UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Methods and Applications of Data Mining in Business Domains

Amrit, C.; Abdi, A.

DOI
10.3390/app131910774
Publication date
2023
Document Version
Final published version
Published in
Applied Sciences (Switzerland)
License
CC BY

Link to publication

Citation for published version (APA):
Amrit, C., & Abdi, A. (2023). Methods and Applications of Data Mining in Business Domains. *Applied Sciences (Switzerland)*, *13*(19), Article 10774. https://doi.org/10.3390/app131910774

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Oct 2026

applied sciences

MDPI

Editorial

Methods and Applications of Data Mining in Business Domains

Chintan Amrit [1,2,*] and Asad Abdi [3]

[1] Department of Operations Management, Amsterdam Business School, University of Amsterdam, P.O. Box 15953, 1001 NL Amsterdam, The Netherlands
[2] LM Thapar School of Management, Dera Bassi Campus, Chandigarh 140507, India
[3] Department of Computing and Mathematics, Faculty of Science and Engineering, University of Derby, Derby DE22 1GB, UK; a.abdi@derby.ac.uk
* Correspondence: c.amrit@uva.nl

1. Introduction

This Special Issue invited researchers to contribute original research in the field of data mining, particularly in its application to diverse domains, like healthcare, software development, logistics, and human resources. We were especially interested in how the data mining method was modified to cater to the specific domain in question. The challenge is that the more complex a domain is the harder it is to make good predictions, as more implicit domain knowledge is required that is not always available [1]. This is especially true in the case of complex domains where there are soft factors, like the interaction of the conflicting and cooperating objectives of the stakeholders [2,3], and system dynamics play a significant role [4]. In a business context, the challenge is that one would like to see (i) how the algorithms can be repeatable in the real world, (ii) how the patterns mined can be utilized by the business, and (iii) how the resulting model can be understood and utilized in the business environment [1]. Furthermore, the idea is to identify the variables that impact the goal variable but to do so with the data, interestingness, deployment, and general domain (business) constraints of the domain [1,5].

One of the methods to analyze a complex domain is using a method called intelligence meta-synthesis [6,7]. Intelligence synthesis is the collection and creation of perceived or understood (i.e., not necessarily objective) information. Meta-synthesis is the collection and creation of knowledge and information from collected intelligences [1]. The goal of this approach is to design and develop predictive models that could eventually be incorporated into a business intelligence dashboard. As a result, one would (i) understand the nature and origin of data that allows the system user to determine the quality of the data to perform the data cleaning; (ii) understand the factors in the domain that influence the predicted variable, leading the developer to determine which variables need to be included in the predictive model; (iii) develop predictive models that are usable and interesting within the domain in terms of predictive power, integrating with existing infrastructure, and integrating with business rules and processes; and finally (iv) use the predicted data to find the optimal business processes in the particular domain. There are also research works that have built on top of intelligence meta-synthesis, such as the study published by the authors of [1].

The main goal of this Special Issue was to bring together researchers, participants, academic scientists, and contributors to share their experiences, present, and discuss their ongoing and latest research results that cover several aspects of original research concerning existing theoretical and methodological contributions, as well as the development of new methods/approaches in the field of data mining in different business domains.

check for updates

Citation: Amrit, C.; Abdi, A. Methods and Applications of Data Mining in Business Domains. *Appl. Sci.* **2023**, *13*, 10774. https://doi.org/10.3390/app131910774

Received: 22 September 2023
Accepted: 26 September 2023
Published: 28 September 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

2. Categorized Overview of Papers (Based on the Areas of Focus)

2.1. Category 1: Retail and Customer Analysis

These papers collectively contribute to the field of data-driven applications and algorithm development. The first paper presents a data-driven recommendation system for retail, emphasizing tailored product recommendations and marketing strategies [8]. The second paper addresses the challenge of handling imbalanced data streams with concept drift, introducing the HSDW-MI algorithm [9]. The third paper focuses on electric load classification in industrial scenarios, highlighting the effectiveness of an equilibrium optimizer-based feature selection method [10]. These papers collectively demonstrate the versatility and impact of data-driven approaches in enhancing decision making and addressing real-world challenges across various domains.

2.2. Category 2: Marketing and Business Decision Support

These papers collectively contribute to the field of data-driven applications and predictive modeling across various domains, like a Decision Support System for the Marketing domain [11]. They offer an advanced model for detecting financial statement fraud using XGBoost [12], introduce innovative neural network models for stock price prediction [13], analyze factors influencing tourist offer prices [14], develop predictive models for healthcare patient influx [15], and propose intelligent decision forest models for customer churn prediction in the telecom industry [16]. They also address customer churn prediction in non-contractual B2B settings [17], improve legal judgment prediction through graph neural networks [18], enhance car sales forecasts using online sentiment data and deep learning [19], introduce a reinforcement learning framework for options trading [20], and predict the charge of a legal case using a novel graph convolutional network [21]. These studies showcase the versatility and practical applications of data-driven techniques in diverse fields, underscoring their importance for informed decision making and predictive accuracy.

2.3. Category 3: Business Process Optimization and Automation

In addition to these papers collectively contributing to the field of data-driven applications and predictive modelling across a variety of domains, they offer specific approaches, such as innovative approaches to optimizing product layouts in supermarkets using sequential pattern mining and optimization algorithms [22]. Furthermore, they introduce a novel supervised learning algorithm for financial risk assessment [23] and explore the development trajectory of radio frequency identification (RFID) applications through academic citation and text mining analysis [24]. Another paper presents a deep learning framework for predicting important trading points in the stock market [25], focusing on high-margin opportunities. Additionally, they provide insights into factors influencing the adoption of business intelligence systems (BISs) in small and medium-sized enterprises (SMEs) [26], and conduct a systematic literature review on AI-based methods for automating business processes and decision support [27]. These studies collectively demonstrate the versatility and practical applications of data-driven techniques across diverse fields, highlighting their significance for informed decision making and optimization.

3. Overview of Significant Findings That Have Emerged from the Papers

Across these papers, the common themes and trends of artificial intelligence, technology, and data-driven applications have emerged. Predictive analytics and machine learning techniques, such as deep learning and neural networks, are prevalent for applications like stock price prediction, demand forecasting, and decision automation. Data-driven decision making has been emphasized, with data analysis and mining playing a crucial role in improving processes. Challenges related to data quality and integration have been highlighted, especially in the healthcare and manufacturing contexts. Natural language processing (NLP) was identified as a recurring theme, applied for sentiment analysis, text mining, and chatbots. Supply chain optimization and ethical considerations in AI applications are also prominent, as are hybrid AI models and the significance of data visualization. The

automation of business processes aligns with the broader trend of digital transformation. In summary, these papers showcase the broad adoption of AI and data-driven methods, emphasizing data quality, ethics, and AI's potential to enhance decision making and optimize processes while addressing challenges like data integration and privacy concerns.

4. Overall Perspective

This collection of papers spans diverse topics in the field of data mining and its applications in business. Together, they significantly contribute to this field by addressing various challenges and providing solutions in domains, such as recommendation systems for retail, handling multiclass imbalanced data streams, feature selection for industrial load classification, data-driven marketing decision support systems, financial statement fraud detection, stock price prediction, factors influencing prices, healthcare resource management, customer churn prediction, legal judgment prediction, supermarket product layout optimization, options trading with reinforcement learning, few-shot charge prediction, RFID applications, stock trading point prediction, business intelligence adoption in SMEs, and AI-based methods for business processes. These papers collectively showcase the adaptability and effectiveness of data mining techniques, making substantial contributions to the broader realm of "*Methods and Applications of Data Mining in Business Domains*".

5. Conclusions

These diverse papers within this Special Issue offer valuable insights and innovations across various domains. Readers are encouraged to explore the full papers to delve deeper into these research findings. In the field of retail, innovative recommendation systems and customer segmentation techniques promise to enhance sales and customer loyalty. Furthermore, novel algorithms for handling imbalanced data streams bring robustness to multiclass classification in dynamic environments. In the industrial sector, precision in electric load classification is emphasized, and the application of equilibrium optimization showcases potential advancements in energy management. For finance professionals, the introduction of dynamic nearest neighbor classifiers hints at improved risk assessment methods. Meanwhile, researchers can gain critical insights into the evolving landscape of radio frequency identification technology. In the realm of stock market predictions, the hybrid convolutional recurrent neural network and deep learning frameworks offer potential avenues for investors. Lastly, the systematic literature review on AI-based methods in business processes highlights the importance of automation and decision support systems in the corporate world, paving the way for future advancements. Although we have included 20 research articles, we think that the challenge of applying data analytics to different business domains is a nascent field that requires further exploration, with clear guidelines and insights into the applicable methods and techniques for not only different domains but also across domains.

Author Contributions: A.A.: writing—original draft preparation, C.A.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. van der Spoel, S. Prediction Instrument Development for Complex Domains. Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 2016.
2. van der Spoel, S.; Amrit, C.; van Hillegersberg, J. Predictive analytics for truck arrival time estimation: A field study at a European distribution centre. *Int. J. Prod. Res.* **2017**, *55*, 5062–5078. [CrossRef]
3. Checkland, P.B. Soft systems methodology. *Hum. Syst. Manag.* **1989**, *8*, 273–289. [CrossRef]
4. Forrester, J.W. System dynamics, systems thinking, and soft OR. *Syst. Dyn. Rev.* **1994**, *10*, 245–256. [CrossRef]
5. Cao, L.; Yu, P.S.; Zhang, C.; Zhao, Y. *Domain Driven Data Mining*; Springer: Berlin/Heidelberg, Germany, 2010.
6. Gu, J.; Tang, X. Meta-synthesis approach to complex system modeling. *Eur. J. Oper. Res.* **2005**, *166*, 597–614. [CrossRef]
7. Cao, L. Domain-driven data mining: Challenges and prospects. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 755–769. [CrossRef]

8. Chen, A.H.-L.; Gunawan, S. Enhancing Retail Transactions: A Data-Driven Recommendation Using Modified RFM Analysis and Association Rules Mining. *Appl. Sci.* **2023**, *13*, 10057. [CrossRef]
9. Han, M.; Li, A.; Gao, Z.; Mu, D.; Liu, S. Hybrid Sampling and Dynamic Weighting-Based Classification Method for Multi-Class Imbalanced Data Stream. *Appl. Sci.* **2023**, *13*, 5924. [CrossRef]
10. Zhou, M.; Yao, X.; Zhu, Z.; Hu, F. Equilibrium Optimizer-Based Joint Time-Frequency Entropy Feature Selection Method for Electric Loads in Industrial Scenario. *Appl. Sci.* **2023**, *13*, 5732. [CrossRef]
11. Hou, R.; Ye, X.; Zaki, H.B.O.; Omar, N.A.B. Marketing Decision Support System Based on Data Mining Technology. *Appl. Sci.* **2023**, *13*, 4315. [CrossRef]
12. Ali, A.A.; Khedr, A.M.; El-Bannany, M.; Kanakkayil, S. A Powerful Predicting Model for Financial Statement Fraud Based on Optimized XGBoost Ensemble Learning Technique. *Appl. Sci.* **2023**, *13*, 2272. [CrossRef]
13. Li, C.; Qian, G. Stock Price Prediction Using a Frequency Decomposition Based GRU Transformer Neural Network. *Appl. Sci.* **2022**, *13*, 222. [CrossRef]
14. Kołakowska, A.; Godlewska, M. Analysis of Factors Influencing the Prices of Tourist Offers. *Appl. Sci.* **2022**, *12*, 12938. [CrossRef]
15. Cubillas, J.J.; Ramos, M.I.; Feito, F.R. Use of Data Mining to Predict the Influx of Patients to Primary Healthcare Centres and Construction of an Expert System. *Appl. Sci.* **2022**, *12*, 11453. [CrossRef]
16. Usman-Hamza, F.E.; Balogun, A.O.; Capretz, L.F.; Mojeed, H.A.; Mahamad, S.; Salihu, S.A.; Akintola, A.G.; Basri, S.; Amosa, R.T.; Salahdeen, N.K. Intelligent Decision Forest Models for Customer Churn Prediction. *Appl. Sci.* **2022**, *12*, 8270. [CrossRef]
17. Mirkovic, M.; Lolic, T.; Stefanovic, D.; Anderla, A.; Gracanin, D. Customer Churn Prediction in B2B Non-Contractual Business Settings Using Invoice Data. *Appl. Sci.* **2022**, *12*, 5001. [CrossRef]
18. Zhao, Q.; Gao, T.; Zhou, S.; Li, D.; Wen, Y. Legal Judgment Prediction via Heterogeneous Graphs and Knowledge of Law Articles. *Appl. Sci.* **2022**, *12*, 2531. [CrossRef]
19. Ou-Yang, C.; Chou, S.-C.; Juan, Y.-C. Improving the Forecasting Performance of Taiwan Car Sales Movement Direction Using Online Sentiment Data and CNN-LSTM Model. *Appl. Sci.* **2022**, *12*, 1550. [CrossRef]
20. Wen, W.; Yuan, Y.; Yang, J. Reinforcement Learning for Options Trading. *Appl. Sci.* **2021**, *11*, 11208. [CrossRef]
21. Wang, P.; Zhang, X.; Cao, Z. Few-Shot Charge Prediction with Data Augmentation and Feature Augmentation. *Appl. Sci.* **2021**, *11*, 10811. [CrossRef]
22. Kaewyotha, J.; Songpan, W. Multi-Objective Design of Profit Volumes and Closeness Ratings Using MBHS Optimizing Based on the PrefixSpan Mining Approach (PSMA) for Product Layout in Supermarkets. *Appl. Sci.* **2021**, *11*, 10683. [CrossRef]
23. Camacho-Urriolagoitia, O.; López-Yáñez, I.; Villuendas-Rey, Y.; Camacho-Nieto, O.; Yáñez-Márquez, C. Dynamic Nearest Neighbor: An Improved Machine Learning Classifier and Its Application in Finances. *Appl. Sci.* **2021**, *11*, 8884. [CrossRef]
24. Su, W.-H.; Chen, K.-Y.; Lu, L.Y.Y.; Wang, J.-J. Knowledge Development Trajectories of the Radio Frequency Identification Domain: An Academic Study Based on Citation and Main Paths Analysis. *Appl. Sci.* **2021**, *11*, 8254. [CrossRef]
25. Yu, X.; Li, D. Important Trading Point Prediction Using a Hybrid Convolutional Recurrent Neural Network. *Appl. Sci.* **2021**, *11*, 3984. [CrossRef]
26. Alsibhawi, I.A.A.; Yahaya, J.B.; Mohamed, H.B. Business Intelligence Adoption for Small and Medium Enterprises: Conceptual Framework. *Appl. Sci.* **2023**, *13*, 4121. [CrossRef]
27. Gomes, P.; Verçosa, L.; Melo, F.; Silva, V.; Filho, C.B.; Bezerra, B. Artificial Intelligence-Based Methods for Business Processes: A Systematic Literature Review. *Appl. Sci.* **2022**, *12*, 2314. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.