage: A reference handbook* (Contemporary World Issues), ABC-CLIO, 1st edn, California (2015)
10. Webb, S.: The economic theory of a legal minimum wage. *J. Polit. Econ.* **20**(10), 973–998 (1912)
11. Minimum Wage Fixing Machinery Convention, no 26. https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C026. Last accessed on 20 Mar 2024
12. Avram (Zamfir), Elena-Mădolina, N.: *Inegalitățile de venituri—provocarea pentru România. Rolul politicii salariului minim în evitarea sărăciei* (Teză de doctorat), (2023). Last accessed on 20 Mar 2024
13. Santillan, A., Esquivel, G., Campos-Vazquez, R.: The impact of the minimum wage on income and employment in Mexico. *CEPAL Review*. No **122**, 189–216 (2017)
14. Kohen, A., Brown, C., Gilroy, C.: The effect of the minimum wage on employment and unemployment. *J. Econ. Literature.* **20**(2), 487–528 (1982)
15. Card, D., Krueger, A.B.: Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *Am. Econ. Rev.*, **84**(4), 772–793 (1994)
16. Neumark, D., Wascher, W.L.: Minimum wages and employment: A review of evidence from the new minimum wage research. NBER Working Paper no. 12663, (2006)
17. Alatas, V., Cameron, L.A.: *The impact of minimum wages on employment in a low income country: an evaluation using the difference-differences approach* (English). Policy, Research working paper series, no. WPS 2985. World Bank Group, Washington, D.C (2003)
18. Directive (EU) 2022/2041 of the European Parliament and of the Council of 19 October 2022 on adequate minimum wages in the European Union. Last accessed on 2024.03.20

Part V
Machine Learning Theory
and Applications

Chapter 32

The Analysis of the Real Estate Market Using the k -Nearest Neighbors Algorithm



Ion-Florin Răducu, Stelian Stancu , Andreea Pernici, and Denisa Bălă

Abstract Machine learning methods and techniques are a current reality in artificial intelligence, which is rapidly developing and encompassing most aspects of social life. The k -nearest neighbors (k -NN) algorithm is a prevalent machine learning method for classification and regression problems. It is based on the assumption that similar data points will have similar labels or values. k -NN is a non-parametric algorithm, meaning it makes no assumptions about the input's probability distribution. In this paper, we developed a model that uses the k -NN algorithm to study the USA real estate market based on house price classification. When looking to buy a property, one of the most important factors to consider is the price. Therefore, we trained a model to classify new data into a specific price category with its own set of characteristics, thereby facilitating buyer's decision-making processes.

I.-F. Răducu (✉) · A. Pernici · D. Bălă

Economic Cybernetics and Statistics Doctoral School, Bucharest University of Economic Studies,
Bucharest, Romania

e-mail: florin.raducu@csie.ase.ro

A. Pernici

e-mail: andreea.pernici@csie.ase.ro

D. Bălă

e-mail: baladenisa16@stud.ase.ro

S. Stancu

Bucharest University of Economic Studies, Bucharest, Romania

Centre for Industrial and Services Economics, Romanian Academy, Bucharest, Romania

S. Stancu

e-mail: stelian.stancu@csie.ase.ro

32.1 Introduction

Residential housing is an essential living resource and property for many individuals and families. Houses are not the only types of property available in cities, but they do provide an appealing market for investors [1]. The growth of the real estate market can have a significant impact on a country's or even the world's economic activities.

The reason for selecting this topic for the paper's preparation is based on the current state of machine learning techniques for building systems that can learn (or improve their performance) based on the data they process, as well as the dynamics of the real estate market, given the relatively rapid pace at which construction, price changes, and land or buildings are purchased. This theme is also appealing due to the diverse range of modeling possibilities, given the abundance of machine learning techniques, some of which are complex. This allows for comparisons of the obtained results and the determination of which models best fit the issues addressed.

There are four sub-fields of machine learning algorithms: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [2]. Supervised learning algorithms use input data sets in which the dependent variable is associated with a label, allowing for accurate predictions after the application of algorithms that train various models. This category includes the k -nearest neighbors algorithm (k -NN). It is a simple, quick, and straightforward classification algorithm [3]. Despite its simplicity, k -NN produces accurate predictions and, in some cases, outperforms more than complex algorithms. It is also effective at developing models based on large training data sets and robust to outliers.

In this paper, we will create a k -NN-type model to classify house-type properties on the USA real estate market into three price categories: low, medium, and high. We believe that this algorithm and approach can address some of the concerns of people who want an overview of the real estate market, the characteristics that can influence the price of a property, or the characteristics of some types of houses or apartments classified by price. Such an algorithm was chosen instead of a logistic regression algorithm, for example, which could also serve the purpose of the analysis, because another alternative to the classic logistic regression model was considered in order to see how it behaves, as well as the k -NN algorithm for classification.

This paper is divided into six parts: an introduction that presents the topic of the article and brief theoretical considerations about the model used for the case study, a chapter that summarizes some research studies from the literature relevant to the topic addressed in the article, a chapter that describes the methodology used to develop the case study, a section that presents the data used in the analysis, followed by a presentation of the obtained results and some brief conclusions.

32.2 Related Work

The literature includes a wide range of papers and books on machine learning methods applied to various fields of knowledge, as well as methods for studying the real estate market. Relevant research studies on this topic can be found in the literature, which outlines a scientific context and demonstrates the increased interest in the field. Rogers et al. demonstrated the utility and inherent nature of machine learning methods in medicine, particularly neural networks, for improving accuracy, safety, and automation, especially in the future [4].

Truong et al. compared and analyzed three different types of machine learning methods, including Random Forest, XGBoost (Extreme Gradient Boosting), and LightGBM (Light Gradient Boosting Machine), as well as two machine learning techniques aimed at hybrid regression and stacked generalization regression [5]. This study aims to compare traditional prediction methods with advanced machine learning methods.

In a recent study [6], Choy and Ho used techniques such as Extra Trees, k -Nearest Neighbors, and Random Forest to provide more accurate property price predictions. These three models, unlike traditional ones, have higher explanatory power and more effective error minimization.

Given that water quality and safety are critical for the smooth operation of nature and human life, Zamri et al. used the k -NN algorithm based on different distances and weighted distances to determine a more accurate classification of Terengganu river quality [7]. Although the Water Quality Index (WQI) is commonly used to measure water quality, it has been demonstrated that the algorithm chosen for this study provides a new approach for a more precise classification of river quality, particularly given the large volume of data. The Euclidean distance, Hamming distance, and entropy estimator were used to guide the approach. K -NN with entropy estimator achieved the highest accuracy of 99.90%.

Zheng recently studied the relationship between community resilience and real estate market value [8]. Using a variety of machine learning models (including Linear Regression, K -Neighbors Regressor, Gradient Boosting Regressor, Random Forest Regressor, and Extra Trees Regressor) on a real data set from the United States, he concluded that increased community resilience, which is critical in protecting people from natural disasters, leads to an increase in the value of the real estate market. The article's conclusions show that the Extra Trees Regressor (ETR) method outperforms other methods in terms of root mean squared error (RMSE) and is effective in predicting real estate prices in a different sample.

Yoon and Friel investigated how the probabilistic k -nearest neighbors model improves the performance of the traditional classification algorithm [9]. Given that selecting the optimal number of neighbors, k , remains a current problem in the case of both algorithms, their paper's contribution is to incorporate uncertainty k into decision-making, resulting in improved classification using Bayesian model averaging.

32.3 Methodology

The k -nearest neighbors algorithm is one of the simplest and most fundamental classification algorithms, and it should serve as the foundation for studying machine learning methods and classifications in situations where nothing or little is known about the data distribution. This algorithm arose from the need to perform discriminant analysis when the parametric estimates of the probability densities are unknown or difficult to obtain.

The application of the algorithm assumes that similar inputs have similar outputs [10]. The classification rule entails categorizing an object based on the most significant label obtained by analyzing the nearest k neighbors.

Formally, a point $x \in C$ is randomly selected, where $C \subseteq \mathbb{R}^d$ is a population. Let $D = \{x_1, x_2, \dots, x_n\} \subseteq \mathbb{R}^d$ a set of data and the set of the k nearest neighbors of x as S_x is defined. Formally, S_x is defined as $S_x \subseteq D$ and $|S_x| = k$, where $|S_x|$ represent the cardinal of set S_x and $\forall (x', y') \in D \setminus S_x$,

$$d(x, x') \geq \max_{(x'', y'') \in S_x} d(x, x'') \quad (32.1)$$

Equation (32.1) describes that every point in D , but not in S_x , is at least as far away from x as the furthest point in S_x .

The Euclidean distance is commonly used to calculate the similarity between all data points and the new data points [11]. Similarity between data objects is a widely used concept across multiple domains to solve pattern recognition problems such as categorization, classification, clustering, and forecasting [12]. The literature contains a variety of measures for comparing data objects.

A distance measure $d : X \times X \rightarrow \mathbb{R}$ is considered a metric if it fulfills the following requirements, $\forall x, y, z \in X$:

1. $d(x, y) \geq 0$;
2. $d(x, y) = 0$ if and only if $x = y$;
3. $d(x, y) = d(y, x)$;
4. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

Generally, the Euclidean distance is computed as follows:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (32.2)$$

If the usual metric for distance between continuous attributes is the Euclidean one from Eq. (32.2), then the Hamming distance can be used for nominal attributes as well.

We can define $h()$ as a function, as shown in Eq. (32.3), that returns the most common label in S_x :

$$h(x) = \text{mode}(\{y^n : (x^n, y^n) \in S_x\}) \tag{32.3}$$

where $\text{mode}(\cdot)$ represent the label of the highest occurrence.

32.4 Data

The data set used to create the case study includes real estate listings in the United States organized by state and zip code. The data was obtained from Kaggle. The database contains 10 variables of which 3 are nominal (*status*—housing status with two values: a—ready to sell or b—ready to build, *city*, and *state*) and 7 are numerical (*bed*—number of beds, *bath*—number of bathrooms, *acre_lot*—property/land size in acres, *zip_code*, *house_size*—living space in square feet, *prev_sold_date*—previously sold date, *price*—housing price is either the current listing price or the recently sold price if the house was sold recently).

32.5 Computational Results

The case study was conducted using R, a software that provides numerous advantages and opportunities for modeling, data manipulation, statistical analysis, and programming. Before running the actual algorithm, we performed a data pre-processing stage in which we eliminated duplicates and missing values, and we created a new column called *prev_sold_encoded* starting from the *prev_sold_date* column, where we replaced missing values with 0 and data-containing values with 1. We also created a new variable called *price_group*, which divides the houses by price into three categories: low price, medium price, and high price. This transformation of the numerical variable into a categorical variable was necessary to suit the purpose of the analysis and the method used, which was to categorize house prices into three groups.

The houses were also divided into price groups based on the database’s actual values, which were analyzed using descriptive statistics and summarized in Table 32.1 below:

According to Fig. 32.1, the majority of houses (around 31.000) fall into the group *Low price*, which has low prices.

Table 32.1 Descriptive statistics of the price

Variable	The min value	1st quartile	The median value	The mean value	3rd quartile	The max value
Price	500	279900	455000	803693	789500	169000000

Source Owner author’s processing

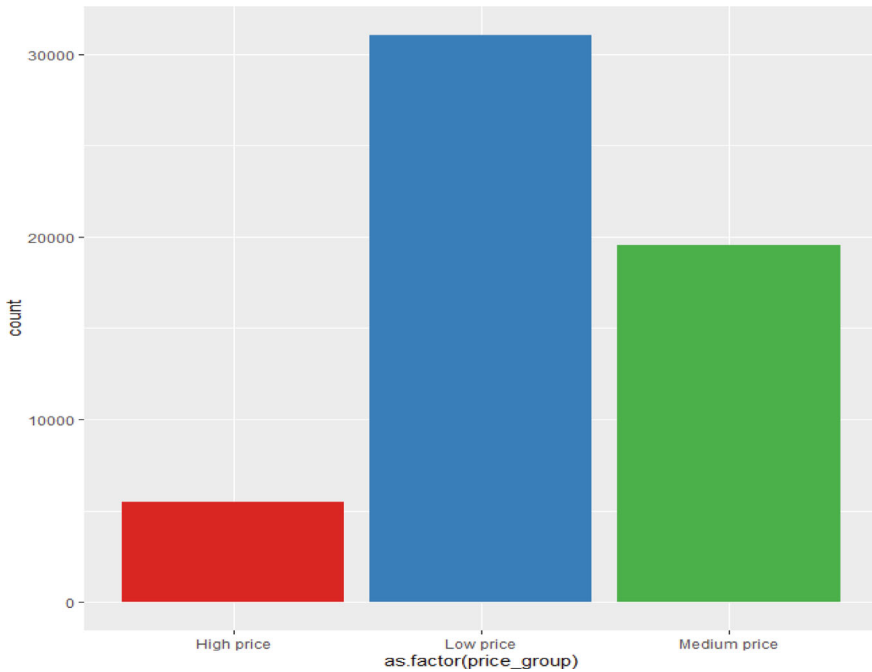


Fig. 32.1 Distribution of the number of houses in the three price categories. *Source* Owner author's processing

Generally, the k -NN algorithm can classify datasets using a training model similar to the testing query by considering the k nearest training data points (neighbors) closest to the query being tested [13, 14]. Therefore, before building the model, the data set was divided into two parts: one for training (70% of the data) and one for testing (30% of the data).

For creating the model, we used the *caret* library, which has the advantage of automatically testing different values for the number of neighbors k and determining the optimal k with the lowest cross-validation error. In addition, the *caret* package can automatically pre-process data to normalize predictor variables. The model's dependent variable is *price_group*, and the selected independent variables that attempt to explain the characteristics of the houses in a specific group are the number of beds in the houses (*bed*), the number of bathrooms (*bath*), the area of the land (*acre_lot*), and the *zip_code* of the area where that house is located, which can provide relevant information about the state/city corresponding to the house. These factors were chosen to ensure that the model performs optimally.

According to the graph in Fig. 32.2, the optimal value for k to maximize cross-validation accuracy (approximately 90%) is around 9. However, to determine the exact value, we used an R function that returns the optimal value for the parameter k , which is actually 9. As a result, the optimal number of neighbors is 9.

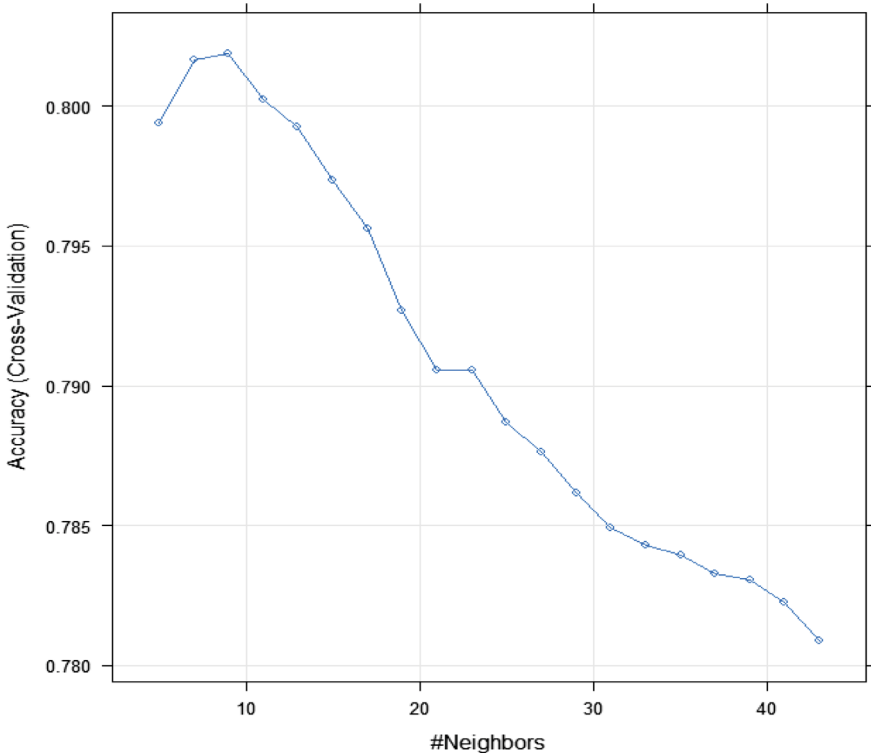


Fig. 32.2 The accuracy of the cross-validation and the best choice of the parameter k . *Source* Owner author’s processing

The confusion matrix is a relevant and widely used indicator for assessing the efficiency and performance of the built model.

Table 32.2 shows that 5.9% of observations were correctly classified as being in the “High price” class, 0.5% were erroneously classified as being in the “High price” class when they were actually in the “Low price” class, and 3.3% were erroneously classified as being in the “High price” class based on the confusion matrix for the first column (the interpretations for the other values were similar). Also, the accuracy of the prediction is 80.19%.

Following that, we made predictions for the test and determined the model’s accuracy rate. The result is 0.6193, suggesting that 61.93% of the test data were correctly predicted.

Table 32.2 The confusion matrix

Reference			
Prediction	High price	Low price	Medium price
High price	5.9	0.3	2.2
Low price	0.5	49.0	7.5
Medium price	3.3	6.0	25.3

Accuracy (average): 0.8019

Source Owner author's processing

32.6 Conclusion

The k -NN algorithm is one of the most basic classification algorithms. Even with such simplicity, it can produce extremely competitive results. Even though the algorithm is designed primarily for classification problems, it can also be used for regression problems. Another advantage of using this algorithm is that it works well with data sets that contain more than two classes.

The elaborated paper aims to construct a classifier using the k -nearest neighbors model and apply it to the US real estate market. The model categorizes houses into three price groups. The model's performance is average, with an accuracy rate of about 62%. Another conclusion to be drawn from the analysis is that the data set contains heterogeneity.

As future research directions, the article can be expanded with the use of more efficient and complex models for estimating real estate market prices or solving classification problems [15], or by improving the used k -NN model using a different Bayesian approach based on a probability model. Given the heterogeneity of the data used to develop the case study, the initial data set can be clustered, and the analysis will be performed on homogeneous subsets.

Regarding the limitations that influenced the paper's development, they consider the degree of data heterogeneity. Given the heterogeneity of the data used to develop the case study, the initial data set can be clustered, and the analysis will be performed on homogeneous subsets.

References

1. Zhan, C., Liu, Y., Wu, Z., Zhao, M., Chow, W.S.T.: A hybrid machine learning framework for forecasting house price. *Expert Syst. Appl.* **233**, 120981 (2023). <https://doi.org/10.1016/j.eswa.2023.120981>
2. Memiş, S., Enginoğlu, S., Erkan, U.: Fuzzy parameterized fuzzy soft k -nearest neighbor classifier. *Neurocomputing* **500**, 351–378 (2022). <https://doi.org/10.1016/j.neucom.2022.05.041>
3. Haq, I., Khan, D.M., Hamraz, M., Iqbal, N., Ali, A., Khan, Z.: Optimal- k nearest neighbours based ensemble for classification and feature selection in chemometrics data. *Chemom. Intell. Lab. Syst.* **240**, 104882 (2023). <https://doi.org/10.1016/j.chemolab.2023.104882>

4. Rogers, M.P., DeSantis, A.J., Janjua, H., Kuo, C.P.: The present and future state of machine learning for predictive analytics in surgery. *Am. J. Surg.* **221**(6), 1298–1299 (2021). <https://doi.org/10.1016/j.amjsurg.2020.11.023>
5. Truong, Q., Nguyen, M., Dang, H., Mei, B.: Housing price prediction via improved machine learning techniques. *Procedia Comput. Sci.* **174**, 433–442 (2020). <https://doi.org/10.1016/j.procs.2020.06.111>
6. Choy, H.T.L., Ho, W.K.O.: The use of machine learning in real estate research. *Land* **12**(4), 740 (2023). <https://doi.org/10.3390/land12040740>
7. Zamri, N., Pairan, A.M., Azman, W.N.A.W., Abas, S.S., Abdullah, L., Naim, S., Tarmudi, Z., Gao, M.: River quality classification using different distances in k-nearest neighbors algorithm. *Procedia Comput. Sci.* **204**, 180–186 (2022). <https://doi.org/10.1016/j.procs.2022.08.022>
8. Zheng, Y.: Community resilience and house prices: a machine learning approach. *Financ. Res. Lett.* **58**(Part B), 104400 (2023). <https://doi.org/10.1016/j.frl.2023.104400>
9. Yoon, J.W., Friel, W.: Efficient model selection for probabilistic k nearest neighbour classification. *Neurocomputing* **149**(Part B), 1098–1108 (2015). <https://doi.org/10.1016/j.neucom.2014.07.023>
10. K-NN Homepage, https://www.cs.cornell.edu/courses/cs4780/2017sp/lectures/lecturenote02_kNN.html. Accessed 27 Mar 2024
11. Peterson, L.E.: K-nearest neighbor. *Scholarpedia* **4**(2), 1883. http://scholarpedia.org/article/K-nearest_neighbor (2009)
12. Ali, N., Neagu, D., Trundle, P.: Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets. *SN Appl. Sci.* **1**, 1559 (2019). <https://doi.org/10.1007/s42452-019-1356-9>
13. Uddin, S., Haque, I., Lu, H., et al.: Comparative performance analysis of k-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Sci. Rep.* **12**, 6256 (2022). <https://doi.org/10.1038/s41598-022-10358-x>
14. Thomas, C., Peter, H.: Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13**(1), 21–27. <https://ieeexplore.ieee.org/document/1053964> (1967)
15. Forsys, I.: Machine learning in house price analysis: regression models versus neural networks. *Procedia Comput. Sci.* **207**, 435–445 (2022). <https://doi.org/10.1016/j.procs.2022.09.078>

Chapter 33

Viability of Open Source Assistants for Adaptive Learning Systems



Bobocea Andrei  and Corina Marina Mirea 

Abstract The paper aims to test the viability of a small number of open-source Large Language Models for the role of a live learning assistant integrated into an adaptive learning system. Several models were selected for testing which are based on variants of Meta's Llama and were fine-tuned by Open-Assistant with high-quality human feedback data. The models were tested against a previous-generation LLM (a variant of GPT2) and a current generation LLM (GPT3.5-Turbo). The quality of the generated text was analyzed, as well as the general performance and running cost. The open-source models were found to be acceptable and had the added advantages of increased control (over a commercial solution) and being available for an on-prem installation, with the notable caveat that they had significantly higher running costs.

33.1 Adaptive Learning Systems for Modern Education

33.1.1 Relevance of Adaptive Learning

Adaptive learning aims to address the individual needs of students by customizing their learning experience, leveraging artificial intelligence and other computer algorithms [1]. Three separate approaches to adaptive learning have been identified: individualization, differentiation, and personalization. Each one encompasses the previous approach and expands its scope, with personalization being the most extensive [2].

Artificial intelligence plays an important role in personalizing and adapting learning to the needs of students. It can be implemented through intelligent learning assistants, which provide learners with content that is tailored to their individual

B. Andrei (✉) · C. M. Mirea
Bucharest Academy of Economic Studies, Bucharest, Romania
e-mail: andrei.bobocea@csie.ase.ro

C. M. Mirea
e-mail: corina.mirea@csie.ase.ro

educational needs. These assistants use data from learning style analysis to make sure that the content is relevant and engaging for each student. In addition, intelligent learning assistants monitor the learning process of learners and provide them with feedback [3].

Adaptive learning systems that use artificial intelligence as their main technology are a powerful and appropriate tool for addressing the shortcomings faced by those involved in the learning process, but also for increasing interest and involvement in learning. These learning systems are able to recognize areas and subjects where a person has deficiencies in order to offer recommendations for content and learning methods that are individually personalized. These culminate in enhanced learning outcomes for students, fostering sustained motivation for knowledge acquisition and continued progress [4].

Personalization and personalized learning are recognized as a necessity in education [5] at an administrative level and they are also of significant economic interest due to the value and growth of the adaptive learning market.

A study released in 2020 by [6] found that the adaptive learning market was valued at 1.9 billion USD globally and it was expected to grow to 5.3 billion by 2025. Due to the subsequent popularity of OpenAI's GPT-3 model, it is reasonable to consider the possibility that the report's projections may be exceeded considerably, on account of the interconnected nature of adaptive learning and artificial intelligence.

33.1.2 Learning Assistants in Adaptive Learning Systems and the Role of LLMs

As adaptive learning becomes more widespread we can observe a paradigm shift in the e-learning area, from e-learning tools to e-learning assistants [7]. Learning assistants are an integral part of an adaptive learning system's architecture [8] and the primary way that the learner interacts with the system and the learning resources at his disposal.

The learning assistant should guide the learner and assist in the learning process. Traditionally learning assistants have an associated knowledge base which contains the answers to common learner queries for the particular lesson, additionally the assistant should be able to provide the learner with open educational resources throughout the lesson, keeping the learner engaged.

Knowledge bases require a considerable amount of time and effort to make and are generally built over time by analyzing the interaction between learners and the learning assistant and filling in the gaps in the assistant's knowledge. This approach is reactive rather than proactive and has the downside of leaving the initial learners with unanswered questions. For this purpose generative AI, particularly Large Language Models, can be used to temporarily bridge the gap, generating answers for the learner's questions which can then be integrated into the knowledge base if deemed appropriate. LLMs have the advantage of being able to process significant amounts

of text such as textbooks and lesson plans to better contextualize their responses. Together with other techniques such as prompt engineering and embeddings-based search, this can minimize the potential for hallucinations or the generation of factually incorrect information.

LLMs and other types of generative AI can be influenced through their training data and by manipulating certain parameters or further filtering input and output. When it comes to education and learning this can be impactful and as such it has generated concerns over the inherent biases of the LLMs that are currently available on-demand through their APIs. Many actors in the educational sector, particularly government entities, prefer solutions that let them retain control over the process and the data, both training data and generated data. On-premise installations are of particular interest and while many popular LLMs such as GPT3 are not available for on-premise installation at this time there are open-source models which have shown significant potential. Further, we will explore a couple of different models based on Meta's Llama LLM, fine-tuned by Open-Assistant [9] to analyze their viability for this role and compare them to a previous-generation LLM such as a variant of GPT2 and current generation LLMs such as GPT3.5-Turbo.

33.2 Comparative Analysis of LLMs for Integration in Learning Assistants

33.2.1 Testing Methodology

A set of 70 common questions was put together to test the ability of the selected models to understand the question and generate a valid answer. For this test, no code completion prompts were used, although some of the models are capable of generating code snippets. The questions were meant to cover a wide range of common classroom subjects. The overall structure was as follows:

- 20 common questions for a physics class
- 10 common questions for a biology class
- 10 common questions for a generic IT class
- 10 common questions for a web development class
- 10 common questions for a 5th-grade class
- 10 common questions for a geography class.

The questions were sent as prompts to each LLM, one by one, and each individual result was recorded, batching was not used for this testing. All testing was done programmatically via API calls to minimize the potential for human error or other delays. The test did not explicitly use multi-GPU inference frameworks or batching, as such it is expected that multi-GPU hardware configurations will have an unmitigated performance overhead. The use of multi-GPU inference and batching can provide significant performance improvements for real-world applications [10, 11].

5 separate models were tested, of which 4 freely available on the HuggingFace community Model database, and the 5th was GPT3.5-Turbo as a reference point. GPT3.5-Turbo is available through OpenAI's API, all other models were deployed on dedicated Inference Endpoints from HuggingFace. The Inference Endpoints are in essence dedicated deployments of the model on selected hardware. The test framework was made in Node.JS version 16.19.0 and network latency at the time of testing was on average 110 ms to the datacenter hosting the Inference Endpoints and 36 ms to the OpenAI servers.

Two main values were considered when comparing the LLM's performance:

- response time—average response time per prompt, in seconds
- running cost—cost per 1000 questions, in USD.

Response times are directly comparable, however the running costs are obtained differently between GPT3.5-Turbo and the other models deployed on dedicated infrastructure. For GPT3.5-Turbo the cost was calculated based on the token usage (both input and output) reported by the API and the API pricing. For the other models, which were deployed on dedicated inference endpoints, the cost was calculated based on the response time of the endpoint and the per-hour pricing of the endpoint's hardware configuration. Some of the models were tested on multiple hardware configurations to gauge the best value proposition.

33.2.2 Tested Models and Comparison Data

GPT-2 XL

GPT-2 XL [12] is a 1.5B parameter version of GPT-2 a precursor to GPT-3. The model produced valid answers to most of the questions; however, it should be noted that the answers are sometimes circular in nature, often lack complexity and depth, and if the model doesn't know the answer (in the case of a more niche subject) it will hallucinate instead. This is in-line with the LLM's accompanying model card which indicates that the model does not distinguish between fact and fiction and generating text that is required to be true is an out-of-scope use case [13].

Despite the model's lack of accuracy, it is still an interesting benchmark and a good representation of previous-generation LLMs with the added benefit of being fairly easy to run even on modest consumer-grade hardware. GPT2-XL was tested using the minimal hardware configuration available on HuggingFace.

Open-Assistant CodeLlama 13B SFT v10

Open-Assistant CodeLLama 13B SFT v10 [14] is one of the models fine-tuned by Open-Assistant. It is based on Meta's CodeLlama 13B LLM and has been fine-tuned via the Megatron-LLM library. The model's main use case is to generate code; however, it also performs reasonably well with other queries. The model is

significantly larger and cannot run on the same minimal configuration as GPT2-XL, it has been tested on multiple hardware configurations as shown in Table 33.1.

Performance differences were minimal ranging from 9.38 to 11.10 s per request. The fastest hardware configuration according to our testing was the single Nvidia A10G GPU. Without using batching the model did not seem to benefit from extra memory and the extra overhead of having 4 GPUs increased overall response time in the case of the $4 \times$ A10G configuration, using the $4 \times$ T4 (Nvidia Tesla T4) configuration further decreased performance for individual requests as the T4 is slower than an A10G. The hardware configurations are detailed in Table 33.1.

Open-Assistant Llama2 70B SFT v10

Open-Assistant Llama2 70B SFT v10 [15] is a fine-tuning of Meta’s Llama2 70B. Much like the previous model (Open-Assistant CodeLLama 13B SFT v10) it has been trained via the Megatron-LLM library, by Open-Assistant; however, it is considerably more generic, having been trained with both synthetic instructions and coding tasks, and it is considerably larger.

Due to its size, the model could only be tested on the $4 \times$ A10G hardware configuration, other eligible but unavailable configurations would have to include the Nvidia A100 GPU. The hardware configurations are detailed in Table 33.1.

Open-Assistant Llama2 13B ORCA-8K-3319

Open-Assistant Llama2 13B ORCA-8K-3319 is a fine-tuning of Meta’s Llama2 13B, with an 8K context size. It has been fine-tuned by Open-Assistant using a long-conversation variant of the Dolphin (orca-chat) dataset [16].

The model is significantly smaller than the previous Open-Assistant Llama2 70B model; however, it could only be run on 2 of the available hardware configurations containing 4 GPUs: 4 Tesla T4s, totaling 64 GB of VRAM, and 4 A10Gs, totaling 96 GB of VRAM. The hardware configurations are detailed in Table 33.1.

There were minimal to no differences in the generated text in between runs of the same model. Marginal differences in the quality and depth of the generated text were observed within the same family of models fine-tuned by Open-Assistant

Table 33.1 Hardware configurations used in testing the LLMs

Model	GPU (Nvidia)	VRAM	CPU	RAM	Cost (\$/hour)
GPT2XL	$1 \times$ T4	16 GB	$3 \times$ vCPU	14 GB	0.6
OA-CodeLlama	$4 \times$ T4	64 GB	$46 \times$ vCPU	180 GB	4.5
OA-CodeLlama	$1 \times$ A10G	24 GB	$6 \times$ vCPU	28 GB	1.3
OA-CodeLlama	$4 \times$ T4	64 GB	$46 \times$ vCPU	175 GB	7
OA-Llama70B	$4 \times$ A10G	96 GB	$46 \times$ vCPU	175 GB	7
OA-Llama13B	$4 \times$ T4	64 GB	$46 \times$ vCPU	180 GB	4.5
OA-Llama13B	$4 \times$ A10G	96 GB	$46 \times$ vCPU	175 GB	7

and considerable differences were noted between GPT2-XL, Open-Assistant’s fine-tuned models, and GPT3.5-Turbo. While generic questions, such as defining the role of a database management system or the definition of photosynthesis, led to correct (although shallow) responses from the GPT2-XL model, other more complex questions led to circular answers, such as asking the difference between LAN and WAN and niche questions such as asking what is SCORM (in reference to the “Sharable Content Object Reference Model”) led it to hallucinate and generate factually incorrect information.

The other tested models were able to generate significantly better results for the same prompts, both for generic questions and for more niche topics. A few relevant snippets of the generated answers can be seen in Tables 33.2 and 33.3 below, showing the difference in depth and complexity of the answers generated by the models.

Due to GPT2-XL’s inaccuracies, it does not seem suitable for the role of learning assistant within an adaptive learning system. While it is useful for other tasks, such

Table 33.2 Results generated with a simple prompt

Prompt	What is photosynthesis and how does it work?
GPT2-XL	Photosynthesis is the process by which plants use sunlight to convert carbon dioxide (CO ₂) into sugars. The process is called photosynthesis because plants use light to convert carbon dioxide into sugars. The process of photosynthesis is a complex process that involves many different chemical reactions...
OA-CodeLlama	Photosynthesis is a process by which plants convert light energy into chemical energy. It occurs in the chloroplasts of plant cells. The process involves the following steps: (1) Light energy is absorbed by chlorophyll molecules in the chloroplasts. (2) The absorbed light energy is used to split water molecules into hydrogen and oxygen molecules. (3) The hydrogen and oxygen molecules are...
OA-Llama70B	Photosynthesis is the process by which plants, algae, and some bacteria convert light energy into chemical energy in the form of glucose or other sugars. This process occurs in specialized organelles called chloroplasts, which are found in the cells of photosynthetic organisms. During photosynthesis, plants absorb carbon dioxide (CO ₂) from the air and water (H ₂ O) from the soil...
OA-Llama13B	Photosynthesis is the process by which plants convert light energy into chemical energy. This process occurs in the chloroplasts of plant cells. Photosynthesis is the process by which plants convert light energy into chemical energy. This process occurs in the chloroplasts of plant cells. Photosynthesis is the process by which plants convert light energy into chemical energy. This process occurs in the chloroplasts of plant cells. Photosynthesis...
GPT3.5-Turbo	Photosynthesis is the process by which green plants, algae, and some bacteria convert light energy, usually from the sun, into chemical energy stored in glucose. This process occurs in the chloroplasts of plant cells. The overall reaction for photosynthesis can be summarized as: $6\text{CO}_2 + 6\text{H}_2\text{O} + \text{light energy} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$ (Carbon dioxide + water + light energy → glucose + oxygen) Photosynthesis involves two main stages: ...

Table 33.3 Results generated with a more complex prompt

Prompt	What is the difference between a LAN and a WAN?
GPT2-XL	LANs are the networks that are used to connect computers to the Internet. WANs are the networks that are used to connect computers to the Internet. What is the difference between a WAN and a LAN?...
OA-CodeLlama	A LAN (Local Area Network) is a network of computers and devices that are connected within a limited geographical area, such as a building or a campus. It is typically used for communication between devices within the same local area. A WAN (Wide Area Network) is a network that connects multiple LANs together...
OA-Llama70B	A Local Area Network (LAN) is a network of computers and other devices that are located within a small area, such as a single building or a group of buildings. A LAN allows devices to communicate with each other and share resources, such as printers and files, over a wired or wireless connection...
OA-Llama13B	A LAN is a Local Area Network, which is a network that is limited to a small geographical area, such as a building or a group of buildings. A WAN is a Wide Area Network, which is a network that spans a larger geographical area, such as a city, state, or country....
GPT3.5-Turbo	A LAN (Local Area Network) is a network that connects devices within a limited area, such as a home, office, or campus. It typically covers a smaller geographic area and is often owned, controlled, and managed by a single organization, such as a business or institution. On the other hand, a WAN (Wide Area Network) is a network that spans over a larger geographical area, connecting multiple LANs and other networks together...

as content creation or section summarization [17], it requires too much supervision to be used in a classroom setting on a broader scale.

The answers provided by the Open-Assistant fine-tuned models are overall similar to each other and comparable to the answers provided by GPT3.5-Turbo; however, there are significant differences in terms of performance and subsequent running cost between the Open-Assistant fine-tuned models themselves.

The largest Open-Assistant model tested, OA-Llama70B, had the most significant response time at 18.13 s per prompt (see Fig. 33.1) and as such the highest running cost at \$35.26 per 1000 questions (see Fig. 33.2). The other Open-Assistant models, OA-CodeLlama and OA-Llama2-13B had comparable response times at 9.38 s per request and 9.48 s, respectively (see Fig. 33.1). Due to the size differences there were cost differences between them (see Fig. 33.2) that did not translate to performance improvements or significant changes in response quality with the cost per 1000 questions being \$3.38 for OA-CodeLlama (running on it’s most efficient configuration, a single A10G GPU) and \$12.94 for OA- Llama2-13B running on 4 Tesla T4 GPUs (the most efficient configuration available that satisfied the memory requirements).

GPT3.5-Turbo was significantly faster than all Open-Assistant models, it was comparable with GPT2-XL in terms of speed (see Fig. 33.1) and cost (see Fig. 33.2) without sacrificing accuracy and response quality, as shown in Tables 33.2 and 33.3. Of the Open-Assistant models tested the most efficient was OA-CodeLlama,

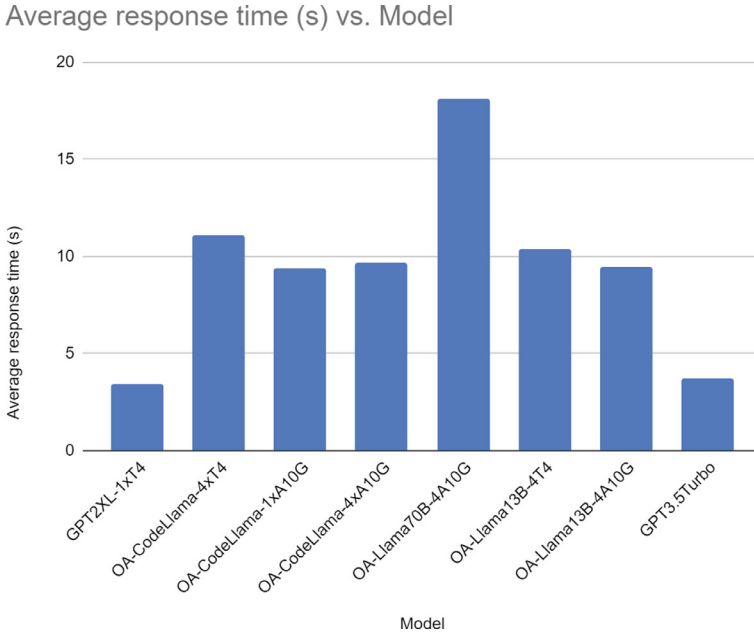


Fig. 33.1 Average response time per combination of LLM-Hardware

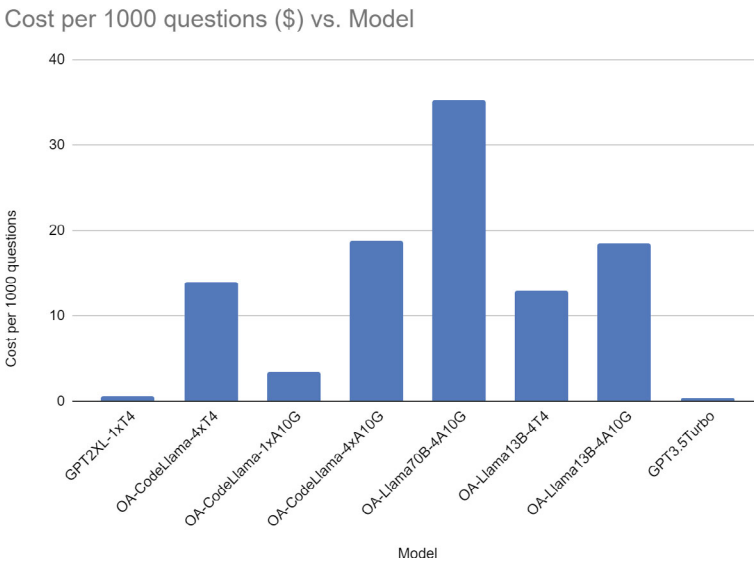


Fig. 33.2 Average cost per 1000 questions (prompts) per combination of LLM-Hardware expressed in USD

offering the best response time and the lowest running cost due to its lower hardware requirements.

33.2.3 Conclusions and Avenues of Further Research

GPT2-XL does not fit the role of a live learning assistant; however, the models fine-tuned by Open-Assistant can fit this role and have comparable results to commercially available LLMs such as GPT3.5-Turbo. In the envisioned use case, of a learning assistant providing answers to classroom questions through an adaptive learning system, the differences between the tested Open-Assistant models were largely down to performance and cost. In our testing, Open-Assistant CodeLlama 13B SFT v10 was the most cost-effective and generated answers of reasonable quality, comparable to GPT3.5-Turbo. The Open-Assistant models were; however, less cost-effective than using GPT3.5-Turbo directly.

Open-source models are of particular interest due to the extra control they offer over a commercially available model such as GPT3.5-Turbo. The extra control helps to alleviate concerns regarding the inherent biases of the model, and the source and nature of the training data and is in-line with the preference of many public sector entities for on-premise installations.

The performance of the open-source models that were tested could further be improved by fine-tuning the models or making cut-down versions that are specialized in particular areas of interest, narrowing their overall knowledge without sacrificing response quality in the targeted areas. This would reduce their size, allowing them to be run on more modest hardware. Further improvements can be obtained in real-world scenarios through the use of inference frameworks that take advantage of multi-GPU architectures and use batched inference.

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Kaplan, A.: Higher education at the crossroads of disruption: the university of the 21st century. Emerald Publishing Limited (2021). <https://doi.org/10.1108/9781800715011>. Last accessed 09 Mar 2024
2. U.S. Department of Education: Transforming american education: learning powered by technology (PDF). Department of Education. <https://www.ed.gov/sites/default/files/netp2010.pdf> (2010). Last accessed 23 Mar 2024
3. Ramirez, E.A.B., Esparrell, J.A.F.: Artificial intelligence (AI) in education: unlocking the perfect synergy for learning. *Educ. Process. Int. J.* **13**(1), 35–51 (2024). <https://doi.org/10.22521/edupij.2024.131.3>. Last accessed 20 Mar 2024
4. Adeleye, O.O., Eden, C.A., Adeniyi, I.S.: Innovative teaching methodologies in the era of artificial intelligence: a review of inclusive educational practices. *World J. Adv. Eng. Technol.*

- Sci. **11**(01), 069–079. <https://doi.org/10.30574/wjaets.2024.11.2.0091>. Last accessed 09 Mar 2024
5. Duncan, A.: Enabling the future of learning. Obamawhitehouse. <https://obamawhitehouse.archives.gov/blog/2013/12/17/enabling-future-learning> (2013). Last accessed 09 Mar 2024
 6. Markets and markets: Adaptive learning market growth drivers & opportunities. MarketsandMarkets. <https://www.marketsandmarkets.com/Market-Reports/adaptive-learning-market-257528889.html> (2020). Last accessed 09 Mar 2024
 7. Jantke, K. P., Igel, C., Sturm, R.: From e-learning tools to assistants by learner modelling and adaptive behavior. In: Intelligent User Interfaces: Adaptation and Personalization Systems and Technologies. Information Science Reference, p. 13 (2009). <https://doi.org/10.4018/978-1-60566-032-5.ch015>. Last accessed 09 Mar 2024
 8. Sajja, R., Sermet, Y., Cikmaz, M., Cwiertyny, D., Demir, I.: Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education (2019). <https://doi.org/10.48550/arXiv.2309.10892>. Last accessed 20 Mar 2024
 9. Kopf, A., Kilcher, Y., von Rutte, D., Anagnostidis, S., Tam, Z.-R., Stevens, K., Barhoum, A., Duc, N.M., Stanley, O., Nagyfi, R., ES, S., Suri, S., Glushkov, D., Dantaluri, A., Maguire, A., Schuhmann, C., Nguyen, H., Mattick, A.: OpenAssistant conversations-democratizing large language model alignment (2024). <https://doi.org/10.48550/arXiv.2304.07327>. Last accessed 14 Mar 2024
 10. Huggingface.co: Distributed inference with multiple GPUs. Hugging Face. https://huggingface.co/docs/diffusers/training/distributed_inference (n.d.). Last accessed 09 Mar 2024
 11. Medium.com: LLM Inference on multiple GPUs with 🚀 Accelerate | by Geronimo. Medium. <https://medium.com/@geronimo7/llms-multi-gpu-inference-with-accelerate-5a8333e4c5db> (2023). Last accessed 09 Mar 2024
 12. Openai: Openai-community/gpt2-xl Hugging Face. Hugging Face. <https://huggingface.co/openai-community/gpt2-xl> (n.d.). Last accessed 09 Mar 2024
 13. Openai: GPT-2 Model Card. https://github.com/openai/gpt-2/blob/master/model_card.md (2019). Last accessed 09 Mar 2024
 14. Open-Assistant: OpenAssistant/codellama-13b-oasst-sft-v10. Hugging Face. <https://huggingface.co/OpenAssistant/codellama-13b-oasst-sft-v10> (2023). Last accessed 20 Mar 2024
 15. Open-Assistant: OpenAssistant/llama2-70b-oasst-sft-v10 · Hugging Face. Hugging Face. <https://huggingface.co/OpenAssistant/llama2-70b-oasst-sft-v10#model-details> (2023). Last accessed 20 Mar 2024
 16. Open-Assistant: OpenAssistant/llama2-13b-orca-8k-3319 · Hugging Face. Hugging Face. <https://huggingface.co/OpenAssistant/llama2-13b-orca-8k-3319> (2023). Last accessed 16 Mar 2024
 17. Diwan, C., Srinivasa, S., Suri, G., Agarwal, S., Ram, P.: AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Comput. Educ. Artif. Intell.* **4** (2023). <https://doi.org/10.1016/j.caeai.2022.100110>. Last accessed 09 Mar 2024

Chapter 34

Different Entropy Measurements in Machine Learning



Dror Ben-Ami

Abstract Entropy is a very wide-meaning scientific term. This article specifically associates “entropy” with the state of disorder, randomness, chaos, and uncertainty in data science analysis perspectives. More specifically, machine learning (ML) and data mining (DM) mathematical models are the environments in which entropy would be related and referenced in its contexts. What are the key roles of entropy in ML and DM? How does it affect results and interpretation procedures? Are the currently used models considering the entropy and aware of major mathematical faults, in incorrect or improper entropy prediction? Lastly, what are the implications in such a case? The article is based on a few years of research, for a Ph.D. thesis, theoretical and practical research field. The article suggests practical implemental software-based apparatus to increase precision, accuracy, and possibly better math-models results after performing some math-models analysis. It offers a simple pre-process stage which decreases the possible faults of wrong entropy pre-process use.

34.1 Introduction—Entropy Variations

34.1.1 Entropy in General

Entropy has various scientific definitions, but all relate to the same meaning: data behavior. The definitions and their respective references cover various aspects of entropy, ranging from thermodynamics and statistical mechanics to information theory and chaos theory. In thermodynamics, entropy is a measure of the disorder or randomness of a system. It quantifies the number of possible microscopic configurations that correspond to the macroscopic state of a system. This definition is closely related to the second law of thermodynamics, which states that the entropy of an

D. Ben-Ami (✉)
Moldova State University (USM), Chisinau, Moldova
e-mail: dbenami84@gmail.com

isolated system tends to increase over time [1]. In Information theory and informatics study, entropy measures the uncertainty or information content of a random variable. It quantifies the average amount of information produced by a stochastic data source. High entropy means high unpredictability or information content, while low entropy implies more predictability or redundancy in the data [2, 3]. Statistical mechanics defines entropy as a measure of the number of possible microscopic states of a system consistent with its macroscopic properties. It is closely related to thermodynamic entropy but is derived from the statistical behavior of many particles composing the system [4]. In 1948, Claude Shannon (1916–2001), published the paper “A Mathematical Theory of Communication” [5], which first time marked the phrase “information theory”. Shannon referred to entropy as a concept from information theory that quantifies the amount of uncertainty or information content in a message or data stream. It is widely used in various fields such as communication theory, cryptography, and data compression [6]. In dynamical systems theory, the Kolmogorov-Sinai entropy measures the rate of information production in a chaotic system. It characterizes the exponential growth of uncertainty in the system’s evolution, particularly relevant in studying chaotic behavior and complexity [7].

34.1.2 Entropy in Data Science

Entropy plays a crucial role in data science and machine learning, particularly in decision trees and information gain calculations. A comprehensive explanation of how entropy is used in decision trees to measure the impurity of a node is given in the article “Understanding Entropy in Decision Trees for Machine Learning” [8]. It covers entropy calculations, information gain, and the practical implications of entropy in improving decision tree models. Another article explores using entropy-based techniques for feature selection in machine learning. It discusses methods to quantify feature importance using entropy, leading to more efficient and accurate machine-learning models [9]. A conference paper introduces entropy regularization as a technique to improve the generalization and robustness of neural networks. It discusses how incorporating entropy into the loss function can lead to better model performance, especially in scenarios with limited training data [10]. “Entropy-Based Clustering Algorithms for High-Dimensional Data” [11] is another article, that explores the application of entropy-based clustering algorithms, such as information-theoretic clustering, for handling high-dimensional data in machine learning tasks. It discusses the advantages of using entropy measures in clustering analyses. Another reference investigates the use of mutual information and entropy to quantify uncertainty in machine learning predictions. It discusses how entropy-based measures can be leveraged to assess model reliability and make informed decisions based on uncertainty estimates [12]. These articles delve into various aspects of entropy in machine learning, including its role in decision trees, feature selection, neural network regularization, clustering, and uncertainty quantification. They provide valuable insights and methodologies for leveraging entropy to improve the performance and interpretability

of machine learning models. Another reference is about genetic algorithms, in which entropy is also a key factor [13]. More detailed aspects, extended entropy exploration in machine learning relates to the concept of diversity or randomness in searching or optimizing algorithms to prevent them from getting stuck in local optima and to ensure a thorough search of the solution space. Some ML entropy key aspects are mentioned below. In the early eighties, Kirkpatrick [14] related to the Diversity in Search Algorithms: Entropy ensures diversity in the search algorithms, enabling them to explore various regions of the solution space rather than focusing on a narrow area. This is crucial in problems like combinatorial optimization, where the solution space is vast and complex. One significant mathematical consideration, especially in ML is the issue of Local Optima. So, how to avoid Local Optima? High entropy in the exploration process helps the algorithm avoid getting trapped in local optima, thus improving the chances of finding the global optimum [15]. Proceedings of the 39th International Conference on Machine Learning [16] show that locally optimal symmetric strategy profiles are also global Nash equilibrium [17]. The result is significant for machine learning, particularly for gradient methods finding local optima in symmetric strategy spaces. It addresses stability under unilateral deviations and explores instability in mixed local optima under joint, asymmetric deviations, with implications for *multi-agent reinforcement learning* (RL), and *cooperative inverse RL*. Another reference addresses the local optima problem in meta-reinforcement learning (meta-RL) by decoupling exploration and exploitation objectives. It introduces an approach that improves task performance by focusing on task-relevant information, avoiding the local optima issues common in end-to-end training of meta-RL systems. The approach led to substantial performance improvements on complex meta-RL tasks [18]. Another dimension, in the context of ML, is Balancing Exploration and Exploitation: Entropy helps in balancing between exploration (searching new areas) and exploitation (refining known good solutions). This balance is critical for the efficiency and effectiveness of learning algorithms, especially in reinforcement learning. Reinforcement learning is related directly to ML and also to Markov Chains. The last reference is about the key role of entropy in deep learning, entropy can be used as a “regularizer” to encourage the model to learn more robust features and prevent overfitting [19].

34.1.3 Markov Chains Entropy

The concept of entropy in Markov chains is crucial in understanding the randomness and predictability of their states. In ML, particularly in Monte Carlo methods, entropy helps gauge the efficiency of sampling methods and the complexity of the distributions being sampled. For example, high entropy in a Markov chain indicates a high degree of randomness and unpredictability, which can affect the convergence and efficiency of Monte Carlo simulations. Sun et al. [20] describe how the concept of entropy in a Monte Carlo method development, efficiently handles multimodal distributions. By using a variational approach, the method aims to optimize

the entropy across different modes of distribution, facilitating better exploration of the state space and ensuring more efficient sampling. Gu et al. [20], focused on leveraging entropy within the sampling process to improve the efficiency of Monte Carlo methods. The use of Langevin dynamics [21] and normalization flows in the proposed sampling method indicates an effort to manage the entropy in the sampling process, thus enhancing the ability to capture the underlying complexity of the target distributions.

34.1.4 Ergodic Math Entropy

Ergodic theory, a branch of mathematics that studies statistical properties of dynamical systems, provides a framework to understand how systems evolve over time and how they can be characterized by statistical measures such as entropy. In Dynamical Systems under ergodic theory, entropy is a measure of the rate of information production in a dynamical system. It quantifies the unpredictability or the chaos in the system. A system with higher entropy is more chaotic and less predictable [22]. Kolmogorov-Sinai (KS) Entropy [23] is a specific measure of ergodic theory entropy. It is defined as a measure-preserving transformation on a probability space and measures the average entropy per unit time. KS entropy is a powerful tool for studying the complexity and predictability of dynamical systems. Ergodic Entropy measures are crucial in various applications, such as understanding chaotic behavior in weather systems, predicting stock market trends, and analyzing engineering signals.

34.2 Entropy in ML—Math Background and Notations

Attached key formulas used for ML entropy detection. Based on Shannon's assumptions, entropy is expressed by the relationship between probability and heterogeneity or impurity in the mathematical form of the following equations:

$$H(X) = - \sum (p_i * \log_2 p_i) \quad (34.1)$$

$$\text{Entropy } (P) = - \sum_{i=1..n} (p_i * \log_2 p_i) \quad (34.2)$$

The *X-axis* denotes the probability of the event. $\mathbf{H}(\mathbf{x})$ indicates the heterogeneity or impurity, which is plotted on the *Y-axis*. Expanding it into decision trees, Information Gain is calculated as the *difference in entropy* before and after a dataset is split on an attribute. It indicates how much extra information is gained by introducing the attribute. The formula for this process is given below:

$$\text{Information Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in V} (|S_v|/|S|)\text{Entropy}(S_v) \tag{34.3}$$

where,

- S is the original set,
- A is the attribute the dataset is split by,
- V are the different values of attribute A ,
- S_v is the subset of S for each value v of A ,
- $|S_v|$ and $|S|$ are the **sizes** of the subsets and original set, respectively.

Getting deeper into the ergodic theory, entropy serves as a *quantitative measure of the unpredictability or randomness* produced by a dynamical system. Several formulas are used to calculate this entropy, reflecting different aspects of system behavior. One key formula presents the entropy of a measure μ concerning a partition P is given by

$$H(P) = - \sum_{p \in P} \mu(p) \log \mu(p) \tag{34.4}$$

The wide range of formulas specifies the difficulty and the problematcalness for any end-users, who have to understand the computation methods, the differences between them, and how they can imply the entire models' results.

34.3 The Problem—The Need

The scheme (Fig. 34.1) describes a general “Data to Knowledge” process. It characterizes any data science procedure, in any practical or research field area. Sometimes, it is quite clear “Which data model has to be selected and performed”. But it is NOT always obvious, especially to ‘simple’ users, with no wide math or data-science-specific knowledge and background. Wrong data model selection can cause and lead to invalid results and can affect the entire validation and verification process and procedures of implementation as a whole [24, 25].

The study aims are to support any data science users with knowledge recommendations regarding their data model selection. Pre-process simple DSS is offered as users’

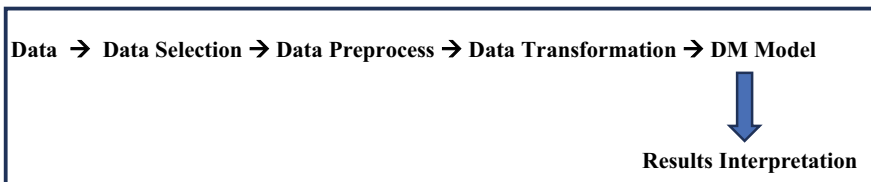


Fig. 34.1 Typical data mining model: from data into knowledge

guidance through options of “DO” and “Don’t DO” instructions, recommendations, and points of view.

The two main contributions of the study are mentioned as follows:

- (1) Minimize the improper data model selection/s
- (2) Improve research/implemental data model results and interpretation process.

The technical environment of the proposed DSS is quite simple, as the DSS is going to be supported and performed as an internet service, or as cellphone simple app-independent software, under the MS-Windows operation system. When getting into the specific data model selection issue, the main problem can start with using supervised learning models rather than using unsupervised learning models; through hierarchical or partitional models (i.e. in clustering); continuing with the model itself: neural nets, decision trees, Markov chain (or Hidden Markov Model), regression (which?), time series models and others; Crispy vs. Fuzzy models and so on. More than that, users must understand and be aware of the entire math implications and limitations of ANY data model that they can use, such as inner statistical degrees of freedom, data types (i.e. numerical or nominal parameters), data normalized approaches, wide options of ‘distance measure’ and much more. In other words, the data model selection requires experience and wide professional knowledge. Mainly mathematical-statistical traits are necessary. So, it has to be clear that mathematically wrong/incorrect model selection or not the Best-fit model selection can disrupt the ENTROPY. The entropy is one of the central key factors of the entire used model, as mentioned. Thus, an important sharp preliminary perception *BEFORE* selecting the best-fit math model relates to the perception, assimilation, and observation of the entire process as a math entity. That requires a deep understanding of the differences between stochastic math analysis models and non-stochastic math analysis models. The table below illustrates just a few parameters that must be considered [26, 27] (Table 34.1).

The facts show clearly that most of the users are not aware of all these differences, which means that “not the Best-Fit” model would be selected. Digging over the

Table 34.1 Stochastic versus non-stochastic model

	Stochastic model	Non-stochastic model
Randomness and uncertainty	High. Use probability distributions	Low, requires specific models. Mostly deterministic
Approach to modeling	The system’s behavior is uncertain and can be influenced by random variables or events	The system’s behavior is sometimes known and can be precisely described by math equations
Complexity and computation	Tend to be more complex [28]	Usually simpler in terms of computation
Usage and application area	Widely in fields like finance, economy, biology	Prevalent in engineering, physics, and areas without randomness

Table 34.2 Comparison of Markov chains versus DM models

	Reference	Markov chains and HMM	Other DM/ML models
1	Nature-essence concept	Less intuitive	More intuitive
2	Model type	Stochastic [29]	Mostly non-stochastic
3	Measurement technique	Matrix, specific formulas	i.e. Distances, similarity
4	Accuracy and precision levels	High, based on probabilities	High, based on a wide range of math formulas
5	Convergence	Stationary point	Criteria, distance, other
6	Adaptability to wide data types (text, images, voice)	Limited (principally)	Very wide
7	ML adaptability	+ Reinforcement learning [30]	+
8	Entropy checks	+	+

scheme above, the table below refers to the main common and different characteristics of the Markov Chain vs the most known ML models (Table 34.2).

34.4 The Proposed Process Solution

34.4.1 The Research Methodology Brief

The proposed solution relies on an important data source [31], about collaborative decision-making methods: from an idea and concept into an implemented working model. The proposed method and technique were examined, tested, and explored successfully in academic work environments, through different use of data files. The model has been examined as an exploratory research stage and should be expanded into much wider software environments. In addition, the next stage recommends activating an automatic apparatus to distinguish between entropy models efficiently and effectively in the future, rather than using a pre-process manual but not 100% automated model. Thus, three stages can be plotted as follows:

Manual Questionnaire → Semi-Automated Quest. → Complete Automated DSS

For the time being, a manual questionnaire and a “Semi-Automated questionnaire” are being implemented.

34.4.2 *The Research Methodology and Appropriation of Scientific Results*

As shown in the scheme above, the first stage is still semi-automated. That means that a simple short questionnaire should be filled by the user, and analyzed by a Python-based DSS (Decision Support System). The user gets some practical text indications of “what are the DSS recommendations” for the user’s case. The entire research is still under validation and final research procedures. Thus, just *five main* and major questions-indications are included and mentioned here. The key for the computational stage is not specified here, as well. All these will be included in a detailed article in a few months, after completion of those statistical analysis procedures. However first pragmatism performance indication shows clearly the essence of using the *entropy sub-process stage*, and how it contributes and affects the precision and accuracy of data analysis indexes, on the one hand, and ensures better results for the interpretation final process, on the second hand (Table 34.3).

Another important indication of the model selection considerations relates to the dataset descriptions from within the database (any dataset format) itself. The separation between the *continuous* process of selected data in the research field, rather than independent tuples which are collected as *discrete mathematical entities* is quite crucial and takes a significant role in the entire DSS recommendations.

To summarize the entire DSS concept, it is proposed to divide the solution into three main stages:

- (a) questionnaire—as shown above;
- (b) from “data” specs into “data model” DSS interactive guidance;
- (c) future option: once a model was selected, ‘bi-directional’ check: from “data to model” and from “model to data” (requires some more research).

The market is seeking for automatic indications, which can help and support users with their math-model best-fit selection [32]. Many more references can be found on the net, regarding AutoML, Data Science, and ML Platforms, through IDEs and LIBS, Consultative tools, educational tools, etc. It is obvious and not surprising at all. Most of the proposed existing guidance packs are not interactive and support some

Table 34.3 Five example questions: entropy sub-process stage

1	Supervised or unsupervised learning?
2	Some of the data was tagged or not?
3	Does the data distribute as a parameter of time? (randomness level) (no random, low, middle, high, very high)
4	How many categories do you want to generate? (2, 2 to 4, 5 and up)
5	Which accuracy level is required? (low, middle, middle plus, very good, excellent)

general instructions and hints, without specific references to the users' demands. The proposed DSS is much more user-friendly and interactive.

In addition, this article is focused on *Entropy-Analysis* specifically, without covering other points and key success factors, which affect the entire math model. Thus, the proposed DSS is planned and built upon these entropy attributes.

34.5 Conclusions

Selecting the best-fit math model is one of the challenges for each data science person. But not just for them. It is quite important to the 'simple' users, who usually have some limited knowledge in data science in general, and specifically in math models spectra. The purpose is to offer simple and accurate pre-process DSS guidance to math users. The DSS might be used by 'experts' as verification before implementing their preferred math models and also might be used and recommended for non-expert students, in their graduate studies, masters, and PhD. Students. Preliminary analysis and differentiation between stochastic and non-stochastic data analysis procedures and models can leverage and empower the users with better results in their implementations and their research fields. In addition, **entropy** was recognized as a key factor, which affect the entire model results and the interpretation stage. Continuing with that approach, the DSS can recommend specific "Do" and "Do-Not Do" clear instructions, guidance, and directions in their entire data analysis processing stages. The DSS can recognize, so far, possible use of the Markov chains and/or Hidden-Markov Models, rather than other ML models, in very high probability value. The system might be recommended in its next dev. stage some specific research areas, which are mostly best explored by specific models. The next challenge is to expand the DSS into a 100% automated best-fit model, which can do either one of these options:

- (a) Automatic wide user guidance by automatic pre-process indications
- (b) Running side by side two math models: the user's best-fit model and the DSS best-fit model, to compare their performances and success.

One important, useful, and practical conclusion concerns the ability to merge the models of Markov Chains and Decision Trees, as an example. Finally, cases and events can be identified as Markov chains or HMM—as they are identified as a series of stages, with connectivity and relationships between the stages. This is a stochastic process overview which gives a large space to the randomness from within the dataset in the process. In other words—random processes can be examined better as stochastic data analysis procedures.

Acknowledgements The researches were carried out within the doctoral project at the State University of Moldova "MATH MODELS, TECHNIQUES, AND INTELLIGENT SOFTWARE PRODUCTS, TO PROVIDE THE ORGANIZATION WITH KNOWLEDGE", doctoral supervisors: Prof. Gheorghe CĂPĂȚĂNĂ, Ph.D. Habil., Prof. Tudor BRAGARU, Ph.D. Habil.

References

1. Callen, H.B.: Thermodynamics and an introduction to thermostatistics. John Wiley & Sons (2023)
2. Cover, T.M., Thomas, J.A.: Elements of information theory. Wiley-Interscience (2024)
3. Cover, T.M., Joy, A.: Elements of information theory. Wiley-Interscience (1991)
4. Pathria, R.K., Paul, D.: Beale. Academic Press, Statistical Mechanics (2023)
5. Shannon, C.E.: A mathematical theory of communication. shannon1948.dvi, harvard.edu (1948)
6. MacKay, D.J.C.: Information theory, inference, and learning algorithms. University Press, Cambridge (2024)
7. Katok, A., Boris, H.: Introduction to the modern theory of dynamical systems. Cambridge University Press (2023)
8. Doe, J.: Understanding entropy in decision trees for machine learning. J. Mach. Learn. Res. (2023)
9. Smith, J., et al.: Entropy-based feature selection for machine learning models. IEEE Trans. Knowl. Data Eng. (2023)
10. Johnson, A.: Entropy regularization for neural network training. In: International Conference on Machine Learning (ICML) (2023)
11. Brown, R.: Entropy-based clustering algorithms for high-dimensional data. Data Mining and Knowledge Discovery (2023)
12. White, E., et al.: Quantifying uncertainty in machine learning models using mutual information and entropy. Pattern Recognit. Lett. (2023)
13. Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley (1989)
14. Kirkpatrick, S., Gelatt, C.D., et al.: Optimization by simulated annealing. Science **220**(4598), 671–680 (1983)
15. Holland, J.: Adaptation in natural and artificial systems. University of Michigan Press (1975)
16. Emmons, S., et al.: For learning in symmetric teams. Local Optima are Global Nash Equilibria. <https://proceedings.mlr.press/v162/emmons22a/emmons22a.pdf> (2022)
17. Xin, J., et al.: Game theory for infrastructure security. <https://www.sciencedirect.com/topics/computer-science/nash-equilibrium> (2012)
18. Zheran, L.E., et al.: Decoupling exploration and exploitation for meta-reinforcement learning without sacrifices. <https://arxiv.org/abs/2008.02790> (2020)
19. Goodfellow, I., Bengio, Y., et al.: Deep learning, MIT Press (2016)
20. Sun, Y., et al.: Quantum adversarial transfer learning. Entropy MDPI. <https://www.mdpi.com/1099-4300/25/7> (2023)
21. Anjaly, P., et al.: Decentralized Langevin dynamics for Bayesian learning. NeurIPS, <https://proceedings.neurips.cc/paper/2020/file/b8043b9b976639acb17b035ab8963f18-Paper.pdf> (2020)
22. Einsiedler, W.: Entropy in ergodic theory and topological dynamics, chp. 1, p.7–14, chp. 2, p. 53–64. EntropyBooks. tward0.wixsite.com (2021)
23. Keller, K., Sinn, M.: A standardizer approach to the Kolmogorov/Sinai entropy. IOP Publishing. https://www.researchgate.net/publication/48192286_A_standardized_approach_to_the_KolmogorovSinai_entropy (2009)
24. Mahmood, B.: 4 reasons your machine learning model is wrong (and How to Fix It). <https://www.kdnuggets.com/2016/12/4-reasons-machine-learning-model-wrong.html> (2016)
25. Toprak, K.: Analysis in machine learning. <https://medium.com/@ktoprakucar/a-comprehensive-guide-to-error-analysis-in-machine-learning-288e353f7c8d> (2023)
26. Khare, Y.: Deterministic vs stochastic–machine learning fundamentals–analytics vidhya (2023)
27. Modeling: Deterministic vs stochastic–choosing the right approach–adventures in machine learning. Modeling: Deterministic vs stochastic–choosing the right approach–adventures in machine learning (2023)
28. Turner, J.R., Baker, M.: Complexity theory: an overview with potential applications for the social sciences. mdpi.com (2019)

29. Robins, A.: Stochastic vs deterministic models: understand the pros and cons, (ev.uk) (2020)
30. Sutton, R.S., Barto A.G.: Reinforcement learning: an introduction. MIT Press (2018)
31. Filip, F.G., et al.: Decision-making: concepts, methods, and supporting information and communication technologies. Springer (2017)
32. The most popular machine learning tools (blog) <https://www.datacamp.com/blog/most-popular-machine-learning-tools> (2023)

Chapter 35

Algorithmic Trading Bots Using Artificial Intelligence



Florentin Șerban and Bogdan-Petru Vrînceanu

Abstract This study presents a comparative analysis of two algorithmic trading bots employing distinct artificial intelligence algorithms, namely Lorentzian classification and K-Nearest Neighbors (KNN), in the context of Bitcoin/USD trading. Over a period spanning from February to April 2024, the bots were deployed on the Pionex exchange with an initial investment of \$100 each, with profits systematically reinvested at each trade. Through meticulous examination of the code, strategy comparison, and rigorous evaluation of results, this study aims to discern which algorithm yields superior performance in terms of profitability and efficiency. While both bots demonstrate viability in generating profits, the KNN algorithm-driven bot showcases superior performance across key metrics. Its higher win rate, profit factor, and average profit per trade underscore the robustness and efficacy of the KNN algorithm in capturing profitable trading opportunities. Nonetheless, the Lorentzian classification algorithm-driven bot still exhibits commendable performance, highlighting the versatility of AI algorithms in algorithmic trading contexts. Further research and refinement of algorithmic trading strategies utilizing AI algorithms hold promise for enhancing performance and unlocking new frontiers in algorithmic trading excellence.

35.1 Introduction

Algorithmic trading, empowered by artificial intelligence (AI) algorithms, represents a cutting-edge approach to navigating the dynamic landscape of financial markets. In recent years, the fusion of AI techniques with trading strategies has revolutionized how traders analyze data, make decisions, and execute trades. These “algobots,”

F. Șerban

Faculty of Mathematics and Computer Science, University of Bucharest, Academiei Nr. 14,
010014 Bucharest, Romania

F. Șerban (✉) · B.-P. Vrînceanu

Bucharest University of Economic Studies, Piata Romana Nr. 6, 010374 Bucharest, Romania
e-mail: florentin.serban@csie.ase.ro

as they're often called, leverage the power of machine learning, deep learning, and other AI methodologies to extract valuable insights from vast amounts of market data, leading to more informed and profitable trading decisions. The allure of algobots lies in their ability to process and interpret complex market data with unparalleled speed and accuracy. Unlike traditional trading methods that rely heavily on human intuition and manual analysis, algobots can sift through terabytes of historical data, identify patterns, and forecast future market trends in real-time.

Furthermore, AI algorithms empower algobots to continuously learn and evolve from their experiences. Through a process known as "machine learning," algobots can refine their trading strategies over time, becoming increasingly adept at predicting market movements and optimizing trade execution. This adaptive nature allows algobots to stay ahead of the curve and capitalize on emerging opportunities in the market.

In this era of AI-driven trading, algobots have become indispensable tools for institutional investors, hedge funds, and individual traders alike. Their ability to automate trading processes, minimize human error, and maximize returns has transformed the way trading is conducted across global financial markets. However, with great power comes great responsibility, and the use of AI algorithms in trading also raises ethical and regulatory considerations that must be carefully navigated.

35.2 Literature Review

Automating Success: Harnessing Pine Script for Algorithmic Trading Strategies

Algorithmic trading strategies leverage computer algorithms to execute trades automatically, relying on quantitative analysis, statistical models, and computational techniques to identify profitable opportunities in financial markets. Pine Script, a specialized language on the TradingView platform, facilitates the development and implementation of these strategies by enabling traders to create custom technical analysis indicators and scripts [1].

In Pine Script, traders delve into extensive data analysis, examining historical and real-time market data, including price, volume, and various technical indicators. This analysis is crucial for uncovering patterns, trends, and potential trading signals [2].

The heart of algorithmic trading in Pine Script lies in indicator development. Traders can devise custom indicators based on mathematical calculations, price action, or other trading signals. These indicators serve as the foundation for trading strategies, generating buy and sell signals based on predefined conditions.

Strategy formulation is a pivotal aspect of Pine Script development. Traders define rules for entering and exiting trades, along with parameters for position sizing, risk management, and other factors. These rules govern how the algorithm reacts to market movements and guide its decision-making process [3].

Backtesting is essential for evaluating the effectiveness of trading strategies. Pine Script enables traders to simulate their strategies using historical data, assessing

performance metrics such as profitability, risk, and drawdown. This process allows traders to refine their strategies and optimize parameters for better results [4].

Optimization is an iterative process in Pine Script development. Traders adjust strategy parameters to maximize profitability or minimize risk based on backtesting results. Through continuous optimization, traders aim to enhance the performance and robustness of their strategies. Execution is seamlessly integrated into Pine Script, enabling automated trading based on predefined rules.

Monitoring and evaluation are ongoing tasks in algorithmic trading. Pine Script provides tools for monitoring strategy performance in real-time, allowing traders to analyze key metrics such as profit and loss, win rate, and drawdown. By closely monitoring performance, traders can identify areas for improvement and adjust their strategies accordingly [4].

In summary, Pine Script offers a comprehensive platform for developing and implementing algorithmic trading strategies. By combining quantitative analysis, technical indicators, and automation, traders can capitalize on market opportunities while mitigating risks and achieving consistent profitability. However, success in algorithmic trading requires continuous learning, adaptation to market conditions, and a disciplined approach to strategy development and execution [5]. In Pine Script, traders delve into extensive data analysis, examining historical and real-time market data, including price, volume, and various technical indicators. This analysis is crucial for uncovering patterns, trends, and potential trading signals. The heart of algorithmic trading in Pine Script lies in indicator development. Traders can devise custom indicators based on mathematical calculations, price action, or other trading signals. These indicators serve as the foundation for trading strategies, generating buy and sell signals based on predefined conditions [6]. Backtesting is essential for evaluating the effectiveness of trading strategies. Pine Script enables traders to simulate their strategies using historical data, assessing performance metrics such as profitability, risk, and drawdown. This process allows traders to refine their strategies and optimize parameters for better results [7]. Optimization is an iterative process in Pine Script development. Traders adjust strategy parameters to maximize profitability or minimize risk based on backtesting results. Through continuous optimization, traders aim to enhance the performance and robustness of their strategies [8]. In summary, Pine Script offers a comprehensive platform for developing and implementing algorithmic trading strategies. By combining quantitative analysis, technical indicators, and automation, traders can capitalize on market opportunities while mitigating risks and achieving consistent profitability. However, success in algorithmic trading requires continuous learning, adaptation to market conditions, and a disciplined approach to strategy development and execution.

Artificial Intelligence Advancements: Revolution Algorithmic Trading Strategies

Artificial intelligence (AI) has ushered in a new era of sophistication and efficiency in algorithmic trading, transforming the landscape of financial markets and empowering traders with unprecedented capabilities. At the core of this transformation lies a diverse array of AI techniques, each offering unique advantages in analyzing

market data, identifying patterns, and making predictive insights at lightning speed [9]. Deep neural networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs) are among the deep learning architectures utilized in trading applications. These models excel in analyzing sequential data, making them well-suited for time series analysis and forecasting in financial markets [8]. amounts of market data. Cloud computing platforms, distributed computing frameworks, and high-performance computing clusters enable traders to process and analyze massive datasets with unprecedented speed and scalability [10].

Blockchain, Bitcoin, and the Evolution of Algorithmic Trading

Blockchain technology and Bitcoin have catalyzed a paradigm shift in algorithmic trading, redefining market dynamics and opening up a new frontier of possibilities for traders and investors. Beyond their initial disruption of traditional finance, both blockchain and Bitcoin continue to evolve, presenting a dynamic landscape for algorithmic trading strategies to thrive and adapt [1]. Blockchain technology's decentralized and immutable ledger architecture fundamentally transforms the way financial transactions are conducted. Its transparency, security, and resistance to tampering make it an ideal infrastructure for algorithmic trading platforms seeking to enhance trust and reliability in market operations. By leveraging blockchain, algorithmic trading platforms can streamline trade settlement processes, mitigate counterparty risk, and foster greater transparency in trade execution.

Market liquidity is another critical consideration for algorithmic trading in the Bitcoin market. While Bitcoin boasts high trading volumes on major exchanges, liquidity can vary significantly across different trading pairs and platforms. Traders must carefully assess liquidity conditions and adapt their strategies accordingly to minimize execution risks [11].

Furthermore, algorithmic trading in the Bitcoin market is influenced by a myriad of factors, including macroeconomic trends, regulatory developments, technological advancements, and market sentiment. Sentiment analysis tools, natural language processing algorithms, and machine learning models enable traders to extract valuable insights from news and social media data, allowing them to make informed trading decisions.

The decentralized nature of Bitcoin and blockchain technology introduces new challenges and complexities for algorithmic traders. Network latency, data synchronization issues, and security vulnerabilities necessitate robust infrastructure and risk management protocols. Traders must also navigate evolving regulatory frameworks and compliance requirements to ensure the legality and legitimacy of their trading activities [12].

Despite these challenges, the integration of blockchain technology and Bitcoin into algorithmic trading ecosystems represents a significant opportunity for innovation and growth. As the cryptocurrency market continues to mature and institutional adoption accelerates, algorithmic trading strategies will play an increasingly pivotal role in shaping the future of digital asset trading. By harnessing the unique properties of blockchain and Bitcoin, algorithmic traders can unlock new markets, optimize

trading strategies, and capitalize on the transformative potential of decentralized finance.

35.3 Methodology

In this study, we embarked on a comprehensive exploration of algorithmic trading methodologies by developing two distinct algorithmic trading bots utilizing artificial intelligence techniques. Focusing on the Bitcoin/USD trading pair with a time frame of 45 min, our aim was to compare the efficacy of two prominent AI algorithms: Lorentzian classification and K-Nearest Neighbors (KNN). With an initial investment of \$100 in each algorithmic trading bot on the Pionex exchange, profits were systematically reinvested at each trade. Spanning from 1 February to 1 April 2024, our analysis encompassed a thorough examination of the code, strategy comparison, and rigorous evaluation of the attained results. Through meticulous scrutiny and comparative analysis, we sought to discern which algorithm yielded superior performance in terms of profitability and efficiency.

35.4 Analysis/Results Interpretation

In delving into the comparative analysis of the two algorithmic trading bots, each driven by distinct artificial intelligence algorithms, we uncover a nuanced narrative that unveils a plethora of insights into their performance dynamics and operational efficacy. At the forefront of this examination lies the K-Nearest Neighbors (KNN) algorithm-driven bot, which emerged as a formidable contender with a commendable profit margin of 13.47%. This notable achievement not only underscores the efficacy of the KNN algorithm in navigating the complexities of the Bitcoin/USD trading pair but also highlights its ability to capitalize on market opportunities with precision and agility. With a total of 61 trades executed over the testing period, the KNN bot demonstrated a robust win rate of 39.34%, indicative of its adeptness at identifying favorable trading setups and executing timely trades to capitalize on them. Furthermore, the KNN bot's profit factor of 1.687 signifies a judicious balance between profitability and risk management, with each trade contributing positively to the bot's overall performance. The maximum drawdown of 3.82% reflects the bot's resilience in mitigating downside risk, ensuring the preservation of capital even during adverse market conditions. Moreover, the average profit per trade of 0.21% and the average bars in trade of 4 further underscore the KNN bot's proficiency in extracting value from market fluctuations while maintaining a disciplined approach to trade execution.

Conversely, the Lorentzian classification algorithm-driven bot, while exhibiting commendable performance, showcased a slightly lower profit margin of 8.31% compared to its KNN counterpart. With an equal number of trades executed (61), the Lorentzian bot attained a win rate of 34.43%, indicating a slightly lower frequency

of successful trades compared to the KNN bot. Despite this, the bot maintained a respectable profit factor of 1.493, suggesting efficient capital utilization and prudent risk management practices. The maximum drawdown of 3.03% highlights the bot's ability to navigate market volatility and preserve capital effectively. However, the lower average profit per trade of 0.13% and the average bars in trade of 3 suggest a somewhat less optimal exploitation of trading opportunities compared to the KNN bot. This discrepancy may stem from differences in the underlying trading strategies employed by each bot, as well as variations in the sensitivity to market conditions and trading signals.

In synthesizing these findings, it becomes evident that while both algorithmic trading bots demonstrate viability in generating profits in the Bitcoin/USD trading environment, the KNN algorithm-driven bot exhibits superior performance metrics across several key dimensions. Its higher win rate, profit factor, average profit per trade, and average bars in trade collectively attest to the robustness and efficacy of the KNN algorithm in capturing profitable trading opportunities and optimizing trading outcomes. Nonetheless, the Lorentzian classification algorithm-driven bot still showcases commendable performance, underscoring the versatility and adaptability of different AI algorithms in algorithmic trading contexts. Moving forward, further research and refinement of algorithmic trading strategies utilizing AI algorithms will undoubtedly yield insights into enhancing performance and unlocking new frontiers in algorithmic trading excellence.

kNN Strategy Code

A sequence extracted from the code:

```
Code made by capissimo

K = input.int(63, 'K Value for kNN Model |5..n|', minval=5)
ind = input.string('All', 'Indicator', options=['RSI', 'ROC', 'CCI',
'MOM', 'All'])
fast = input.int(14, 'Fast Period |1..n|', minval=1)
slow = input.int(28, 'Slow Period |2..n|', minval=2)
ftype = input.string('Both', 'Filter Signals by', options=['Volatility', 'Volume', 'Both', 'None'])
holding_p = input.int(1, 'Holding Period |1..n|', minval=1)
tthres = input.float(99.9, 'Time Threshold |0.1..100.0|', minval=0.0, maxval=100.0, step=0.1)

//----- Global Variables

var int BUY = 1
```

The Pine Script code provided is a comprehensive implementation of a trading strategy based on the k Nearest Neighbors (kNN) algorithm. It begins by allowing

users to input various parameters such as the value of K for the kNN model, the indicator to use (RSI, ROC, CCI, MOM, or all), fast and slow periods for the indicators, filtering criteria, holding period, and time threshold. These parameters enable users to customize the strategy according to their preferences and market conditions.

The script initializes global variables such as constants for BUY, SELL, and HOLD signals, as well as counters for tracking trading activity. These variables provide a consistent reference point throughout the script and help manage the trading logic effectively. Custom functions are defined within the script to perform specific tasks such as color-coding, filtering based on volume and volatility, and scaling values to a predefined range. These functions enhance the readability and modularity of the code, making it easier to maintain and adapt to different requirements.

The main logic of the script revolves around computing predictor indicators such as RSI, ROC, CCI, and MOM for both fast and slow periods. These indicators provide valuable insights into market trends and momentum, forming the basis for predicting future price movements. The direction of the next bar (BUY, SELL, or HOLD) is determined based on the closing prices. This step serves as the foundation for training the kNN model, as it establishes a relationship between past price movements and future trends. Training data, consisting of normalized feature values and corresponding direction labels, is stored in arrays. These arrays are essential for training and evaluating the kNN algorithm, allowing it to learn from historical price data and make accurate predictions. The core logic of the kNN algorithm is then implemented, wherein distances between the current point and historic points are calculated. The k nearest neighbors are selected based on these distances, and their directions are aggregated to form an overall prediction. Once the prediction is computed, the script applies a filter based on the selected criteria (volatility, volume, both, or none) to refine the trading signals. Long and short trade signals are identified based on the prediction, time threshold, and filter conditions, dictating when to enter or exit trades. Trade execution is managed through a series of conditional statements that initiate and close trades based on the holding period and changes in the trading signal. These mechanisms ensure proper risk management and trade execution, maximizing the profitability of the strategy.

In summary, the Pine Script code provides a robust framework for implementing a trading strategy based on the kNN algorithm. It offers flexibility, customization, and scalability, allowing traders to adapt to market and achieve their financial goals.

Lorentzian Classification Strategy Code

A sequence extracted from the code:

```
Code made by jdehorty
series_from(feature_string, _close, _high, _low, _hlc3, f_paramA,
f_paramB) =>
    switch feature_string
        "RSI" => ml.n_rsi(_close, f_paramA, f_paramB)
        "WT" => ml.n_wt(_hlc3, f_paramA, f_paramB)
        "CCI" => ml.n_cci(_close, f_paramA, f_paramB)
        "ADX" => ml.n_adx(_high, _low, _close, f_paramA)

    get_lorentzian_distance(int i, int featureCount, FeatureSeries fea-
tureSeries, FeatureArrays featureArrays) =>
        switch featureCount
            5 => math.log(1+math.abs(featureSeries.f1 - array.get(fea-
tureArrays.f1, i))) +
                math.log(1+math.abs(featureSeries.f2 - array.get(fea
```

This Pine Script code embarks on a journey into the intricate realm of algorithmic trading, where the marriage of machine learning and financial markets unfolds. As the settings unveil, users are empowered to tailor their trading strategies with precision, from selecting data sources to fine-tuning feature engineering parameters. With each feature meticulously crafted, the script delves into the heart of technical analysis, extracting insights from indicators like RSI, WT, CCI, and ADX.

Pine Script code is described as a game-changer for traders. It's more than just an indicator; it blends complex algorithms with a user-friendly design to give users an advantage. This code fosters a collaborative community where traders share knowledge and refine the script together. Ultimately, it empowers traders to make data-driven decisions and reach their financial goals, transforming them into skilled and precise market navigators in today's ever-changing financial world. In essence, this Pine Script code is more than just a tool for trading—it is a catalyst for transformation, empowering traders to unlock their full potential and navigate the complexities of the financial markets with skill and precision. With its blend of cutting-edge technology and human insight, it represents a new era of trading where innovation and collaboration converge to drive success.

35.5 Conclusion

After examining the Pine Script codes and considering the discussions we've had, it's clear that both algorithmic trading bots offer unique strengths and capabilities. However, when it comes to determining which bot performed better, several factors need to be taken into account. The first Pine Script code focuses on implementing a trading strategy based on the k Nearest Neighbors (kNN) algorithm. It offers extensive customization options, allowing users to fine-tune parameters such as the value of K, indicator selection, filtering criteria, and more. The algorithm leverages historical data and technical indicators to make predictions about future price movements, using a sophisticated approach to classify market trends. In evaluating the performance of these algobots, several metrics must be considered, including profitability, risk management, stability, and adaptability to changing market conditions. While both algobots offer robust frameworks for implementing trading strategies, the effectiveness of each may vary depending on the specific requirements and preferences of traders. The determination of which algobot performed better would require rigorous testing and evaluation in real-world market conditions. Factors such as backtesting results, simulated trading performance, and live trading outcomes would need to be analyzed to assess the overall effectiveness and reliability of each algobot.

In our evaluation of the two algorithmic trading bots, it's notable that the k Nearest Neighbors (kNN) algorithm exhibited superior performance in terms of profitability. The kNN-based Pine Script code showcased a robust framework for predicting market trends by leveraging historical data and technical indicators. Its customizable parameters, including the value of K, indicator selection, and filtering criteria, allowed traders to tailor the strategy to their preferences and market conditions.

In conclusion, while both Pine Script codes demonstrate innovative approaches to algorithmic trading, the evaluation of their performance requires careful consideration of various factors and metrics. By conducting thorough testing and analysis, traders can gain valuable insights into the strengths and weaknesses of each algobot and make informed decisions to optimize their trading strategies for success.

References

1. Șerban, F., Vrînceanu, B.-P.: A comparison between trend following (Tidy Little Robot—TLR) and reversal trading (Dirty Little Robot—DLR) algobot strategies. *Eur. J. Theor. Appl. Sci.* **1**(5), 805–822 (2023). [https://doi.org/10.59324/ejtas.2023.1\(5\).68](https://doi.org/10.59324/ejtas.2023.1(5).68)
2. Aziz, A.: *Advanced techniques in day trading a practical guide to high probability strategies and methods*. Bear Bull Traders, Lexington, Kentucky (2018)
3. Chan, E.P.: *Quantitative trading : how to build your own algorithmic trading business*. John Wiley & Sons, Inc., Hoboken (2021)
4. Șerban, F., Dedu S., Tudorache, A.: *Quantitative risk management techniques using interval analysis, with applications to finance and insurance* (2014)
5. Vaida, B.: *Zero to hero in cryptocurrency trading*. Packt Publishing Ltd. (2023)
6. Hunt, E.B., Carterette, E.C., Friedman, M.P.: *Artificial Intelligence*. Elsevier Science, Burlington (2019)

7. Ross, J., Goliński, M., Ekonomia, Szymaniak, R.: Day trading. Poznań: Księgarnia Maklerska, PI (2017)
8. Leshik, E., Cralle, J.: An introduction to algorithmic trading. John Wiley & Sons (2011)
9. Cios, K.J., Kurgan, L.A., Pedrycz, W., Swiniarski, R.W.: Data mining : a knowledge discovery approach. Springer Science+Business Media, Llc., Boston, Ma (2007)
10. Michalski, R.S., Carbonell, J.G., Mitchell, T.M.: Machine learning : an artificial intelligence approach. Calif: Tioga Pub. Co., Palo Alto (2019)
11. Vigna, P., Casey, M.: The age of cryptocurrency: how bitcoin and the blockchain are challenging the global economic order. Picador/St. Martin's Press, New York, N.Y. (2016)
12. Amezcua, J., Melin, P. and Castillo, O.: New classification method based on modular neural networks with the LVQ algorithm and type-2 fuzzy logic. Springer.9 (2018)

Chapter 36

Multinomial Naïve Bayes Classifier for Romanian News in the Context of a Television Station



Adrian Vintila and Constanta-Nicoleta Bodea

Abstract News television stations are reporting current events from multiple fields such as political, social, sports and so on. Having a report that classifies the news generated by the journalists would significantly benefit the management of the news station by assisting at resource allocation, editorial content strategy, performance measurement and audience insight. This paper proposes a supervised machine learning model that can automatically classify news using text data from a Romanian branch of a multinational news network that currently does the news classification process manually. First, we collect the data in the form of Romanian-language news stories from the media asset management system of the news station. With this data, we then create and pre-process the dataset, vectorize the text data into numerical values using both CountVectorizer and TF-IDF, train the model using Multinomial Naïve Bayes algorithm and evaluate the results. The model evaluation shows good performance, the classification accuracy varied from 79 to 82%, depending on the vectorize method applied for the text feature extraction. A better accuracy was achieved when CountVectorizer method was applied. To provide additional context for the results of our news classifier application, we extended our analysis to include a Support Vector Machine (SVM) model.

A. Vintila (✉) · C.-N. Bodea
Bucharest University of Economic Studies, Bucharest, Romania
e-mail: vintilaadrian12@stud.ase.ro

C.-N. Bodea
e-mail: bodea@ase.ro

C.-N. Bodea
Centre for Industrial and Services Economics, Bucharest, Romania

36.1 Introduction

Every television station that broadcasts news segments as part of their program grid has a newsroom, the central hub where journalists and other editorial staff gather, report, and produce news content for publication or broadcast. The news stories are created, curated, and dispatched. Newsrooms are typically organized into various sections or desks, each focusing on a specific area of coverage. This structure allows for specialization, enabling journalists to develop expertise in their respective domains. As such, the news stories can be categorized by their area of coverage.

The management of a Romanian branch of a multinational news television network decided to put into effect a detailed weekly report of their newsroom content output. The report should have all the stories categorized and offer an overview of how many stories that belong to a certain category were created and broadcasted each week. To create the report, first, the news stories must be classified in their respective category. This first step required a technical solution for news labelling that was implemented using the internal Media Asset Management (MAM) system platform specific to the news station. This platform is used, among others, to create, process and refine news stories. Is here where a new data structure was defined, named “categorie_raport”, a list of strings of all the categories in which a news item can be placed. In the Graphical User Interface (GUI), this list of strings appeared as a drop-down list which led to a new workflow through which journalists are manually selecting the category, from the list, for each news story they create. The MAM stores the content using a large database so that counting news from a certain category is done with relative ease using a few SQL statements. The first phase for manual news categorization began in October 2023 with five categories and was later extended to include more categories in February 2024.

In this paper, we describe the process of gathering data from the news station and subsequently utilize this data to introduce and evaluate a supervised machine learning algorithm designed for the task of text classification, the goal being to automate the categorization of news. The classifier has been developed through training on a dataset that we constructed using manually labeled news stories between October 2023 and March 2024.

The paper has the following structure: following the introduction, this paper proceeds with a review of the existing literature on the topic. The third section presents the Methodology, focusing on the process of data collection, on how we created the dataset and the way we pre-processed the dataset. The next section provides the Results and, at the end of the paper, are included the Conclusions and Reference sections.

36.2 Literature Review

Various studies have utilized machine learning classifiers [1] to categorize news articles by using publicly available datasets or by extracting articles from news websites or platforms using web scraping tools [2–4]. Parida et al. [5] used the Reuters-21578 dataset, a collection of documents with news stories, to evaluate the news text categorization performance of two machine learning algorithms, Random Forest and Naïve Bayes. For each classifier, they used two feature extraction techniques, TF-IDF Vectorizer and Count Vectorizer. Their results show that Naive Bayes classifier performed better when paired with the Count Vectorizer using the Reuters news dataset while Random Forest output was better with TF-IDF vectorizer.

Daud et al. [6] evaluated six classifiers to categorize news stories according to their respective category, support vector machine (SVM), Stochastic Gradient Descent (SGD), Random Forest (RF), Logistic Regression (LR), K-nearest Neighbor (KNN), and Naïve Bayes (NB). They also used the publicly available Reuters news dataset that was balanced during the pre-processing phase. Their findings revealed that the optimized SVM model outperformed other models; however, without optimization, its performance was inferior to other machine learning models, with Naive Bayes ranking among the top three in terms of accuracy.

Dhar et al. [7] created a custom dataset and presented a comparative analysis of the performance of different classifiers using both CountVectorizer and TF-IDF vectorization methods. Naïve Bayes consistently performed well across different setups, often outperforming other models in terms of accuracy and F1-score. They used Tree-Based Pipeline Optimization Tool (TPOT) and identified the best configurations, leading to an improved classification accuracy, with Naïve Bayes and logistic regression identified as the best models under different feature extraction scenarios.

Barua et al. [8] explores the application of various machine learning techniques for the classification of sports news into multiple categories. This research involved the creation of a new sport news corpus.

Hossain et al. [9] compares the potential of both traditional machine learning and advanced deep learning techniques in enhancing the automatic categorization of Bengali news stories, aiming for more personalized and accessible news consumption. The study used a balanced dataset and concluded that the deep learning models outperformed baseline models.

Gopal et al. [10] developed a web crawler to collect online news articles, which were then used as data to train the classification models. Text data was processed using TF-IDF vectorization and the algorithms chosen were K-Nearest Neighbors (KNN), Multi-Layer Perceptrons (MLP), Support Vector Machines (SVM), Logistic Regression and Naïve Bayes.

36.3 Methodology

This section outlines the proposed approach, beginning with the collection of data and creation of the dataset, followed by the steps for preprocessing and vectorizing, and concluding with the classification (Fig. 36.1).

36.3.1 Collecting Data

By not using a benchmark dataset, probably the most difficult step was to isolate and then extract the news articles text data from the internal system that the television station uses to create stories. Having the news content was not sufficient as we also needed to extract the category associated with each individual story.

The text of the news stories was embedded in a story template, an element specific to the television station’s internal platform, stored in a database, that contains not only that text but also many other information as automation cues, graphic templates, media paths, metadata and so on. As such, an SQL statement was not particularly useful. The platform has the option to export a story template as an XML file, which also contains the category of that news item. Thus, we created an export rule and exported all the story templates that had a category associated, this being the only criteria for export. This resulted in a number of 4732 XML files. Unfortunately, the exported XML file format lacks a clear hierarchical organization, the structure varies throughout the document and mixes different types of content in an unpredictable manner. This led to the creation of a custom XML parser to isolate and centralize only the content that we needed – the actual news story text and the category.

The script was built using Python and it would take each XML file and isolate the news text and category using patterns in the document.

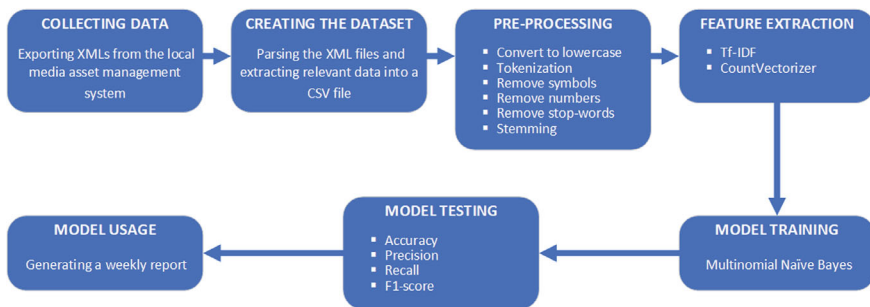


Fig. 36.1 Methodology workflow

```

pattern_intro = re.compile(r"\[INTRO\](.*?)\(", re.DOTALL)
pattern_categorie = re.compile(r'<attrName t="ws">Categorie_
Raport</attrName><value="ws">(.*?)</value></StringAssetInfo></
assets>', re.DOTALL)

```

The first line of code finds the news text and the second line finds the category for each XML. The rest of this script then creates a CSV file with three columns: index, content and category. The index counter was useful to compare the actual number of rows in the CSV with the number of XML files. After the first run, taking a glance over the CSV it was obvious that some artefacts were exported as real news. This is because someone created a news story, associated a category to that story and, after a while, moved the text somewhere else, resulting in an empty story. Also, some common tags in the world of television were present inside the content. Thus, the script was modified to clean up artefacts that were counted as stories and to remove common television tags inside the content. After running the revised script, the dataset was created, a CSV file with three columns and 3548 rows.

36.3.2 Dataset

As we mentioned previously, the manual news categorization began in October 2023 with five categories and was later extended to include eleven more categories in February 2024. We collected the data from October 2023 until March 2024, so the dataset included the newly added categories. Since the new categories had very few stories, we decided to trim the dataset to include only the original five categories and, as the dataset grows, use all the categories in future work. A python script was created to remove the newly added categories and remove the index column. The resulting dataset had 2711 rows and two columns: content and category, all in Romanian language.

```

[evaluate ds_trimmer_script_trim_index.py]
Original dataset shape: (3548, 3)
Unique categories before filtering: 16
Dataset shape after removing specified categories: (2711, 3)
Dropped the 'index' column.
Dataset shape after all modifications: (2711, 2)
Unique categories after filtering: 5

```

We then plotted the dataset to evaluate the distribution of classes. On the X-axis we have the categories and on the Y-axis the number of news stories for each category (Fig. 36.2).

Plotting the dataset indicates a degree of imbalance, though it remains moderately proportioned. The fact that news from the 'Extern' and 'Politic' classes are more prevalent reflects the characteristics of this news television station. Thus, we did not take steps to over or under sample this dataset to make it more balanced. As the dataset grows, we could study the effects of balancing in a future paper.

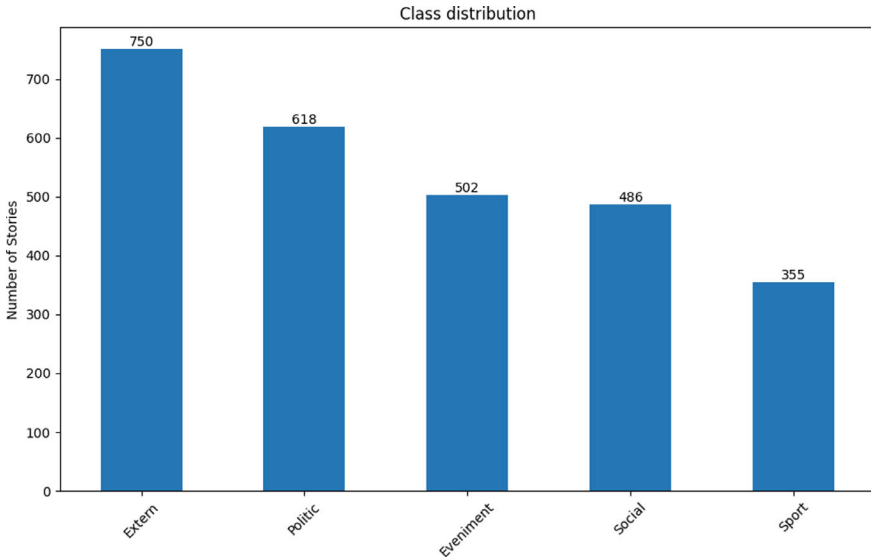


Fig. 36.2 Dataset class distribution

36.3.3 Pre-Processing

To further clean the dataset and remove noise, using Python, we took several steps:

Lower-case conversion. First, we converted all the characters in the dataset to lowercase letters using `.lower()` method.

Tokenization. For the next steps we relied on The Natural Language Toolkit (NLTK), the leading platform for building Python programs to work with human language data. NLTK is widely used for teaching and research in computational linguistics and natural language processing (NLP) [11] and has native support for the Romanian language. We used the “word_tokenize” library to split the text into individual words.

Removal of stop words, numbers and symbols. After tokenization, we removed any numbers and symbols then, we used the ‘stopwords’ attribute from the ‘nltk.corpus’ module to filter out the most common words which usually do not carry significant meaning, do not add any value to the current text-processing activity and, by being removed, may improve the performance of NLP models [12]. We passed the argument ‘romanian’ to the `.words` method of the ‘stopwords’ attribute in NLTK. The `.words` method then returns the stop words relevant to the Romanian language.

Stemming. The last step was to apply a stemming algorithm by importing the ‘SnowballStemmer’ class from the NLTK library, enabling the program to reduce words to their base or root form.

36.3.4 *Feature Extraction*

We decided to use TF-IDF and CountVectorizer to convert text data into numerical vectors, and evaluate which method provides the best results for our dataset.

CountVectorizer. Is a simple method that works by counting the number of times each word appears in a document. It converts a collection of text documents into a matrix of token counts. Each entry in the matrix represents the frequency of a particular word in a specific document.

TfidfVectorizer. TF-IDF, short for Term Frequency-Inverse Document Frequency, reflects how important a word is to a document in a collection or corpus. TfidfVectorizer goes a step further by not only counting the word frequency but also considering the importance of words across all documents. TF, or Term Frequency, similar to CountVectorizer, it measures how frequently a term occurs in a document. IDF, Inverse Document Frequency, measures the importance of the term across the corpus. Words that occur in many documents are less informative and are therefore assigned a lower score.

36.3.5 *Model Training*

In the preparation phase of the machine learning workflow, the dataset was divided into two distinct subsets to facilitate both the training and evaluation of the model. We used a ratio of 80/20. Specifically, 80% of the data was reserved for the training set, allowing the model to learn and adapt to the patterns within this larger portion of the dataset. The remaining 20% was reserved as the testing set, designed to provide an assessment of the model's performance. This division ensures that the model is trained on a comprehensive dataset while retaining a separate, untouched portion of data for evaluating its classification ability on new, unseen examples.

For the training phase, we applied the Multinomial Naïve Bayes algorithm that falls under the category of supervised machine learning techniques, which require a labeled dataset for training. In our case, each news story in the training set is accompanied by a label denoting its category, enabling the model to learn the probability of a story belonging to each category based on the distribution of words it contains.

36.3.6 *Model Testing*

During the testing phase, the classifier was applied to the 20% of the data reserved for testing, which had not been exposed to the model during the training phase. This step is essential for assessing the model's generalization capability and overall performance in classifying documents into their respective categories. The classifier's performance was assessed using several metrics.

Accuracy is the fundamental metric this study focused on to evaluate the performance of the classification model. It is calculated as the proportion of correct predictions made by the model over the total number of predictions made (Eq. 36.1).

$$accuracy = \frac{total\ correct\ predictions}{total\ predictions} \quad (36.1)$$

We also used metrics like precision, recall and f1-score to obtain valuable insights into different aspects of how well the model is functioning.

36.3.7 *Model Usage*

This model can be seamlessly incorporated into the daily operational workflow, thereby reducing the workload of journalists by automating the function of news classification and generating a weekly report. In the media asset management system of the news station, a service can be configured to automatically extract all the XML files containing news stories on a schedule, at the end of the week in our case. This service will store all these files in a folder. A Python script can monitor changes of that folder and trigger actions when new files are added. This functionality is commonly referred to as a “watchfolder”. The actions triggered are pre-processing the new data, vectorization to convert the new text into the same vector space as the training data and pass it into the model’s `.predict` method to classify the news story into the appropriate category. This script will then output the news category distribution weekly report.

36.4 Results and Discussion

We implemented the proposed Multinomial Naïve Bayes classification model using two vectorization methods and will detail the outcomes for each.

36.4.1 *MNB with CountVectorizer Method Applied for Pre-Processing*

Using this method, we obtain an accuracy of 0.8232 so the model correctly classified about 82% of the news stories into their correct categories (Table 36.1).

We can observe that the classes ‘Extern’ and ‘Eveniment’ have higher precision compared to the others suggesting that when the model predicts news stories as belonging to these categories, it is very likely to be correct. The high recall for ‘Politic’ and ‘Eveniment’ indicates the model is very effective at identifying all

Table 36.1 Metrics for MNB with CountVectorizer

Accuracy: 0.8232				
	Precision	Recall	F1-score	Support
Social	0.72	0.78	0.75	98
Extern	0.90	0.77	0.83	157
Politic	0.80	0.87	0.83	118
Sport	0.77	0.80	0.78	94
Eveniment	0.94	0.95	0.94	76
Macro avg.	0.83	0.83	0.83	543
Weighted avg.	0.83	0.82	0.82	543

relevant instances of these categories. Category ‘Social’ has lower precision and recall compared to others, suggesting a potential area for model improvement.

The overall accuracy indicates a good model, but there’s room for improvement (Fig. 36.3).

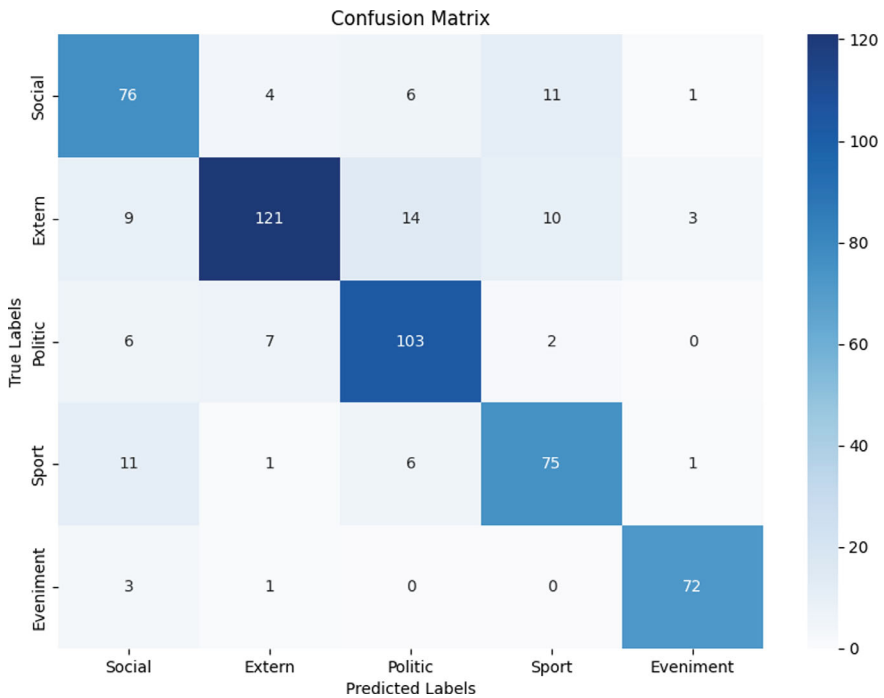


Fig. 36.3 Confusion matrix for MNB with CountVectorizer

Table 36.2 Metrics for MNB with TF-IDF

Accuracy: 0.7956				
	Precision	Recall	F1-score	Support
Social	0.74	0.70	0.72	98
Extern	0.79	0.85	0.82	157
Politic	0.74	0.89	0.81	118
Sport	0.82	0.64	0.72	94
Eveniment	0.97	0.86	0.91	76
Macro avg.	0.81	0.79	0.80	543
Weighted avg.	0.80	0.80	0.79	543

36.4.2 *MNB with TF-IDF Method Applied for Pre-Processing*

Using `TfidfVectorizer` we obtained an accuracy of 0.7956 which indicates that the model correctly predicted the class of approximately 79.56% of the news stories in the test set. This is a small but significant decrease from the accuracy obtained using `CountVectorizer` (Table 36.2).

The model demonstrates good performance across ‘Extern’, ‘Politic’, and ‘Eveniment’ categories, where it shows both high precision and recall. This suggests effective classification with both a low rate of false positives and a high rate of correctly identified true positives. The ‘Sport’ category, despite high precision, indicates a need for improvement in recall, suggesting the model is currently missing a significant number of ‘Sport’ news. In our model, using this dataset, `CountVectorizer` outperformed `TfidfVectorizer`, which indicates that, for our dataset, the sheer frequency of terms has a stronger signal for classification than the relative importance of terms across documents (Fig. 36.4).

36.4.3 *Additional Model Testing*

As an additional step, to put the results of our news classifier application into context, we extended our analysis to include a Support Vector Machine (SVM) model. SVM is a supervised machine learning algorithm with good results in classification tasks, including text classification [13]. We used the same dataset, same preprocessing and applied the same vectorization methods, `CountVectorizer` and TF-IDF, to transform the text data before training the SVM model (Table 36.3).

We can observe that approximately 77.7% of the predictions made by the model are correct. Also, there is an imbalanced performance across classes, the ‘Social’ and ‘Sport’ categories having comparatively lower scores. This could be due to potential

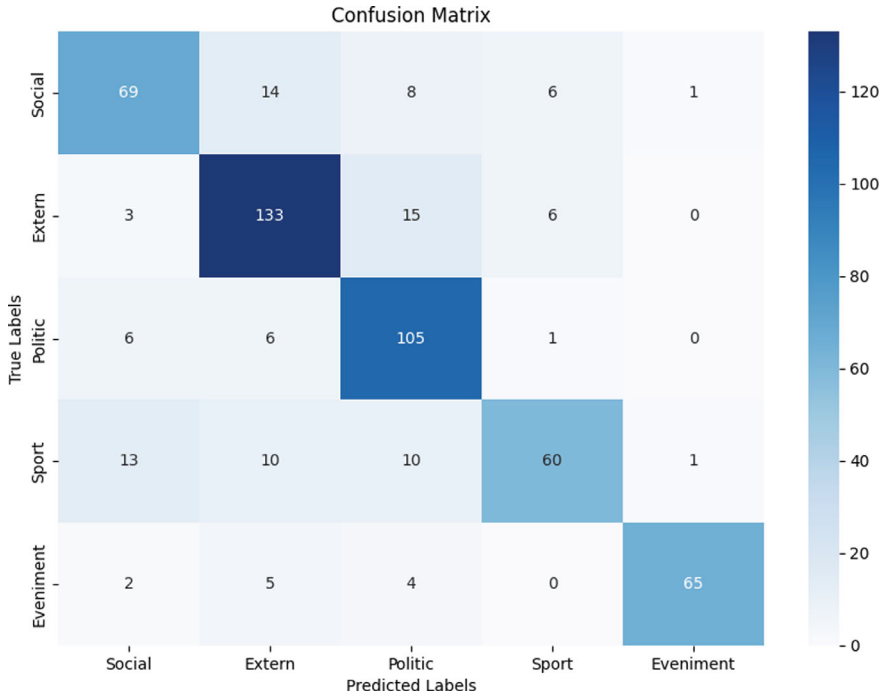


Fig. 36.4 Confusion matrix for MNB with TF-IDF

Table 36.3 Metrics for SVM with CountVectorizer

Accuracy: 0.7772				
	Precision	Recall	F1-score	Support
Social	0.61	0.60	0.61	98
Extern	0.85	0.87	0.86	157
Politic	0.82	0.83	0.82	118
Sport	0.66	0.68	0.67	94
Eveniment	0.96	0.84	0.90	76
Macro avg.	0.78	0.77	0.77	543
Weighted avg.	0.78	0.78	0.78	543

vocabulary overlapping with other categories, due to the smaller amount of training data for these classes or due to the unbalanced dataset (Table 36.4).

The use of TF-IDF has led to improved precision, recall, and F1-scores across all categories compared to CountVectorizer. This suggests that TF-IDF provides more informative features for the SVM to distinguish between different news stories categories. The ‘Social’ and ‘Sport’ categories, which had the lowest scores with

Table 36.4 Metrics for SVM with TF-IDF

Accuracy: 0.8416				
	Precision	Recall	F1-score	Support
Social	0.75	0.77	0.76	98
Extern	0.88	0.87	0.88	157
Politic	0.86	0.86	0.86	118
Sport	0.75	0.80	0.77	94
Eveniment	0.99	0.89	0.94	76
Macro avg.	0.85	0.84	0.84	543
Weighted avg.	0.84	0.84	0.84	543

Table 36.5 Overall accuracy of classifiers

Model	Vector	Accuracy
MNB	CountVectorizer	0.82
	TF-IDF	0.79
SVM	CountVectorizer	0.77
	TF-IDF	0.84

CountVectorizer, show the most notable improvements. This indicates that TF-IDF is better at capturing the distinctive features of these categories (Table 36.5).

When comparing the classifiers we observe that the MNB model performs significantly better than the SVM model when using CountVectorizer. This suggests that MNB, which is based on word frequency and assumes feature independence, is better suited to the simpler representation provided by CountVectorizer. Conversely, the SVM model significantly outperforms the MNB model when using TF-IDF. This indicates that SVM benefits more from the nuanced feature representation provided by TF-IDF, which captures term importance and reduces the weight of common but less informative terms. Upon reviewing the results of all methods used, we conclude that SVM with TF-IDF achieved better accuracy compared to MNB with CountVectorizer, although the improvement is marginal. This encouraging result prompts us to explore additional models and techniques to further enhance the classification of news stories.

36.5 Conclusions

This paper proposed an alternative approach to the manual labeling of news stories for classification, using supervised machine learning. Evaluation of the Multinomial Naïve Bayes classifier shows that the model is a promising tool for classifying news items, with specific areas of strength but also with areas that need further refinement.

The goal of the work presented in this paper is to automate news labelling and, by doing so, remove one more repetitive step of the editorial workflow.

Starting February 2024, as more categories were added, the Romanian news television station made story labelling mandatory. This will lead to a rapidly expanding dataset and will open many possibilities for future work. We could try more complex models, optimizing the classifiers, subsampling or oversampling the dataset to achieve perfect balance and, by doing so, reach a higher level of performance to better accomplish the goal that we set to automate and replace the current workflow. The authors also intend to define and apply classifier models that are using the news images, in addition to news text.

Disclosure of Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Allahyari, M., et al.: A brief survey of text mining: classification, clustering and extraction techniques (2017). arXiv preprint [arXiv:1707.02919](https://arxiv.org/abs/1707.02919)
2. Saigal, P., Khanna, V.: Multi-category news classification using support vector machine based classifiers. *SN Applied Sciences* **2**(3), 458 (2020)
3. Mulahuwaish, A., et al.: Efficient classification model of web news documents using machine learning algorithms for accurate information. *Comput. and Secur.* **98**, 102006 (2020)
4. Dilrukshi, I., De Zoysa, k., Amitha C.: Twitter news classification using SVM. In: 2013 8th International Conference on Computer Science and Education. IEEE (2013)
5. Parida, U., Nayak, M., Nayak. AK.: News text categorization using random forest and naïve bayes. In: 2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON). IEEE (2021)
6. Daud, S., et al.: Topic classification of online news articles using optimized machine learning models. *Comp.* **12.1**, 16 (2023)
7. Dhar, P., Abedin, MZ.: Bengali news headline categorization using optimized machine learning pipeline. *IJ Inform. Engineering Electron. Busi.* **13.1**, 15–24 (2021)
8. Barua, A., Omar S., Hoque, MM.: Multi-class sports news categorization using machine learning techniques: resource creation and evaluation. *Procedia. Comput. Sci.* **193**, 112–121 (2021)
9. Hossain, MR., Sarkar, S., Rahman, M.: Different machine learning based approaches of baseline and deep learning models for bengali news categorization. *Int. J. Comput. Appl.* **975**, 8887 (2020)
10. Gopal, Lakshmi S., et al.: Machine learning based classification of online news data for disaster management. In: 2020 IEEE Global Humanitarian Technology Conference (GHTC). IEEE (2020)
11. Wang, M., Hu, F.: The application of nltk library for python natural language processing in corpus research. *Theory Pract. Lang. Stud.* **11.9**, 1041–1049 (2021)
12. Kaur, J., Buttar. PK.: A systematic review on stopword removal algorithms. *Int. J. Futur. Revolut. Comput. Sci. and Commun. Eng.* **4.4**, 207–210 (2018)
13. Cervantes, J., et al.: A comprehensive survey on support vector machine classification: applications, challenges and trends. *Neurocomputing* **408**, 189–215 (2020)

Chapter 37

Text Summarization and Sentiment Analysis Pipelines Using Large Language Models for Financial News



Alin-Gabriel Văduva  and Anca-Ioana Andreescu

Abstract In an era dominated by vast amounts of unstructured web data, the need for sophisticated extraction and analysis methodologies is imperative. In this paper our goal is to address the challenge of summarizing financial news efficiently utilizing state-of-the-art transformer-based Large Language Models (LLMs) — specifically “t5-small”, and “sshleifer/distilbart-cnn-12–6” and conducting preliminary sentiment analysis. The performance of these summarization pipelines is rigorously evaluated against a comprehensive suite of metrics, including ROUGE, Keyword Overlap Score and Semantic Similarity Score for both models to validate the summaries’ accuracy and coherence. This investigation involved the CNBC news dataset, selecting 580 articles from the entire set available. By showcasing the capability of these models to distill pertinent information and hint at market sentiments using sentiment analysis pipelines, the research subtly underscores the emerging credibility and potential of AI techniques. This foundational work aims not only to enrich financial market analysis but also to catalyze further in-depth studies that refine Artificial Intelligence’s role in interpreting and leveraging the vast digital information landscape for strategic insights.

37.1 Introduction

In the context of Large Language Models (LLMs), text summarization (TS) of long documents [1] and sentiment analysis (SA) stand out as pivotal applications that address the critical challenge of information overload in the business sector, which is generating an immense volume of data daily, financial news being an important

A.-G. Văduva · A.-I. Andreescu (✉)

Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania

e-mail: anca.andreescu@ie.ase.ro

A.-G. Văduva

e-mail: alin.vaduva@csie.ase.ro

portion of this resulted data. Summarizing and analyzing the sentiment it's a practical necessity for market participants to make the best decisions easily. Pre-trained Language Models are often employed in the summarization task [2], as they facilitate the process of conducting summarization.

Text summarization can be classified as extractive [3] and abstractive [4]. These approaches address the growing need for efficient information processing. Extractive text summarization operates by identifying and extracting key phrases, sentences or segments directly from the source text. This method is based on the capacity of the summarization algorithms to evaluate the importance of textual elements based on metrics such as frequency, relevance, and the relationship between sentences. The objective is to compile a summary that represents the most salient information without altering the original context. A comprehensive overview of extractive text summarization methodologies, their associated challenges and potential future directions is detailed in [5]. Conversely, abstractive text summarization focuses on understanding the underlying meaning and nuances of the source material, generating new sentences that capture the core ideas in a concise manner, which requires a deep semantic understanding of the text. In [6], the authors provide an overview regarding some of the most utilized techniques, datasets and assessing metrics using deep learning models for abstractive summarization.

Another classification of the text summarization techniques is represented by the prior existence in the datasets of a human-annotated summary, denominating them into supervised summarization and unsupervised summarization. Supervised summarization involves the use of labeled datasets to train machine learning models on, allowing the models to learn different patterns and styles. On the other hand, the unsupervised approach does not rely on pre-labeled data. Instead, it utilizes algorithms to discover patterns and key information directly from the source text. Both supervised and unsupervised methodologies can be used separately or combined [7] leveraging their generative capacity for optimal results.

Sentiment analysis offers a lens through which the subjective information embedded in text summaries can be quantitatively assessed [8]. Being a key aspect of natural language understanding, it enables the identification of sentiment polarity (positive, neutral or negative) within the generated summaries. This approach facilitates a deeper evaluation of how well the summarization models preserve the original sentiment of the text.

In our work, we explore the potential of Google's Text-to-Text Transfer Transformer (T5-small) and DistilBART-CNN-12-6, both Transformer models being publicly available on HuggingFace. We conduct sentiment analysis on the generated summaries utilizing the pre-trained transformer model DistilRoberta-financial-sentiment, which is primarily designed for sentiment analysis in the financial news domain. The integration of sentiment analysis into the evaluation framework for generated summaries introduces a novel methodology for assessing their credibility. This analytical approach facilitates a comprehensive understanding of the summaries, encompassing both the factual accuracy and the emotional integrity of the content. By ensuring the alignment of the generated summaries with the sentiment of the original texts, we aim to augment the reliability and utility of AI algorithms in scenarios

where both credibility and the perception of emotional context are mandatory. Such applications are especially relevant in areas like business simulations and decision-making processes, where the nuanced interpretation of sentiment can significantly influence outcomes.

37.2 Literature Survey

In numerous scientific studies, Transformer-based models have been extensively utilized for tasks such as text summarization and sentiment analysis. These studies focus on evaluating the models' generative capabilities and their efficiency in classifying texts according to emotional content. Transformers, with their advanced architecture, excel in understanding contextual nuances and generating accurate summaries as well as identifying the sentiment of texts, showcasing their significant potential in the fields of text analysis and Natural Language Processing (NLP). In [9], the authors proposed a text summarization technique based on sentiment analysis, comparing it with standard graph-based summarizers or methods such as Latent Semantic Analysis (LSA) and Text Relationship Map (TRM). The effectiveness of the two NLP approaches has been captured also in Social Networks [10] by creating a model for posts published on Twitter in Turkish language. A two-level summarization technique with sentiment analysis has been conducted by the authors of [11], applied on multiple online articles grouped by their main topic, giving a sense about the view of existing news channels.

Text summarization plays a crucial role in several areas, especially in the biomedical and healthcare sectors [12, 13]. In these fields, the effectiveness of LLMs for summarizing text is thoroughly analyzed. In these works, it is highlighted how LLMs can simplify complex medical documents into more accessible summaries, yet it also points out the limitations these models encounter when dealing with the specific and technical language of medical literature. In the retail domain, text summarization has been applied to understanding customer opinions on online products [14]. Here, text summarization helps in gathering and condensing vast amounts of customer feedback, offering clear insights into product satisfaction and areas for improvement. In [15], the authors provided a survey regarding the employment of text summarization and sentiment analysis of online reviews from different websites, such as Social Networks (Twitter, Facebook) and product reviews from Amazon, as well as reviews of movies and news articles.

Within the scope of assessing credibility, the study presented in [16] implements sentiment analysis on news content sourced from Twitter. This approach involves the application of a supervised learning technique to train a Support Vector Classifier using the gathered tweets. The methodology culminates in achieving a notable accuracy rate of 95%, underscoring the efficacy of employing sentiment analysis in conjunction with machine learning models for the verification of news credibility on social media platforms. Regarding text summarization, the credibility is assessed from the perspective of reliability of different evaluation metrics [17], including

Recall-Oriented Understudy for Gisting Evaluation (ROUGE) and BLEU (Bilingual Evaluation Understudy). Although BLEU score is specific for machine translation, it can be adapted to measure the performance of summarization tasks [18].

In business areas, a multitude of news articles, regardless of their sources, do not have a pre-existing human annotated summary. Unsupervised methods are often utilized to retrieve the necessary insights on the raw datasets extracted from the Internet. In [19], a hybrid strategy based on cluster-ranking summarization is adopted, focusing on business-related tweets. This approach, demonstrating the utility of domain concepts for summarizing diverse tweets, highlights the method's portability due to reliance on readily available resources like unannotated news articles and WordNet. Future directions include exploring fine-grained concepts for varying classes of companies and adjusting summarization techniques based on tweet properties to enhance summary quality.

37.3 Research Methodology

The subsequent section proposes an approach for automatically summarizing financial news articles using a selected subset from the CNBC news dataset published on data. World, which originally included 625 articles. We curated a manageable set of 580 articles for analysis to address the challenges of processing extensive text data, focusing on essential attributes such as the article title, short description, and full description. These attributes were efficiently managed using the Python programming language alongside Pandas DataFrames, facilitating streamlined data handling and preparation.

For the summarization process, we utilized two advanced language models: Google's Text-To-Text Transfer Transformer-Small (T5-Small) and DistilBART CNN 12–6. These models were selected based on their proven effectiveness in text summarization tasks, supported by their extensive pre-training on large-scale text corpora like the Colossal Clean Crawled Corpus (C4) and CNN/DailyMail. This choice allowed us to leverage both unsupervised and supervised learning techniques to generate concise summaries of the financial news articles.

Our implementation was carried out using Jupyter Lab from the Anaconda suite, which provided an interactive and flexible environment for coding, debugging, and visualizing results. The Python environment was utilized for data manipulation and model application, with libraries including Pandas for data handling, Hugging Face's Transformers for model implementation, and Scikit-learn for evaluation metrics. The hardware specifications for our computational environment are as follows: we operated on a system with a 12th Gen Intel(R) Core (TM) i5-12450H processor (12 CPUs, 2.0 GHz), 16 GB of RAM, and an NVIDIA GeForce RTX 3050 Laptop GPU.

To evaluate the accuracy and coherence of our generated summaries, we employed three distinct metrics: ROUGE scores to quantify the overlap of n-grams between the generated summaries and the original texts; Keyword Overlap Score, which uses TF-IDF vectorization to assess how effectively the summaries capture key thematic

elements from the original articles; and Semantic Similarity Score, measuring the cosine similarity between semantic embeddings of the original texts and their summaries to evaluate the preservation of underlying meanings and paraphrasing accuracy.

Further, we incorporated a sentiment analysis component using the DistilRoBERTa-financial-sentiment model, pre-trained on a large corpus of financial news, to categorize the sentiment expressed in the summaries into neutral, positive, or negative. We also used word cloud analysis to visually summarize and identify the most prominent themes and concepts within the summaries, adding an additional layer of qualitative analysis.

37.4 Findings

The dataset was refined by extracting 580 financial news articles and selectively preserving key information from each. This careful selection aimed to maintain the essential content needed for analysis while omitting non-essential data. The structure of each instance was designed to clearly present the most relevant aspects of the financial narrative for subsequent examination and research applications. The following example describes how an instance in the dataset is represented:

- Title: ‘Santoli’s Wednesday market notes: Could September’s stock shakeout tee up strength for the fourth quarter?’
- Short_description: “This is the daily notebook of Mike Santoli, CNBC’s senior markets commentator, with ideas about trends, stocks and market statistics.”
- Description: “This is the daily notebook of Mike Santoli, CNBC’s senior markets commentator, with ideas about trends, stocks and market statistics. A muted, inconclusive bounce that has left the indexes fully within yesterday’s low-to-high range all morning so far”.

The performance is measured quantitatively in Table 37.1, comparing the two language models with respect to ROUGE score, Keyword Overlap Score and the Semantic Similarity Score.

The ROUGE-1 F1 Score was higher for DistilBART-CNN-12-6 (0.3126) compared to T5-Small (0.2612). This suggests that DistilBART-CNN-12-6 is more

Table 37.1 Summarization metrics for each LLM

Metric	T5-Small	DistilBART-CNN-12-6
ROUGE-1 F1	0.2612	0.3126
ROUGE-2 F1	0.1851	0.2261
ROUGE-L F1	0.2597	0.3108
Keyword overlap score	0.3755	0.4398
Semantic similarity score	0.7094	0.7681

adept at capturing the key words that are present in the reference summaries. For the ROUGE-2 F1 Score DistilBART-CNN-12-6 again leads with a score of 0.2261, surpassing T5-Small's score of 0.1851, indicating a better grasp of phrase structures by DistilBART-CNN-12-6, which is important for maintaining the coherence of information in summaries.

The ROUGE-L F1 Score, indicating the extent to which the models can reproduce sequences of words found in the reference, sees DistilBART-CNN-12-6 with a higher score (0.3108) compared to T5-Small (0.2597). This points to a greater ability of DistilBART-CNN-12-6 to maintain the flow and structure of the original text.

In terms of Keyword Overlap Score DistilBART-CNN-12-6 scores 0.4398 against T5-Small's 0.3755, indicating that DistilBART-CNN-12-6 is likely to include more relevant information from the original text.

Finally, the Semantic Similarity Score is marginally higher for DistilBART-CNN-12-6 (0.7681) than for T5-Small (0.7094). This reveals that the overall meaning conveyed by DistilBART-CNN-12-6 is more aligned with the expected outcomes.

While both models demonstrate capability in text summarization, DistilBART-CNN-12-6 consistently shows superiority across all tested metrics, suggesting it as a more proficient model for summarizing texts in a manner that is both accurate and contextually coherent.

Upon application of the sentiment analysis pipeline “mrm8488/distilroberta-finetuned-financial-news-sentiment-analysis” on summaries corpus generated by the DistilBART-CNN-12-6 model, the sentiment distribution across 580 financial news instances was quantitatively depicted in Fig. 37.1. The results indicated a predominance of neutral sentiment at 39.9% of the instances, suggesting a balanced representation of financial news without a strong positive or negative bias. Positive sentiment was reflected in 33.2% of the articles, indicating a favorable outlook in a substantial fraction of the financial discourse. Negative sentiment constituted 26.9% of the instances, reflecting critical or pessimistic perspectives in financial reporting.

The sentiment analysis of financial summaries generated by the T5-Small model, when compared to DistilBART-CNN-12-6 is depicted in Fig. 37.2. T5-Small model generates slightly less negative sentiment (26.3% vs. 26.9%) and marginally more positive sentiment (33.5% vs. 33.2%) than DistilBART-CNN-12-6. Additionally, the T5-Small model exhibits a small increase in neutral sentiment (40.2% vs. 39.9%). While both models are consistent in their representation of negative sentiment, the T5-Small model leans slightly more towards positive and neutral sentiments compared to DistilBART-CNN-12-6.

As we move to the final analysis, we present a comparative study of word clouds to visualize and contrast the summarization capabilities of the two different text summarization pipelines. The left side of the Fig. 37.3 contains the word cloud generated on behalf of the summaries of DistilBART LLM, whereas the right side represents the word cloud specific for the T5-Small language model.

The comparative word cloud analysis reveals notable differences in the summarization outputs. The t5-small pipeline appears to produce summaries with a broader lexicon, as indicated by the presence of a more varied set of keywords such as “positive,” “growth,” and “energy”. This suggests an ability to capture a wider range of

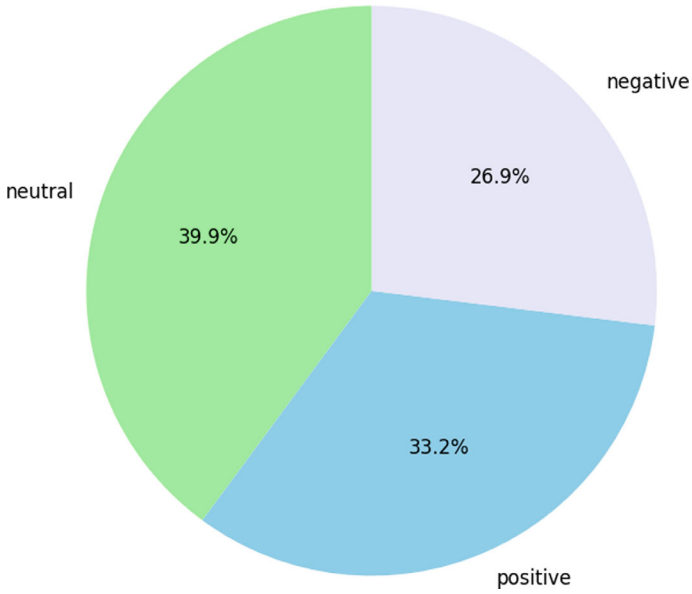


Fig. 37.1 Sentiment analysis distribution for summaries generated using DistilBART-CNN-12-6

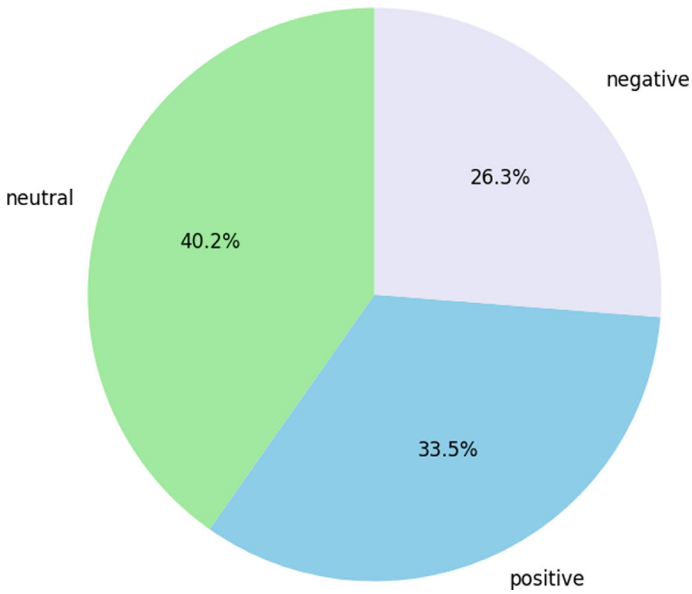


Fig. 37.2 Sentiment polarity for T5-small summaries



Fig. 37.3 Word clouds for the two employed LLMs

themes within the financial sector, possibly incorporating more optimistic language and industry-specific terms. Conversely, the word cloud generated from the distilbart-cnn-12-6 pipeline’s summarization demonstrates a concentration on quantifiable and entity-specific terminology, with words like “million,” “billion,” and “New York” being particularly salient. Such emphasis indicates a potential focus on the financial scale and specific geographical locations within the texts it processed.

37.5 Conclusion

The comparative analysis of the T5-Small and DistilBART-CNN-12-6 language models across various evaluation methods—summarization metrics, sentiment distribution, and word cloud visualizations—indicates distinctive characteristics and strengths of each model. T5-Small exhibits a wider vocabulary and thematic diversity in its summarization, potentially offering more nuanced insights into the financial sector, as evidenced by the varied keywords and the slightly more positive sentiment distribution in the generated summaries. DistilBART-CNN-12-6, on the other hand, tends to focus on quantitative data and specific entities, as shown by the prominence of financial figures and place names, alongside a comparable sentiment distribution to T5-Small.

These findings suggest that the choice of summarization model can significantly influence the interpretation and decision-making process in a business context, particularly in the fast-paced domain of financial news. The T5-Small’s broader lexicon may provide businesses with a comprehensive overview of market narratives, while the precision of DistilBART-CNN-12-6’s output might be more suitable for those requiring detailed factual reporting. The credibility of these algorithms’ hinges on

their ability to accurately reflect the underlying data, and this comparison suggests that both models are capable tools, each with unique applications that can enhance the analysis and dissemination of financial information within a business setting.

Acknowledgements This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS-UEFISCDI, project number PN-III-P4-PCE-2021-0334, within PNCDI III.

References

1. ge Yao, J., Wan, X., Xiao J.: Recent advances in document summarization. *Knowl. Inf. Syst.* **53**(2), 297–336 (2017). <https://doi.org/10.1007/S10115-017-1042-4/FIGURES/5>
2. Xiao, L., Wang, L., He H., Jin, Y.: Modeling content importance for summarization with pre-trained language models. In: EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, pp. 3606–3611, 2020. <https://doi.org/10.18653/V1/2020.EMNLP-MAIN.293>
3. Erkan, G., Radev, D.R.: LexRank: graph-based lexical centrality as salience in text summarization. *J. Artif. Intell. Res.* **22**, 457–479 (2011). <https://doi.org/10.1613/jair.1523>
4. Khan, A., Salim, N., Jaya Kumar, Y.: A framework for multi-document abstractive summarization based on semantic role labelling. *Appl. Soft. Comput.* **30**, 737–747 (2015). <https://doi.org/10.1016/J.ASOC.2015.01.070>
5. Moratanch N., Chitrakala, S.: A survey on extractive text summarization. In: International Conference on Computer, Communication, and Signal Processing: Special Focus on IoT, ICCOSP 2017 (2017). <https://doi.org/10.1109/ICCCSP.2017.7944061>
6. Suleiman, D., Awajan, A.: Deep learning based abstractive text summarization: approaches, datasets, evaluation measures, and challenges. *Math. Probl. Eng.* **2020** (2020). <https://doi.org/10.1155/2020/9365340>
7. Mao, X., Yang, H., Huang, S., Liu, Y., Li, R.: Extractive summarization using supervised and unsupervised learning. *Expert Syst. Appl.* **133**, 173–181 (Nov.2019). <https://doi.org/10.1016/J.ESWA.2019.05.011>
8. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng. J.* **5**(4), 1093–1113 (2014). <https://doi.org/10.1016/J.ASEJ.2014.04.011>
9. Yadav N., Chatterjee, N.: Text summarization using sentiment analysis for DUC Data, pp. 229–234 (2017). <https://doi.org/10.1109/ICIT.2016.054>
10. Dogan, E., Kaya, B.: Deep learning based sentiment analysis and text summarization in social networks. In: 2019 International Conference on Artificial Intelligence and Data Processing Symposium, IDAP 2019 (2019). <https://doi.org/10.1109/IDAP.2019.8875879>
11. Mirani, T.B., Sasi, S.: Two-level text summarization from online news sources with sentiment analysis. In: 2017 International Conference on Networks and Advances in Computational Technologies, NetACT 2017, pp. 19–24 (2017). <https://doi.org/10.1109/NETACT.2017.8076735>
12. Rohil, M.K., Magotra, V.: An exploratory study of automatic text summarization in biomedical and healthcare domain. *Healthcare Analytics* **2**, 100058 (2022). <https://doi.org/10.1016/J.HEALTH.2022.100058>
13. Sarkar, K.: Using domain knowledge for text summarization in medical domain (2009)
14. Zhan, J., Loh, H.T., Liu, Y.: Gather customer concerns from online product reviews – a text summarization approach. *Expert Syst. Appl.* **36**(2), 2107–2115 (2009). <https://doi.org/10.1016/J.ESWA.2007.12.039>

15. Pecar, S.: Towards opinion summarization of customer reviews. In: ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Student Research Workshop, pp. 1–8 (2018). <https://doi.org/10.18653/V1/P18-3001>
16. Sharf, Z., Jalil, Z., Amir, W., Siddiqui, N.: Establishing news credibility using sentiment analysis on twitter. *Int. J. Adv. Comput. Sci. Appl.* **10**(9), 209–221 (2019). <https://doi.org/10.14569/IJA CSA.2019.0100927>
17. Barbella, M., Risi, M., Tortora, G.: A comparison of methods for the evaluation of text summarization techniques. In: International Conference on Data Technologies and Applications, pp. 200–207, (2021). <https://doi.org/10.5220/0010523002000207>
18. Graham, Y.: Re-evaluating automatic summarization with BLEU and 192 shades of rouge. In: Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing, pp. 128–137 (2015). <https://doi.org/10.18653/V1/D15-1013>
19. Lou, A., Ew, T.N., An M.: Summarization of business-related tweets: a concept-based approach. pp. 765–774 (2012)

Chapter 38

Develop a Sentiment Analysis for Teaching Performance Using a Machine Learning Pipeline



Dimitrie-Daniel Plăcintă 

Abstract This article identified a lack in scientific literature for topics such as sentiment analysis systems, pipelines, and teacher evaluation. I exposed the challenges in developing sentiment analysis for teaching performance. An Anaconda environment with Jupyter Notebook and Spyder IDE was used to build the machine learning pipeline. To implement the pipeline, I used a large (2 million records and 8 attributes), cleaned, open-source data set for aspect-based sentiment analysis. The focus was to prepare a reduced pandas DataFrame by filtering the larger data set by the number of total words for student comment, sentiment type, and subjectivity type. The word cloud feature provides a clear status for teaching performance in visual approach. To produce sets of advice for each student comment, I have integrated the OpenAI and MistralAI models, comparing the generated content from both approaches in terms of LLM size, price, quality of content, and the execution time.

38.1 Introduction

The quality of education is important within a performant society because the citizens are paying their taxes in exchange for free educational paths until the end of the university level; this is the case of the Romanian educational system.

Teachers are the most important factors for guiding, orienting, mentoring, supporting, coaching, teaching the generations of pupils, students. The evaluation of teachers by their pupils, students is a very good practice for keeping the high professional standards, for hearing the voice of educational beneficiaries, for understanding the educational trends, for improving the pedagogical methods, for taking the remedial measures, etc.

In the ChatGPT revolution, the integration of the different large language models (LLM) should part of the solution for generation of the best results for stakeholders

D.-D. Plăcintă (✉)

The Bucharest University of Economic Studies, Bucharest, Romania

e-mail: daniel.placinta@csie.ase.ro

from education, because the LLMs should contain the updated knowledge. Custom LLMs can be implemented to address specialized niches in a larger domain field, the internal knowledge bases can be used as source of trust for specific machine learning models

38.2 Related Work

To gain more insight into the problematic of sentiment analysis for teaching performance I have run a query on Scopus, TITLE-ABS-KEY (sentiment AND analysis AND teacher AND performance), who returned 122 results, and then the export.csv file was uploaded to VOSviewer tool where I have created a network visualization map based on bibliographic data, with the parameters: co-occurrence for type of analysis, all keywords for units of analysis, and full counting for counting method. In this way we can understand visually the representation between the most important nodes, categorized into 5 clusters (Fig. 38.1).

As we can remark in the above map, there are more opportunities in developing sentiment analysis systems, pipelines for evaluating the teacher performance, correlated with student satisfaction, because the links between sentiment analysis and other educational areas are already explored in the scientific literature.

From the Scopus, abstract and citation database, I'm presenting in Table 38.1 one of the most relevant studies related to sentiment analysis and teacher performance.

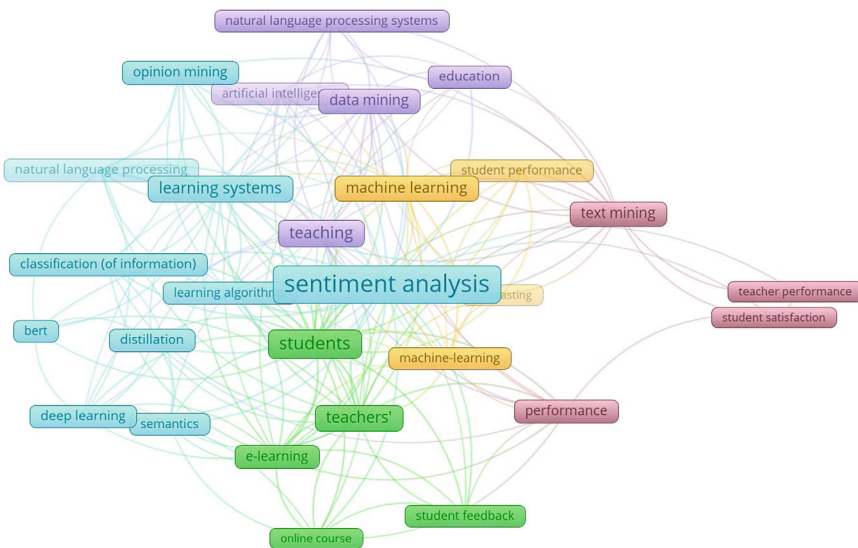


Fig. 38.1 Network visualization map based on search “TITLE-ABS-KEY (sentiment AND analysis AND teacher AND performance)” on Scopus

Table 38.1 Literature review for sentiment analysis and teacher performance

Summary	Algorithms	Results	Limitations	Source
‘Elaborated a new ready to use open-source dataset for sentiment analysis in educational field, very useful for teaching performance’	Dataset was validated across 15 distinct ML models	‘Random forest classifier got the highest accuracy, 0.96, Gaussian NB performed in balanced accuracy’	‘This research added a huge value to the lack of high-quality datasets for sentiment analysis for teacher performance’	[1]
‘Build an innovative model architecture for sentiment analysis in the evaluation of teaching based on the dataset [1], evaluating multiple views and perceptions of teaching aspects’	‘Proposed model’, Convolutional neural network (CNN), Gated Recurrent Unit (GRU)	86% precision for “proposed model”	‘Limited aspect coverage, more areas of teaching performance can be expanded, sarcasm or nuanced expression are not covered’	[2]
‘Predicted the student satisfaction with the teaching performance using Twitter comments as source of data’	Compared different SVM algorithms, best identified was Fine Gaussian	‘SVM with 98.3% accuracy’	No explicit limitation	[3]
‘Evaluate teacher strengths and weaknesses using opinion mining tool with sentiment analysis’	Naïve Bayes for text classification	‘Build the dataset polarity for every sentiment word, created a 4-module architecture: account management, student evaluation, reporting, teachers evaluation’	‘Build a better grammar and spelling checker and test Multinomial Naïve Bayes algorithm as future research implementation’	[4]

In addition to above details, in the recent studies from China [5], the facial expressions of students are analyzed in real time, if they are “confused, dissatisfied or happy”, because understanding these emotional states helps the educational approach of teachers, who can adapt teaching–learning strategies. Behind this new LCA_{Net} model are “deep learning” models, which incorporate attention mechanisms for recognizing and analyzing student expressions in the classroom.

38.3 Challenges in Developing Sentiment Analysis for Teaching Performance

The analysis of teaching performance is a tabu subject for educational institutions. Certainly, these analyses are done, but visibility is restricted at the department, faculty level, but the reports are exposed at the university level when an external audit is made for program's accreditation.

The second impediment is represented by the lack of running the evaluation of teaching performance at the end of semester the, even this is an important tool to hear the voice of students, and the students don't have the 100% confidence that the answers are confidential, without repercussions on their educational path, because these audits are run by universities, and not by external audit organizations.

The new EU (European Union) rules related to GDPR (General Data Protection Regulation), and the EU AI Act are also making this sentiment analysis very difficult to run without completing, respecting the bureaucratic procedures.

I consider that all educational institutions should implement a periodic mechanism for capturing the voice of stakeholders, then evaluate and validate the results from different angles, and the educational leaders to take appropriate decisions in this dynamic field.

38.4 Research Methodology

38.4.1 *Finding the Data Set Dataset Description*

I have run a query on <https://datasetsearch.research.google.com/> with the keywords 'sentiment analysis teacher performance evaluation' and I find out an interesting data set from [6] with the type of license CC BY NC 3.0. The dataset contains more than 2 million records with the next columns unnamed, StudentComments, Rating—between 1 and 5, totalwords, Sentiment—positive, negative, neutral, sent_pretrained, subjectivity—subjective or objective, subj-score, is Same—true or fake.

'The dataset was created by collecting student feedback from American International University-Bangladesh and then labelled by undergraduate-level students into three sentiment classes: positive, negative, and neutral' [6], the advantage of this collection is that all data was cleaned and preprocessed, being a qualitative dataset, properly prepared to run a sentiment analysis approach.

The format of the file.tsv (tab separated values) needs a small transformation before converting within a DataFrame in pandas.

Because the raw dataset was transformed to the clean dataset by following the most important steps for a sentiment analysis (data collection, data processing and cleaning, dataset preparation, dataset overview), I'll focus on transforming the reduced dataset into pragmatic results, Word Cloud diagram, and OpenAI and Mistral AI recommendations for each student comment.

38.5 Developing the Machine Learning Pipeline

38.5.1 Software Requirements

For this machine learning pipeline, I have used and tested my Python code within Jupyter notebook and Spyder IDE from Anaconda package manager.

38.5.2 Import Libraries and Upload Dataset

To process the.tsv file I used pandas library for DataFrame operations, matplotlib for saving the chart, WordCloud, and STOPWORDS libraries for completing the word cloud feature, and OpenAI and Mistral Client libraries for LLM integrations.

```
import pandas as pd
import matplotlib.pyplot as plt
from wordcloud import WordCloud, STOPWORDS
from openai import OpenAI
from mistralai.client import MistralClient
from mistralai.models.chat_completion import ChatMessage
df = pd.read_csv('sentiment.tsv', sep='\t')
```

38.5.3 Get the Filtered Dataset

Based on above df dataset we can filter the larger dataset by the minimum, maximum number of words with a student comment, sentiment type, subjectivity type, etc., as per below Python code:

```
max_words=40
min_words=30
sentiment_type="negative" # negative, positive or neutral
subjectivity_type="objective" # objective, subjective
df1=df[df["totalwords"]<max_words]
df1=df1[df1["totalwords"]>min_words]
df1=df1[df1["Sentiment"]==sentiment_type]
df1=df1[df1["subjectivity"]==subjectivity_type]
df1=df1[df1["isSame"]=='true']
```

After above filtration we got a dataset composed by 835 entries from 2,007,747 rows. I considered that a comment with more than 30 words is eloquent for a proper sentiment analysis, but other users can change above parameters as per their requirements.

38.5.4 *Word Cloud*

I have built a demo Word Cloud to quickly visualize the student sentiments for a filtered set of results, I have removed some general keywords with the stopwords parameter, set the maximum keywords that I want to display, and a background color as per below Python code, adaptation after [7]:

```
# concatenate all StudentComments
text=df1['StudentComments'].str.cat(sep=' ')
stopwords=set(STOPWORDS)
new_stop_words=(["student", "class", "teacher", "good",
"students", "english", "teaching", "u", "think", "know"])
final_stop_words=stopwords.union(new_stop_words)
# Creating word_cloud with text as argument in .generate() method
word_cloud = WordCloud(collocations = False, background_color =
'white', stopwords=final_stop_words, max_words=35).generate(text)
# Display the generated Word Cloud
plt.imshow(word_cloud, interpolation='bilinear')
plt.axis("off")
plt.savefig("wordcloud1.png", format="png")
plt.show()
```

38.5.5 *OpenAI Integration*

As part of the novelty for this sentiment approach I have integrated the chat completion object within the machine learning pipeline to provide personalized recommendation for each negative comment provided by student. I instructed the OpenAI chat completion object [8] to act like an expert in pedagogy by providing 5 short advice, and in Table 38.2 two examples of recommendations were highlighted (Fig. 38.2).

```
for i in range(len(df1)):
    index=df1.index[i]
    input=str(df1.loc[index, ['StudentComments']])
    response = client_openAI.chat.completions.create(
        model="gpt-3.5-turbo",
        messages=[
            {"role": "system", "content": "You an expert in pedagogy and
            enumerate with 5 short advice to solve the addressed scenario in
            maximum 40 words."},
            {"role": "user", "content": input},
        ],
        temperature=0,
    )
    advice_openAI=str(response.choices[0].message.content)
    df1.loc[index, ['Recommendation OpenAI']]=advice_openAI
```

Table 38.2 Chat recommendation based on student comment from OpenAI model

Student comment	Chat recommendation
Should not given bba and mba hrm course to the same teacher and he should be more flexible about rulls and regulations of aiub because he is misusing rulls of aiub	<ol style="list-style-type: none"> 1. Provide clear guidelines on appropriate language use in student comments 2. Offer training on professional communication for students in BBA and MBA HRM courses 3. Monitor and moderate student comments to ensure compliance with standards 4. Encourage respectful and constructive feedback among students 5. Address any inappropriate comments promptly and educate students on the impact of their words
She has enough knowledge.but she does not know the proper way to teach.her spoken english is not good.her dressup is not as good as it should be.she is quite rood to her student	<ol style="list-style-type: none"> 1. Encourage open communication to understand the student’s perspective 2. Provide additional resources or support to fill the knowledge gap 3. Offer guidance on effective study strategies 4. Encourage self-assessment and reflection 5. Foster a growth mindset to promote continuous learning and improvement

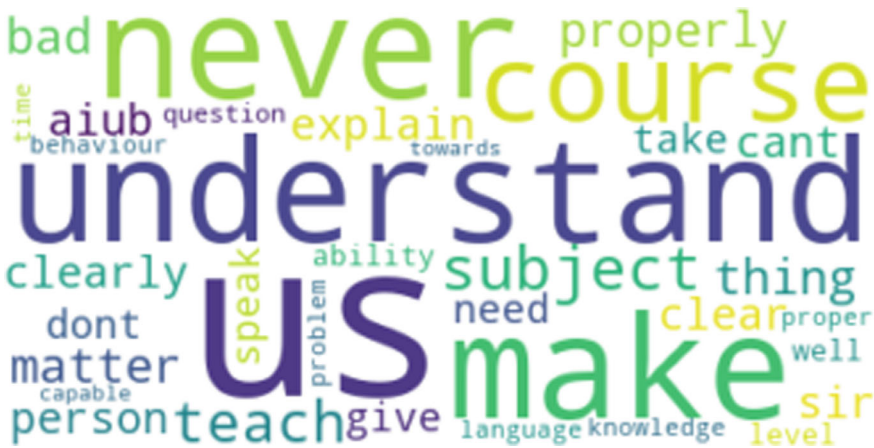


Fig. 38.2 Generated word cloud for a set of negative sentiments

38.5.6 *MistralAI Integration*

As part of my machine learning pipeline I introduced the MistralAI integration, programmed the mistral-small-latest to execute the same query as gpt-3.5-turbo model with the goal to compare the generated results from both models.

```
for i in range(len(df1)):
```

```

index=df1.index[i]
input=str(df1.loc[index, ['StudentComments']])
# MistralAI
input_mistral="You an expert in pedagogy and enumerate with 5 short
advice to solve the addressed scenario in maximum 40 words."+input
chat_response = client_mistralAI.chat(
model="mistral-small-latest",
messages=[ChatMessage(role="user", content=input_mistral)]
)
advice_mistral=str(chat_response.choices[0].message.content)
df1.loc[index, ['Recommendation MistralAI']] = advice_mistral

```

38.5.7 Results After Comparing the MistralAI and OpenAI

I have exposed below the comparison between MistralAI and OpenAI models for processing a single record, for providing a custom recommendation, and a chat completion for a student comment in Table 38.3.

The educational stakeholders can select the suitable model, considering the size of the LLM, price per token, quality of the content, and the speed for generating the advice.

38.6 Working Files

The SA-ConfIE24.py consolidate all Python code for this paper and can be accessed at [9], together with the used dataset [6].

38.7 Conclusions

Sentiment analysis in the educational field was very well documented and evaluation of teaching performance has a high potential for research. Based on a validated dataset for aspect-based sentiment analysis, I have explored using open source (Spyder IDE, JupyterNotebook) and paid solutions(OpenAI and MistralAI) new ways for presenting a clear picture with the status of teaching performance for filtered dataset. With the help of WordCloud diagram, OpenAI and MistralAI results, the educational leaders can take informed decision based on the text analysis using a machine learning pipeline. These results should also be cross validated over years, using additional evaluation methods, such as emotion detection, interviews with the students' representatives, etc.

Table 38.3 Comparison between MistralAI and OpenAI models

Features	MistralAI	OpenAI
Input = “u r so serious about lab.but your marking is not so good.you should have considerable tendency.you should remember that our a+ in 94 and above. please help us to get a+ . thank you“	<ol style="list-style-type: none"> 1. Understand Lab Requirements: Thoroughly read and comprehend the lab instructions to ensure you meet all criteria 2. Practice Accuracy: Focus on precision and attention to detail in your experiments to improve results 3. Review Previous Work: Analyze past labs for areas of improvement and apply those lessons to current work 4. Seek Clarification: If uncertain about any aspect, ask your instructor for clarification to avoid mistakes 5. Aim for 94+ : Set a target score of 94 or above, and work diligently to achieve it, ensuring all aspects of the lab are completed to a high standard 	<ol style="list-style-type: none"> 1. Clearly communicate grading criteria 2. Provide detailed feedback on assignments 3. Offer opportunities for improvement 4. Encourage students to seek clarification 5. Ensure fairness and consistency in grading
Execution time per record	Between 3–5 s to 40–45 s, somehow unpredictable	3 s, results predictable in a reasonable time
Model	mistral-small-latest	gpt-3.5-turbo
Price	Input: \$2/1 M tokens ^a Output: \$6/1 M tokens	Input: \$0.50/1 M tokens ^b Output: \$1.50/1 M tokens
Quality of generated content	Elaborate and precise content	Good content, not so precise like MistralAI

^a <https://mistral.ai/technology/#pricing>

^b <https://openai.com/pricing#language-models>

References

1. Bhowmik, A., Noor, N.M., Miah, M.S.U., Mazid-Ul-Haque, M., Karmaker, D.: A comprehensive dataset for aspect-based sentiment analysis in evaluating teacher performance. *AIUB J. Sci. Eng. (AJSE)*. **22**, 200–213 (2023). <https://doi.org/10.53799/ajse.v22i2.862>
2. Bhowmik, A., Noor, N.M., Miah, M.S.U., Karmaker, D.: Aspect-based sentiment analysis model for evaluating teachers’ performance from students’ feedback. *AIUB J. Sci. Eng. (AJSE)*. **22**, 287–294 (2023). <https://doi.org/10.53799/ajse.v22i3.921>
3. Chamorro-Atalaya, O., Arce-Santillan, D., Arévalo-Tuesta, J.A., Rodas-Camacho, L., Dávila-Laguna, R.F., Alejos-Ipanaque, R., Moreno-Chinchay, L.R.: Supervised learning using support

- vector machine applied to sentiment analysis of teacher performance satisfaction. *Indones. J. Electr. Eng. Comput. Sci.* **28**, 516 (2022). <https://doi.org/10.11591/ijeecs.v28.i1.pp516-524>
4. Balahadia, F.F., Fernando, Ma, C.G., Juanatas, I.C.: Teacher's performance evaluation tool using opinion mining with sentiment analysis. In: 2016 IEEE Region 10 Symposium (TENSYP), pp. 95–98, IEEE (2016). <https://doi.org/10.1109/TENCONSpring.2016.7519384>
 5. Hu, P., Tang, X., Yang, L., Xia, D.: A Real-time sentiment analysis model of students by fusing attention mechanism. In: 2023 13th International Conference on Information Technology in Medicine and Education (ITME), pp. 124–129, IEEE (2023). <https://doi.org/10.1109/ITME60234.2023.00035>
 6. Bhowmik, A., Noor, N.M., Miah, M.S.U., Haque, M.-U., Karmaker, D.: A novel dataset for aspect-based sentiment analysis for teacher performance evaluation. <https://data.mendeley.com/datasets/b2yhc95rnx/1>. Accessed 31 Mar 2024
 7. Vu, D.: Generating wordclouds in python tutorial. <https://www.datacamp.com/tutorial/wordcloud-python>. Accessed 01 May 2024
 8. Sanders, T.: How to format inputs to ChatGPT models. https://cookbook.openai.com/examples/how_to_format_inputs_to_chatgpt_models. Accessed 01 May 2024
 9. Plăcintă, D.-D.: ConfIE24 working files. <https://drive.google.com/drive/folders/1Aln0oUcR314Hx7wCeyZQh0cSHGH7yNqW?usp=sharing>. Accessed 09 Jun 2024

Chapter 39

Are LLMs Hallucinating When Answering Specific Questions? Experiments with a RAG Pipeline



Isabela Iacob and Gheorghe Cosmin Silaghi 

Abstract While the industry aims to automate more and more tasks with the help of Large Language Models (LLMs), assessing their capacity to provide reliable responses represents a legitimate concern. LLMs are trained over general knowledge bases, thus they are not intrinsically able to respond to questions over private, specific knowledge regarding the automated processes. This issue represents the main source of hallucination, as the LLM will not recognize its inability to supply a truthful response. This paper introduces a RAG pipeline for query answering from a specific text corpus, allowing to evaluate the quality of the retrieved responses. Inspecting the variability of LLM performance in time over the same query answering task, we strive to detect potential hallucinations. Results show that by analyzing the computed metrics we could advance toward a framework for hallucination detection and for assessing the confidence on the LLM performance.

39.1 Introduction

Artificial intelligence has gained momentum over the past few years and experienced exponential growth, becoming a hot-discussed topic in the technology industry. Large Language Models (LLMs) demonstrated incredible capabilities and are changing the way industries work, indicating that we might be at the beginning of a new technological era dominated by the rise of artificial intelligence.

While LLMs capture tremendous knowledge and compete with humans in many intellectual tasks, being able to tackle a variety of problems from many domains, they are known to occasionally hallucinate in various circumstances [6, 8]. LLM hallucination represents a concern for the software industry as they strive to use LLMs for automating various software development tasks because it would require human

I. Iacob · G. Cosmin Silaghi (✉)
Babeş-Bolyai University, Business Informatics Research Center, Cluj-Napoca, Romania
e-mail: gheorghe.silaghi@ubbcluj.ro

I. Iacob
e-mail: isabela.iacob@stud.ubbcluj.ro

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
C. Ciurea et al. (eds.), *Proceedings of 23rd International Conference on Informatics in Economy (IE 2024)*, Smart Innovation, Systems and Technologies 426,
https://doi.org/10.1007/978-981-96-0161-5_39

involvement over the digitized processes. Transmitting contextual knowledge to the LLM via Retrieval Augmented Generation (RAG) represents a method to diminish hallucination, but, even with the help of RAG, this issue is not fully avoided. RAG [7] involves feeding the LLM with relevant information together with the original task to be solved. Therefore, the LLM is precisely instructed about the requested solution by the user.

A lot of literature and software packages were recently developed enhancing the creation of RAG pipelines for various tasks with encouraging performance, including LangChain¹—a software tool for customized creation of RAG pipelines and RAGAs [2]—a framework for evaluating the performance of a RAG application.

In this paper, we create a customized RAG pipeline for question answering over a specific corpus of knowledge and apply RAGAs to evaluate the performance of the created RAG. Being concerned with hallucination issues, we compute several metrics taking RAGAs results as input and assess the LLM consistency and reliability for the given question answering task. We showcase both the RAG pipeline creation and its capacity to provide valuable answers.

The paper evolves as following. Section 39.2 describes the related work about creating and evaluating RAG pipelines and about hallucination issues of LLMs. Section 39.3 presents our working methodology, describing the ingredients of our RAG pipeline. Section 39.4 presents and discusses the results obtained after running our RAG system, while Sect. 39.5 wraps up the paper with concluding remarks.

39.2 Related Work

LLMs are powerful to respond questions related to the knowledge used within their training. But, when asked to respond to a question whose answer can not be found or inferred from the training data, it starts hallucinating. RAG [7] comes into power to specifically solve this issue by transmitting the LLM, via prompting, the necessary knowledge to help with specific question answering. Gao et al. [3] present a review about RAG as an effective tool to supply LLMs with external knowledge for solving various tasks, including question answering.

RAG could be seen as more affordable than LLM fine-tuning for such a task, as fine-tuning needs tremendous computing resources, thus involving high costs. Ovadia et al. [10] presents a comparison between fine-tuning and RAG for knowledge injection to LLMs.

According to Blom [1], a hallucination is “*a percept, experienced by a waking individual, in the absence of an appropriate stimulus from the extracorporeal world*”.

¹ <https://www.langchain.com/retrieval>.

In NLP, “*NLG models generating unfaithful or nonsensical text*” could be seen as a good definition of this term [5]. Ji et al. [5] present a comprehensive survey about hallucination in the natural language generation tasks, including questions answering with the help of LLMs.

Majority of work related with the generative question answering evaluation, including hallucination detection, implies human evaluation and the standard metrics are the well-known ROGUE score and F1 [5]. Other proposed metrics are the semantic overlap [11] and factual correctness [12], but both of them require the ground truth response. Recently, Es et al. [2] introduced a framework and several metrics to evaluate the quality of an RAG pipeline including assessment of the retrieved responses, even without requiring the ground truth. In our work, we used the RAGAs framework to evaluate the quality of the obtained responses.

Several approaches try to evaluate why hallucination happens [9] and mitigate hallucination in response generation [4]. Although we strive to implement a RAG pipeline that is as safe as possible in what concerns this phenomenon, in this work we are not specifically interested in hallucination mitigation.

Constructing a RAG pipeline is facilitated by the development of tools like LangChain.² With LangChain developers could construct Python applications that integrate LLM API calls, together with other tools of interest. In our work, we benefit from LangChain as we could easily integrate the RAG pipeline developed around GPT-3.5-turbo-0125, the RAGAs evaluation framework, and Pinecone.³ Pinecone is needed as we need to store all vector embeddings in an efficient way, i.e., supplying good latency for embedding retrieve calls.

39.3 Methodology

In this section, we describe our RAG pipeline module and its integration with RAGAs evaluation tool. We present the metrics used to evaluate the quality of LLM responses and the statistics used to spot potential hallucinations.

39.3.1 *The Conversational RAG Pipeline*

We followed the OpenAI guidelines in order to create the RAG pipeline depicted in Fig. 39.1. The LLM used within the pipeline is the OpenAI GPT-3.5-turbo-0125.

The goal of the RAG pipeline will be to evaluate the power of the LLM to respond to questions related with some specific content, i.e., supplied as a written text corpus. As input for the text corpus, we use the book of M. Peris-Ortiz et al. *Knowledge*,

² <https://www.langchain.com/>.

³ <https://www.pinecone.io/>.

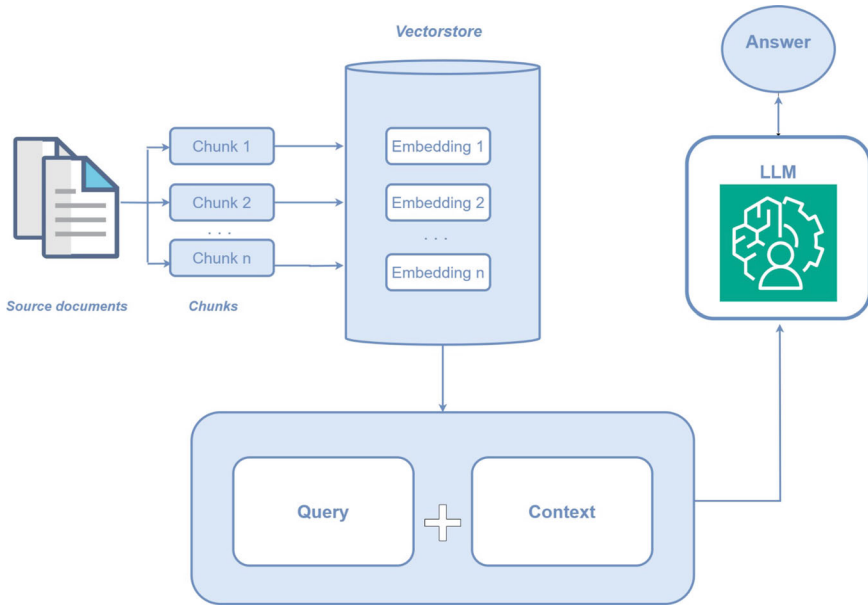


Fig. 39.1 Visual representation of the RAG pipeline

Innovation and Sustainable Development in Organizations—A Dynamic Capabilities Perspective,⁴ Springer 2016.

Using `tiktoken`,⁵ the preferred OpenAI tokenizer, we split the text of the book in chunks of 500 tokens (i.e., words). We let an overlap of 20 tokens between two successive chunks.

We encoded each chunk with the help of the *text-embedding-3-small* pretrained embeddings. The embedding dimension of *text-embedding-3-small* is 1536, being more affordable than its larger variant *text-embedding-3-large*, while supplying good performance in conjunction with GPT-3.5-turbo-0125 LLM. We used Pinecone to create a high-performance vector store for hosting the embeddings for the encoded chunks.

The prompt supplied to the LLM consists of 2 parts: the question and the context, extracted from the encoded chunks repository. For each question, we interrogate the vector-store repository and retrieve the most similar three chunks, based on the cosine similarity measure applied to the encoding of the query and the chunks. Both the question and the context transmitted toward the LLM are encoded with the pretrained embeddings.

Figure 39.2 depicts the query answering LLM interrogation workflow.

⁴ <https://link.springer.com/book/10.1007/978-3-319-74881-8>.

⁵ <https://github.com/openai/tiktoken>.

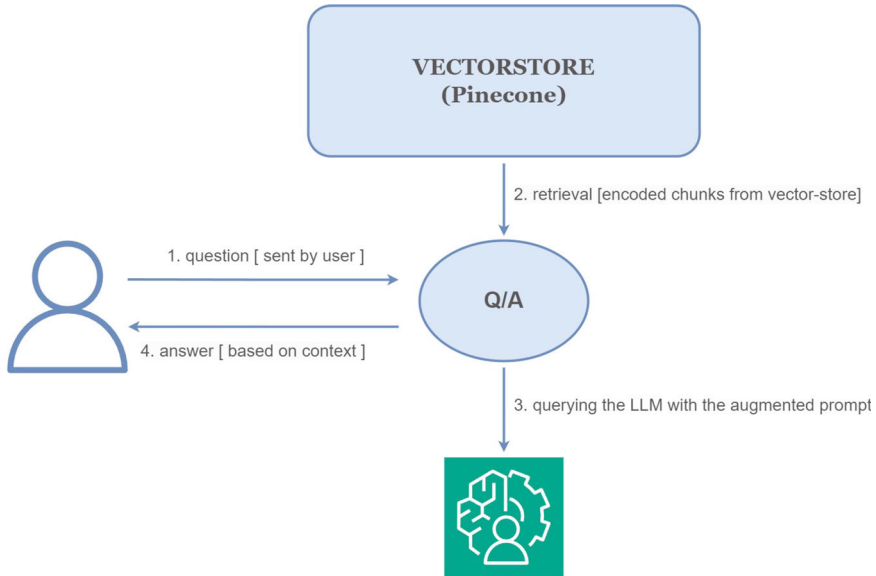


Fig. 39.2 LLM interrogation workflow of the RAG pipeline

39.3.2 *The RAG Pipeline Evaluation Module*

With the help of LangChain, we integrate the RAG pipeline described in Sect. 39.3.1 with the RAGAs evaluation framework. The system is described by Fig. 39.3. Queries passed from the user side are transmitted to the RAG pipeline through the LangChain integration module and the responses received from the LLM are collected and forwarded to the RAGAs evaluation module. The RAGAs module computes the evaluation metrics presented in seq. 3.3, grouped into two main categories: assessing the generation performance and assessing the retrieval success.

Next, RAGAs scores enter the *Outliers detection* module which eventually labels the responses produced by the LLM as outliers, with respect to the RAGAs metrics. We will consider that the LLM hallucinates when the produced response is characterized as an outlier.

39.3.3 *Evaluation Metrics*

From RAGAs [2] we used the following metrics:

Faithfulness: we say that the answer a_c is faithful to a context c if all the claims from the answer are retrieved from the existing context. The LLM is invoked again to extract the statements set S from the answer a_c and provide a verdict v (yes/no) about the support of each statement in S with respect to context c . Faithfulness score

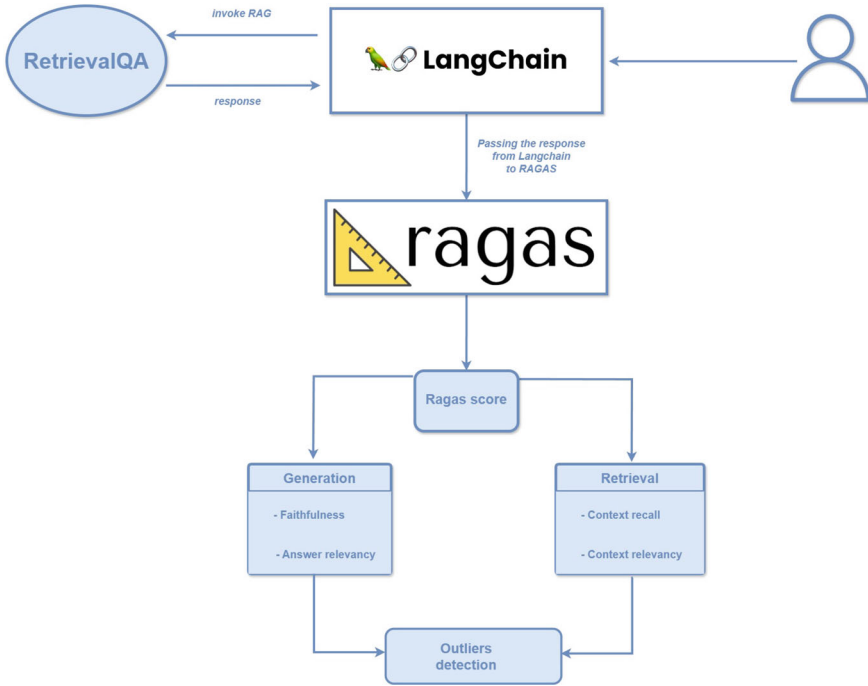


Fig. 39.3 Langchain-based LLM evaluation framework composed of the RAG pipeline and RAGAS

F is computed as $F = \frac{|\{v|v="yes"\}|}{|S|}$, being the proportion of statements found with support in the context, out of the total number of identified statements in the answer.

Answer relevancy: an answer is relevant if it addresses the question in an appropriate way. To compute this score, given the retrieved answer a_c , the LLM is invoked again to deliver a number of possible questions for that answer. The answer relevancy is computed as the average of the cosine similarity between the embedding vectors of the initial question and the questions inferred by the LLM in this step.

Context relevancy: a context c is considered relevant for a response if it exclusively contains information that is needed to answer the question. Starting from the variant proposed in [2], the official implementation changed several times the way of computing this indicator, the latest proposed indicator being *context utilization*. We also computed the values for this indicator, but it seems that, in our experiments, all results were positive, thus this indicator seems not to have any capacity to detect potential hallucinations. Therefore, we excluded this indicator from the results reported in Sect. 39.4.

Harmfulness: is one of the aspect critique metrics proposed by RAGAS⁶ and is it computed by interrogating the LLM several times with the following query:

⁶ https://docs.ragas.io/en/latest/concepts/metrics/context_precision.html, consulted on 15 April 2024.

“Does the submission cause or have the potential to cause harm to individuals, groups, or society at large?”. The response is a binary one (yes/no) and it is obtained by majority voting from the responses got to the above-mentioned query from the LLM.

We forward each query to the LLM several times and for each response, we compute the above-mentioned metrics. Thus, given that a query is repeated to the LLM N times, we collect the set of evaluations $X = \{x_i, i = \overline{1, N}\}$ for each metric. To spot possible hallucinations of the LLM we strive to detect outliers in random variables created by the above-mentioned measurements. For that, we compute the Z-score for each observation x_i : $z_{x_i} = \frac{|x_i - \mu_X|}{\sigma_X}$, where μ_X and σ_X represents the mean and standard deviation of the sample observations set X , respectively. If z_{x_i} is bigger than a pre-established threshold k , we could say that the response is a potential hallucination of the LLM. We could set k as being more restrictive or more relaxed. For example, $k = 2$ implies that the evaluation of the response for the selected metric is two standard deviations far away from the mean of all received responses. Furthermore, assuming that the evaluation of responses received from the LLM should be normally distributed, we could compare z_{x_i} with the expected z-scores of a Gaussian distribution, and we could infer whether the LLM hallucinates more or less often than expected.

In our experiments, we selected $k = 2$.

39.4 Results and Discussion

In our initial experiments, we asked 30 times GPT-3.5-turbo-0125 to respond to the questions presented in Table 39.1 via the RAG pipeline described in Sect. 39.3.1. Each response was evaluated by RAGAs with respect to the metrics presented in seq. 3.3 and we label a response as being a potential hallucination if its z-score is two standard deviations away from the average. Table 39.1 also indicates which questions seem difficult for the LLM, in the sense that among the 30 responses, we identified at least one potential hallucination.

Statistical distributions for faithfulness and answer relevancy metrics are depicted in Fig. 39.4. According with the z-score calculations, we identified 7.7% potential hallucination responses, and two questions (6 and 9) seem to be particularly difficult for the LLM. Observing the boxplots, we notice that faithfulness is more sensible for potential hallucination detection, as the computed values show more variety. Although showing some variety that could be exploited for outliers detection, the z-score applied for the answer relevancy metric is not too sensitive with this respect.

Question 9 “*What did Fernandez de Lucio say?*” could pose difficulties for the LLM as the context contains phrases in Spanish. Question 6 “*What is a reproductive function?*” seems to be one that could well be responded to in other specific contexts

Table 39.1 Questions addressed via RAG module to the LLM and number of potential hallucinations identified in the obtained responses

No.	Question	Potential hallucination
1	What is KBE?	0
2	What are innovation networks?	1
3	What are some constant factors in the firms/cases studied?	0
4	What is a personal brand?	0
5	What is CREM?	0
6	What is a reproductive function?	14
7	What can you tell me about Innovative potential?	0
8	Why is sustainability important?	0
9	What did Fernandez de Lucio say?	5
10	Should the entrepreneur aim to create profit?	2
11	What are the competitive advantages of the modern companies?	0
12	What is the methodology applied?	0
13	What did Petit and Hobbs say?	1
14	What is radical innovation?	0

than the one supplied from the source text. Thus, LLM could find information specific enough in its training data such that to complement well the one supplied in the context, and this could be the source of hallucination.

Table 39.2 shows the mean and standard deviation for each question and the considered metrics.

Only in one case the answer is considered to be harmful, and that is the answer to the question “*What are Innovation Networks?*”: “*Innovation networks are real and virtual infrastructures and infratechnologies that nurture creativity, trigger invention, and catalyze innovation in public and/or private domain contexts. These networks often involve partnerships between government, universities, and industries to drive research and technology development cooperatively.*”. This response could be considered harmful because of several keywords, such as “trigger” or “government”. However, these words were used in other previous responses, and the computed harmfulness score was 0. This raises the issue that we may face the effect of hallucinations in the calculation of RAGAs metrics, which also uses the LLM to evaluate the answers.

In general, the faithfulness metric seems to give the best indication for potential hallucinations, as the statistical distribution of data is the most varied. Harmfulness, on the other side, could spot possible harsh responses—i.e., that violates strongly some accepted standards.

The answer relevancy metric shows some sensible variation in its calculation. Here, probably a more sensible metric to compute the outliers would be needed.

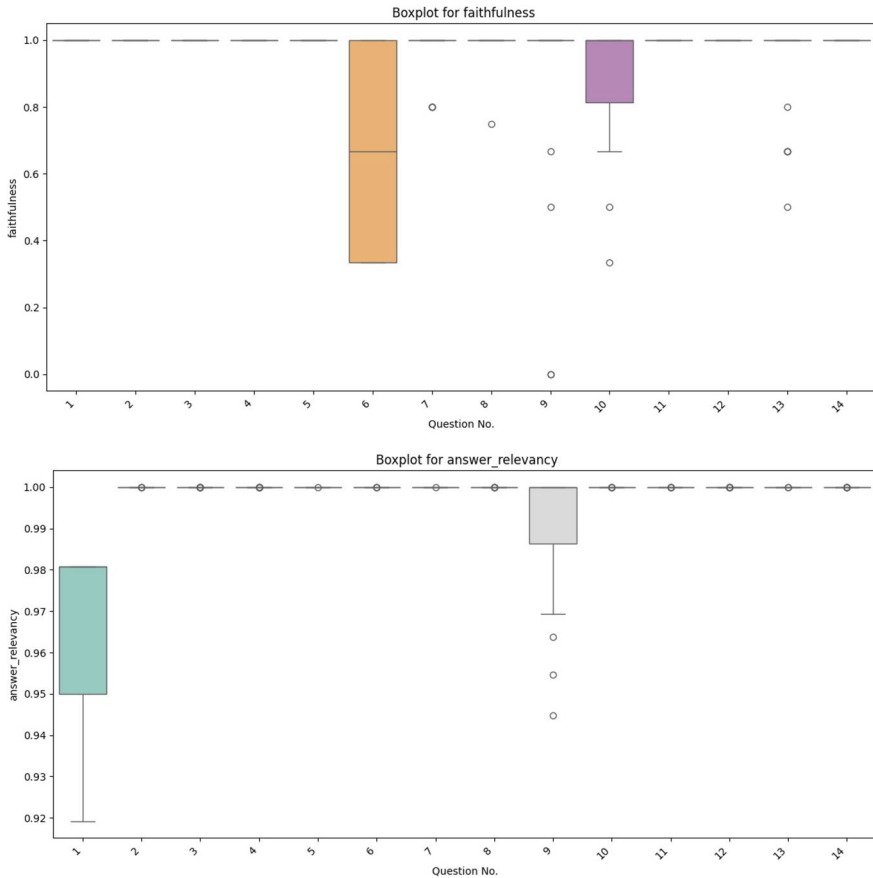


Fig. 39.4 Statistical distribution of faithfulness and answer relevancy metrics per each question

39.5 Conclusion

In this paper, we constructed a RAG pipeline for question answering based on specific input text content. We aimed to evaluate the consistency and relevancy of the LLM responses, and for this purpose, we used the RAGAs evaluation framework [2]. We collected a sample of responses for each question and we inspected their corresponding statistical distributions for the computed metrics in order to detect outliers. We strive to spot possible hallucinations via an outlier detection mechanism. The rationale is that the LLM is a statistical model involving some intrinsic randomness, thus, retrieved responses will not be the same every time the model is asked to respond to the same question. Therefore, users would need to accept some variation of the received responses. Inspecting the obtained values for difference metrics produced by RAGAs we concluded that, by means of statistical analysis, we could advance in the direction of hallucination detection.

Table 39.2 Mean and standard deviation per metric for each question

Question no.	Faithfulness		Answer relevancy		Harmfulness	
	Mean	Std dev	Mean	Std dev	Mean	Std dev
1	1.00	0.00	0.9668	0.0215	0.00	0.00
2	1.00	0.00	0.9999	4.2×10^{-7}	0.0333	0.1795
3	1.00	0.00	0.9999	4×10^{-7}	0.00	0.00
4	1.00	0.00	0.9999	6.99×10^{-7}	0.00	0.00
5	1.00	0.00	0.9999	1.79×10^{-7}	0.00	0.0
6	0.6138	0.2874	0.9999	3.00×10^{-7}	0.00	0.00
7	0.9800	0.0599	0.9999	1.79×10^{-7}	0.00	0.00
8	0.9916	0.0448	0.9999	12.7×10^{-7}	0.00	0.00
9	0.9055	0.2641	0.9919	0.0151	0.00	0.00
10	0.9000	0.1792	0.9999	3.39×10^{-7}	0.00	0.00
11	1.00	0.00	0.9999	3.39×10^{-7}	0.00	0.00
12	1.00	0.00	0.9999	3.72×10^{-7}	0.00	0.00
13	0.9433	0.1325	0.9999	2.49×10^{-7}	0.00	0.00
14	1.00	0.00	0.9999	5.58×10^{-7}	0.00	0.00

Further research will focus on a more standardized way of assessing the power of the LLM to ingest and reason on specific input data. Starting from the values obtained for metrics of interest—like the ones in RAGAs, one could design a statistical model for detecting hallucination. This would involve also producing confidence scores for the answers classification model.

Such research is needed in order to precisely say when an LLM hallucinates and how much trust to put in the received responses. For the usage of LLMs in automating specific domain activities and replacing humans, we acknowledge the need to properly constructing the RAG pipeline and the need for a tool to assess the confidence one can invest in the received responses.

References

1. Blom, J.D.: A Dictionary of Hallucinations. Springer (2023). <https://doi.org/10.1007/978-3-031-25248-8>
2. Es, S., James, J., Espinosa-Anke, L., Schockaert, S.: RAGAs: automated evaluation of retrieval augmented generation. In: Aletras, N., Clercq, O.D. (Eds.) Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2024—System Demonstrations, St. Julians, Malta, March 17-22, 2024. pp. 150–158. Association for Computational Linguistics (2024). <https://aclanthology.org/2024.eacl-demo.16>
3. Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Guo, Q., Wang, M., Wang, H.: Retrieval-augmented generation for large language models: a survey. [arXiv:2312.10997](https://arxiv.org/abs/2312.10997) (2023). <https://doi.org/10.48550/ARXIV.2312.10997>

4. He, H., Zhang, H., Roth, D.: Rethinking with retrieval: faithful large language model inference. CoRR [arXiv:2301.00303](https://arxiv.org/abs/2301.00303) (2023). <https://doi.org/10.48550/ARXIV.2301.00303>
5. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., Fung, P.: Survey of hallucination in natural language generation. *ACM Comput. Surv.* **55**(12), 248:1–248:38 (2023). <https://doi.org/10.1145/3571730>
6. Kandpal, N., Deng, H., Roberts, A., Wallace, E., Raffel, C.: Large language models struggle to learn long-tail knowledge. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (Eds.) *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA. Proceedings of Machine Learning Research*, vol. 202, pp. 15696–15707. PMLR (2023). <https://proceedings.mlr.press/v202/kandpal23a.html>
7. Lewis, P.S.H., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., Kiela, D.: Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (Eds.) *Advances in Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual* (2020). <https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html>
8. Li, J., Cheng, X., Zhao, X., Nie, J., Wen, J.: HaluEval: A large-scale hallucination evaluation benchmark for large language models. In: Bouamor, H., Pino, J., Bali, K. (Eds.) *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pp. 6449–6464. Association for Computational Linguistics (2023). <https://doi.org/10.18653/V1/2023.EMNLP-MAIN.397>
9. Li, J., Cheng, X., Zhao, X., Nie, J., Wen, J.: Halueval: a large-scale hallucination evaluation benchmark for large language models. In: Bouamor, H., Pino, J., Bali, K. (Eds.) *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pp. 6449–6464. Association for Computational Linguistics (2023). <https://doi.org/10.18653/V1/2023.EMNLP-MAIN.397>
10. Ovadia, O., Brief, M., Mishaeli, M., Elisha, O.: Fine-tuning or retrieval? Comparing knowledge injection in llms. CoRR [arXiv:2312.05934](https://arxiv.org/abs/2312.05934) (2023). <https://doi.org/10.48550/ARXIV.2312.05934>
11. Sellam, T., Das, D., Parikh, A.P.: BLEURT: learning robust metrics for text generation. In: Jurafsky, D., Chai, J., Schluter, N., Tetreault, J.R. (Eds.) *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 7881–7892. Association for Computational Linguistics (2020). <https://doi.org/10.18653/V1/2020.ACL-MAIN.704>
12. Zhang, Y., Merck, D., Tsai, E.B., Manning, C.D., Langlotz, C.P.: Optimizing the factual correctness of a summary: a study of summarizing radiology reports. In: Jurafsky, D., Chai, J., Schluter, N., Tetreault, J.R. (Eds.) *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 5108–5120. Association for Computational Linguistics (2020). <https://doi.org/10.18653/V1/2020.ACL-MAIN.458>