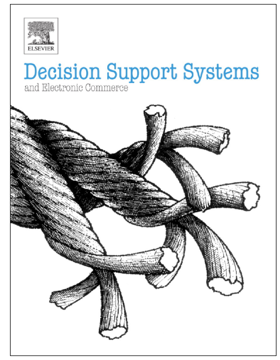
Journal Pre-proof



Explainable AI for enhanced decision-making

Kristof Coussement, Mohammad Zoynul Abedin, Mathias Kraus, Sebastian Maldonado, Kazim Topuz

PII: S0167-9236(24)00109-X

DOI: https://doi.org/10.1016/j.dss.2024.114276

Reference: DECSUP 114276

To appear in: *Decision Support Systems*

Please cite this article as: K. Coussement, M.Z. Abedin, M. Kraus, et al., Explainable AI for enhanced decision-making, *Decision Support Systems* (2024), https://doi.org/10.1016/j.dss.2024.114276

Explainable AI for Enhanced Decision-Making

Kristof Coussement[a], Mohammad Zoynul Abedin[b], Mathias Kraus[c], Sebastian Maldonado[d], Kazim Topuz[e]

[a]*IESEG School of Management Univ. Lille CNRS UMR 9221 - LEM - Lille Economie Management 3 Rue de la Digue 59000 Lille France*
[b]*Swansea University Department of Accounting and Finance School of Management Bay Campus Fabian Way Swansea SA1 8EN Wales United Kingdom* [c]*Universitt Regensburg Fakultt fr Informatik und Data Science Bajuwarenstrae 4 93053 Regensburg Germany*
[d]*University of Chile Department of Management Control and Information Systems School of Economics and Business Santiago Chile*
[e]*The University of Tulsa Collins College of Business School of Finance and Operations Management Tulsa Oklahoma USA*

Abstract

This paper contextualizes explainable artificial intelligence (AI) for enhanced decision-making and serves as an editorial for the corresponding special issue. AI is defined as the development of computer systems that are able to perform tasks that normally require human intelligence by understanding, processing, and analyzing large amounts of data. AI has been a dominant domain for several decades in the information systems (IS) literature. To this end, we define explainable AI (XAI) as the process that allows one to understand how an AI system decides, predicts, and performs its operations. First, we contextualize its current role for improved business decision-making. Second, we discuss three underlying dimensions of XAI that serve as broader innovation grounds to make better and more informed decisions, i.e., data, method, and application. For each of the contributing papers in this special issue, we describe their major contributions to the field of XAI for decision making. In conclusion, this paper further presents a future research agenda for IS researchers in the XAI field.

Keywords: explainable artificial intelligence, interpretability, visualizations

---

Preprint submitted to Decision Support Systems *June 10, 2024*

1. Introduction

Artificial Intelligence (AI) defined as the development of computer systems that are able to perform tasks normally requiring human intelligence by understanding, processing, and analyzing large amounts of data has been a prevalent domain in information systems (IS) for several decades now. An increasing number of businesses rely on AI to achieve outcomes that operationally and/or strategically support (human) decision making in different domains. At present, AI-based machine learning (ML) has become widely popular as a subfield of AI, both in academia as in industry. ML has been widely used to enhance decision making, including predicting organ transplantation risk [1], forecasting remaining useful life of machinery [2], student dropout prediction [3], disentangling human trafficking types [4], money laundering detection [5, 6], among others. In the early days, AI attempted to imitate human decision-making rules with only partial success, as humans often could not accurately describe the decision-making rules they use to solve problems [7]. With the development of advanced AI, exciting progress has been made in algorithmic development to support decision making in various fields including finance [8, 9], marketing [10, 11, 12], education [3], human resource management [13], tourism [14], aviation [15] and others [16].

Recently, advances have heavily focused on boosting the predictive accuracy of AI methods, with deep learning (DL) methods being a prevalent example. For instance, [17] use the text on corporate websites to improve business failure identification using advanced DL-based natural language processing (NLP) methods. The stringent focus on improved prediction performance often comes at the expense of reduced explainability, which leads to decision makers' distrust and even rejection of AI systems [18]. Explainable AI (XAI) describes the process that allows us to understand how an AI system decides, predicts, and performs its operations. Therefore, XAI reveals the strengths and weaknesses of the decision-making strategy and explains the rationale of the decision support system (DSS) [19]. Numerous scholars confirm that XAI is the key to developing and implementing AI in industries such as retail, banking and financial services, healthcare, manufacturing and supply chain [20, 18, 21, 22]. In addition, XAI has also received the attention of governments due to its ability to improve the efficiency and effectiveness of government functionalities and decision supports [23].

In fact, in many cases, understanding why a model makes certain decisions and predictions is as important as its accuracy. Because model explainability helps managers better understand model parameters and apply them more confidently, allowing managers to communicate the analytical rationale more convincingly for their decisions to stakeholders [24]. Among others, exploring the drivers of AI explainability and precise understandability in decision making is one of the main contributions of this special issue. In concrete terms, we explore three main domains of innovation that make up an XAI strategy that helps to ensure transparency during decision-making, that is, data, method and application.

Section 2 zooms in on the dimensions of XAI for enhanced decision making, while Section 3 provides an overview and summarizes the contributions included in this special issue. In addition, Section 4 concludes this article and discusses interesting future research directions in XAI for the IS field.

2. Dimensions of Explainable AI for Enhanced Decision-Making

2.1. Data

The link between XAI and different *types of data*, including observational data with structured and unstructured data, as well as synthetic data, is a vital aspect of the interpretability and transparency of the ML outcomes. Extant IS literature has already dived into often manually crafted variables from structured data that are inherently interpretable and directly improve managerial decision-making. For example, [25] designed direct understandable variables – linked to the user network, user behavior, message readability, and message structure – out of X social media posts to improve the identification of influential customers. However, existing IS literature is less lenient to converting unstructured – textual, audio, or image – data into directly interpretable features, and often the focus lies on optimizing the predictive performance for enhanced decision-making. Furthermore, experimental data plays a crucial role in XAI solutions by allowing controlled experimentation using real-life field testing or the creation of synthetic datasets (e.g. [26]. For example, in the field of causal ML, real-life field tests are a common tool for generating experimental data to check the beneficial impact of different treatments on the final outcome. For example, [10] tests the impact of three different motivational emails sent to members of the online community to measure whether motivational emails have an impact on member participation and answer the question of which motivational email works best for which member. [27] rely on empirical data gathered through experiments, emphasizing the importance of data quality for the validity of their serial mediation model. [28] particularly focusing on decision-making within sharing economy platforms in the Metaverse. It employs a case study to empirically validate its theoretical propositions, contributing to both the academic and practical understanding of the subject. The study contributes to the data aspect of XAI by employing fuzzy logic, specifically through the use of Schweizer-Sklar norms. This approach is designed to manage uncertainty in decision making within the Metaverse sharing economy. Although it does not revolutionize the field of XAI, it does offer a way to handle ambiguous or uncertain data in a specific application context. The synergy between XAI and these types of data is essential to build trust in AI systems, address bias and fairness concerns, and ensure that AI-based decisions are not only accurate but also explainable, interpretable, and actionable for stakeholders and end users. It also facilitates regulatory compliance and ethical considerations in AI applications in various domains.

2.2. Method

ML-based AI solutions require careful consideration when selecting the appropriate algorithms. In the realm of extant IS literature, there is a fundamental distinction between ”glass-box” and ”black-box” algorithms. Glassbox algorithms offer direct interpretability, providing decision makers with insights into the decision-making process. In particular, regression models and decision trees are prevalent examples of glass-box models found in the XAI literature. For example, [3] used a logistic regression model to reveal the key student metrics that influence dropout behavior. They leveraged the beta coefficients of their logistic regression model to visually illustrate the direction and impact of each metric on student attrition. In addition, the advent of Generalized Additive Models (GAMs) has introduced a nuanced perspective that aims to bridge the gap between model interpretability and predictive performance [29]. Particularly capable of handling structured data, GAMs and related spline-base methods have been applied to various prediction tasks, such as house price prediction [30] or customer churn prediction [31].

However, a prevailing assumption exists that black-box models, owing to their potentially superior predictive performance, lead to more informed decisions than their glass-box counterparts. Recent benchmarking studies have reinforced this perception, with ensemble methods, particularly XGBoost, garnering significant attention in the recent IS literature. [32] demonstrated the dominance of XGBoost in effectively predicting financial risks in supply chain finance. [33] find favorable results for XGBoost compared to other ML models in predicting the probability of risk and the frequency of claims in the context of automobile insurances. However, the trade-off is that these black-box models often compromise interpretability.

However, recent advances have been made to elucidate the inner workings of black-box models. Post-hoc interpretation tools such as SHAP [34], ICE [35], and LIME [36] now allow the examination of the impact of various drivers on final predictions. For example, [37] used the SHAP algorithm to provide intuitive insight into the context of readmissions from the emergency department of COVID-19. On the methodological front, [28] incorporates an adapted version of the Ordinal Priority Approach (OPA) into a fuzzy logic context. These methodological choices aim to improve the model's adaptability and transparency, although further research would be needed to generalize these methods to other contexts.

Furthermore, the choice of evaluation metrics plays a pivotal role in aligning AI solutions with the objectives of decision-makers. Historically, the IS literature has relied on statistical metrics such as accuracy, recall, precision, area under the ROC curve (AUC), and H-measure for predictive model evaluation. However, these metrics often lack direct interpretability and may not correlate with standard business metrics and key performance indicators. Hence, a recent surge in research has emerged within the field of profit analytics, focusing on optimizing and evaluating algorithms based on profit-oriented criteria. For example, [38] introduced ProfLogit, an adaptation of conventional logistic regression designed to maximize the expected maximum profit criterion (EMPC). In the context of churn prediction, their work demonstrated that ProfLogit outperforms other models, achieving the highest out-of-sample EMPC performance, as well as the best profit-based precision and recall values overall. In the field of marketing, [39] developed profit-conscious ensemble selection (PCES), demonstrating that the resulting measures offer a managerially meaningful assessment of the business value of a targeting model and the extent to which PCES improves decision quality given that profit is a more interpretable business metrics compared to statistical metrics such as AUC or accuracy.

2.3. Application

The existing literature in the field of IS explores various decision-making scenarios where AI plays a central role. First, it sheds light on mature IS applications that have long benefited from ML advancements but have more recently incorporated emerging trends in the field of XAI. Notable application domains include credit scoring [40], customer churn prediction [31], educational science [41], identifying factors influencing accident severity [42] and fraud detection [43]. For example, [31] introduces a novel ML algorithm, the spline-rule ensemble classifier with structured sparsity regularization, offering an effective balance between prediction accuracy and strong interpretability. [44] explores the use of probabilistic ML to understand motivational climate variables and their impact on student teaching evaluations, improving decision making through XAI.[42] introduces an explainable analytics framework combining descriptive, predictive, and prescriptive analytics to enhance the transparency and decision-making process to understand factors that contribute to the severity of accidents.

Conversely, we observe a growing interest in IS literature towards developing XAI solutions in domains that are not predominantly data-driven and ML-oriented. These newer applications encompass areas such as student attrition prediction [3] and human resource management [45]. For example, [3] visualize the course feedback given by the students using t-SNE. The authors find two segments for students that discuss various topics, including feedback related to the teacher, the course, the facilities, and student issues. In the context of a health crisis like a pandemic, the researchers do not just stop at developing ML models. They take it a step further by integrating it into a real-world setting. Using a blend of advanced analytics ([46] and Bayesian networks [47], they create a decision-making tool that healthcare professionals can actually use. [28] addresses the practical application of XAI by exploring the sharing economy within the emerging field of the Metaverse. The study moves beyond theoretical propositions by including a case study for empirical validation. While the case study is limited in scope, it provides a useful starting point for further research and offers some actionable insights for stakeholders in this specific area.

3. Overview of Contributions to this Special Issue

This section summarizes the papers of the special issue and gives insight into how they contribute to enhanced decision-making through XAI.

Online retail platforms face numerous challenges, such as cyber-attacks, data breaches, and device failures, that jeopardize their operational integrity and data security. [48] propose a robust resilience strategy that integrates explainable deep learning with a blockchain-based consensus protocol to address these threats. By leveraging this combined approach, the strategy facilitates rapid detection of incidents, provides clear explanations of detected vulnerabilities, and improves decision-making processes. To validate the strategy's effectiveness, the authors conducted an experimental study using NAB datasets on real-time online retail systems. The results highlight the strategy's ability to implement efficient and effective cyber resilience measures, thereby enhancing the decision-making capabilities necessary for maintaining business continuity and operational resilience.

[28] contextualize their study in the intersection of the sharing economy and the Metaverse. Transportation applications in the sharing economy reduce car ownership and single-vehicle use, enhancing regional environmental sustainability. The Metaverse integrates these applications with transportation networks, bolstering sustainability. This research supports decisionmakers by creating an explainable multi-criterion decision-making (MCDM) model to prioritize three Metaverse integration alternatives: safety measures, payment systems, and optimizing operations. The authors showcase a case study with twelve criteria in four aspects (economic, user, operational, and advancement) provides expert evaluations. Findings suggest optimizing operations through Metaverse integration is the best approach.

In this study, [49] introduce MERLIN, an innovative XAI method that offers contrastive explanations for two ML models, thereby presenting the idea of model-contrastive explanations. The researchers develop an encoding that enables MERLIN to handle both textual and tabular data, as well as a mix of continuous and discrete features. To demonstrate the efficacy of their method, they test MERLIN on a wide range of benchmark datasets. Additionally, the authors enhance the IS field by releasing MERLIN as a publicly accessible python-pip package.

[50] present a hybrid XAI framework to evaluate cyber-risks arising from correlated phishing attacks. Their framework is divided into several phases. The initial phase determines the probability of expert phishers within a community of attackers with varying levels of expertise. The subsequent phase assesses the likelihood of phishing attacks on a firm despite its investments in IT security and regulatory measures. The third phase differentiates phishing and legitimate URLs using various ML classifiers. Following this, it estimates the joint distribution of phishing attacks using an exponential-beta distribution and measures the expected loss with Archimedean Copula. Finally, the authors provide recommendations for firms by calculating optimal investments in cyber-insurance versus IT security.

[51] introduce an innovative and interpretable Graph Neural Network (GNN) model designed for assessing financial distress. This model incorporates financial indicators, network data, and event information. It not only identifies the propagation effects of events but also elucidates these effects, providing a clear analytical framework. Empirical tests confirm the model’s efficacy, and its interpretability offers significant insights for managerial decision-making.

The persistent pilot shortage in the USAF is being worsened by the rising demand from commercial airlines. To tackle this problem, [52] contribute to the IS literature by proposing a decision support system for selecting pilot candidates using advanced ML techniques. Considering the recent Responsible Artificial Intelligence Strategy released by the US Department of Defense, this study employs interpretable and explainable ML methods to develop traceable and fair models that can be governed responsibly and reliably. These models are utilized to predict candidates' average merit assignment selection system scores based on pre-selection and pre-training information.

The following paper addresses the lag in AI adoption among industry players by proposing XAI as a solution to improve decision-making processes. Using an experimental design, [27] provides empirical evidence on the impact of XAI on supply chain decision making. The study identifies a serial mediation path that involves transparency and agile decision making, showing that XAI enhances transparency and improves cyber resilience during supply chain cyberattacks. A post-hoc text analysis of X posts reveals a positive attitude towards XAI and highlights themes of transparency, explainability, and interpretability. In general, the findings underscore the significant role of XAI in improving decision support systems.

[53] introduce a comprehensive framework for expert-augmented supervised feature selection, covering pre-processing, in-processing, and post-processing aspects of XAI. In the pre-processing stage, the Probabilistic Solution Generator through Information Fusion (PSGIF) algorithm is introduced, using ensemble techniques to enhance a Genetic Algorithm's (GA) capabilities. The framework balances explainability and prediction accuracy with multiobjective optimization models, allowing experts to set acceptable sacrifice levels, thus improving explainability by focusing on relevant features. The in-processing stage incorporates expert opinions into feature selection as a multi-objective problem, optimizing GA parameters through Bayesian optimization. For post-processing, the Posterior Ensemble Algorithm (PEA) assesses the predictive power of features, comparing objective and subjective importance, and is tested on 16 datasets to validate the framework's effectiveness in both single and multi-objective settings.

In the e-commerce sector, [54] explores how textual product descriptions affect consumer choices and sales forecasting. The authors create an AI framework combining text mining, advanced neural networks, and regression analysis. By analyzing nearly 200,000 sales records, they demonstrate the use of SHAP to uncover the significance of specific product description phrases in enhancing sales forecasting and customer acquisition. This study presents a transparent and understandable model that uses textual data to improve e-commerce strategies.

Complying with regulatory frameworks such as GDPR becomes a crucial point when developing AI systems in high-stakes domains. Often, black-box models are not an option in these scenarios. [30] delves into this topic by introducing LitBoost, a tree-based boosting model that improves the interpretability of gradient boosted trees by modifying them to work similarly to GAMs, known for their transparency and simplicity. The paper evaluates LitBoost using a dataset from the Oslo housing market and a synthetic dataset designed to mimic real world heterogeneity. The results demonstrate LitBoost's ability to provide accurate and interpretable predictions, while enabling information sharing between different geographic groups during the training process.

[55] introduces a decision support system tailored for proactive management in railway control rooms. This system offers explainable predictions and recommendations on the automation usage in traffic control, where operators make critical decisions on signal operations. Emphasizing the need for explainability in automation, particularly in settings where split-second decisions are crucial, this decision support system employs SHAP values to ensure consistency in explainability across various ML methods. Implemented as a proof of concept in Infrabel's digital control rooms, the system's effectiveness and impact on operational management were validated through end-user feedback, focusing on the system's operational impact and the perceived value of incorporating agreement levels in explainability.

Shapley values are one of the main tools for model-agnostic post hoc explanations on tabular datasets. In [56], the authors propose a novel approach that combines two different explanations provided by Shapley values; one that explains the model itself and another that explains the model combined with feature dependencies in the data. This approach eliminates the need to choose between the two explanations and enhances the overall explanatory power of Shapley values, offering intuitive and superior insights on real-world datasets.

Outlier detection is a very challenging application for XAI techniques because it typically relies on unsupervised learning. In [57], the authors address the task of data categorization by building on formal concept analysis (FCA). They recognize that different sets of features used for classification can yield varying categorizations, influenced by subjective choices based on interests or preferences, referred to as interrogative agendas. They represent these agendas as sets of features and develop two algorithms: an unsupervised FCA-based algorithm for outlier detection using different agendas and a supervised metalearning algorithm to learn suitable (fuzzy) agendas with different feature weights. They combine these to create a supervised outlier detection algorithm, showing that their approach performs comparably to standard outlier detection algorithms on benchmark datasets and provides both local and global explanations of the results.

Another approach based on Shapley values was proposed in [58]. The authors focus on improving customer churn prediction by utilizing hybrid models that apply model segmentation and classification. The goal of these approaches is to capture non-linearities while ensuring interpretability. The authors introduce a novel model-agnostic approach based on SHAP to interpret these complex models. The proposed method is extensively benchmarked on 14 customer churn datasets for predictive performance, and the interpretability is demonstrated through a case study. The contributions include introducing new hybrid segmented models and proposing a modelagnostic tool for interpreting hybrid segmented models, relevant for decision makers in various predictive modeling tasks beyond customer retention management.

[59] introduce an innovative method utilizing Representational Similarity Analysis (RSA) to evaluate explanation similarity within XAI applications. The study demonstrates that RSA can quantitatively assess the stability of explanations across various ML models and preprocessing techniques. By mitigating explanation drift, RSA enhances the reliability and consistency of XAI systems, thereby supporting more robust decision-making frameworks. The findings indicate that RSA significantly improves the interpretability of ML models, leading to higher trust and better-informed decisions.

[60] propose an advanced predictive clustering model integrated with the Dirichlet Process Mixture Model (DPMM) to handle heterogeneous and longitudinal electronic health records (EHRs). By incorporating attention mechanisms, the model not only improves disease risk prediction but also offers multi-level explainable evidence. Evaluations on real-world datasets confirm the model's effectiveness in capturing detailed patient trajectories and providing actionable insights for healthcare practitioners. Performance evaluations reveal that the model achieves a higher accuracy and interpretability compared to traditional clustering methods.

[61] presents a methodology for generating counterfactual explanations that recommend business changes to improve customer repurchase behavior. By utilizing electronic word of mouth (eWOM) data, the method predicts customer revisit intentions and provides actionable insights for different

customer segments. The empirical evaluation demonstrates the method's superior performance compared to existing approaches, making it a valuable tool for enhancing business decision-making.

These papers collectively showcase innovative methodologies that significantly enhance decision-making through the application of XAI, addressing current challenges and proposing advanced solutions to foster future advancements in the field.

4. Final Notes & Future Research Agenda

This special issue shows that XAI is an important stream of research to further enhance decision making along the data, method and application side. Furthermore, various future research directions emerge that highlight several promising research areas. For example, within the area of e-commerce, future studies could investigate whether the relatively low impact of product descriptions on sales, compared to historical sales data and prices, is due to the intrinsic nature of the data or to the loss of information during text processing [54]. Especially with recent advances in natural language processing, this line of research can be of great interest to a wide IS readership. In the context of human-machine interaction, future research could study the addition of detailed operator information to the model, which is expected to greatly impact the prediction accuracy for future human operator behavior [55]. In the area of methodological innovations, more research could explore the integration of Structured Additive Regression (STAR) models with the idea of training local models to improve predictive accuracy while maintaining interpretability [30]. Furthermore, one of the main challenges in the current XAI landscape is to improve the interpretation of multimodal models, which involve integrating multiple types of data (e.g., text, images, audio, and structured data). By combining diverse data sources, multimodal XAI can provide richer and more comprehensive explanations of the prediction of the model. In this special issue, several methodological advances have been presented mainly for tabular data (see, e.g., [57, 56, 58]), textual data [49], and network data [51], suggesting that explainability is a key element for unimodal decision making. However, there is a pressing need for multimodal methods that can derive interpretable insights from multimodal tasks [62]. For example, [57, 58] show active research on XAI methods based on Shapley values, especially for business decision support systems. These methods can be further extended and combined with other XAI strategies for deep learning to provide integrated solutions for multimodal systems. Building on the work of [59], [60], and [61] future studies should explore the application of RSA to broader AI ecosystems, particularly focusing on the long-term impacts of explanation stability on user trust and system adoption. In addition, future research should investigate the scalability of advanced predictive clustering models and their application to diverse datasets. Enhancing these models with additional layers of explainability, such as incorporating domain-specific knowledge, could further improve their interpretability and utility. In addition, refining counterfactual explanation methods to handle more complex, multi-dimensional datasets and improve their applicability across various business domains. Future work could also focus on developing more personalized counterfactual explanations to meet specific user needs, enhancing decision making in personalized marketing strategies [61].

5. Acknowledgments

The guest editorial team thanks Prof. Victoria Yoon and Prof. Andrew Chen for their trust in this project and their leadership in guiding this special issue. Furthermore, we thank the authors of the papers collected in this special issue for their professionalism in incorporating the Reviewers' and Editors' comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] K. Topuz, F. D. Zengul, A. Dag, A. Almehmi, M. B. Yildirim, Predicting graft survival among kidney transplant recipients: A Bayesian decision support model, Decision Support Systems 106 (2018) 97–109. doi:10.1016/j.dss.2017.12.004.

[2] M. Kraus, S. Feuerriegel, A. Oztekin, Deep learning in business analytics and operations research: Models, applications and managerial implications, European Journal of Operational Research 281 (3) (2020) 628–641. arXiv:1806.10897, doi:10.1016/j.ejor.2019.09.018.

[3] M. Phan, A. De Caigny, K. Coussement, A decision support framework to incorporate textual data for early student dropout prediction in higher education, Decision Support Systems 168 (January) (2023) 113940. doi:10.1016/j.dss.2023.113940.

[4] E. Eryarsoy, K. Topuz, C. Demiroglu, Disentangling human trafficking types and the identification of pathways to forced labor and sex: an explainable analytics approach, Annals of Operations Research (2023) 1–35.

[5] F. Vandervorst, W. Verbeke, T. Verdonck, Data misrepresentation detection for insurance underwriting fraud prevention, Decision Support Systems 159 (2022) 113798. doi:https://doi.org/10.1016/j.dss.2022.113798.

[6] R. Fu, Y. Huang, P. V. Singh, Crowds, Lending, Machine, and Bias, Information Systems Research 32 (1) (2021) 72–92. doi:10.1287/isre.2020.0990.

[7] A. Fu¨gener, J. Grahl, A. Gupta, W. Ketter, Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation, Information Systems Research (2022). doi:10.1287/isre.2021.1079.

[8] S. Fan, N. Ilk, A. Kumar, R. Xu, J. L. Zhao, Blockchain as a trust machine: From disillusionment to enlightenment in the era of generative AI, Decision Support Systems 182 (2024) 114251. doi:https://doi.org/10.1016/j.dss.2024.114251.

[9] K. Idbenjra, K. Coussement, A. De Caigny, Investigating the beneficial impact of segmentation-based modelling for credit scoring, Decision Support Systems 179 (2024) 114170. doi:https://doi.org/10.1016/j.dss.2024.114170.

[10] S. Debaere, F. Devriendt, J. Brunneder, W. Verbeke, T. De Ruyck, K. Coussement, Reducing inferior member community participation using uplift modeling: Evidence from a field experiment, Decision Support Systems 123 (August) (2019). doi:10.1016/j.dss.2019.113077.

[11] A. De Caigny, K. Coussement, K. W. De Bock, Leveraging fine-grained transaction data for customer life event predictions, Decision Support Systems 130 (May 2019) (2020) 113232. doi:10.1016/j.dss.2019.113232.

[12] K. Coussement, S. Lessmann, G. Verstraeten, A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry, Decision Support Systems 95 (2017) 27–36. doi:10.1016/j.dss.2016.11.007.

[13] Y. Jung, Y. Suh, Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews, Decision Support Systems 123 (2019) 113074. doi:https://doi.org/10.1016/j.dss.2019.113074.

[14] V.-H. Trieu, H. Q. Vu, M. Indulska, G. Li, A computer vision-based concept model to recommend domestic overseas-like travel experiences: A design science study, Decision Support Systems 181 (2024) 114149. doi:https://doi.org/10.1016/j.dss.2023.114149.

[15] B. Cankaya, K. Topuz, D. Delen, A. Glassman, Evidence-based managerial decision-making with machine learning: The case of bayesian inference in aviation incidents, Omega (2023) 102906.

[16] H. Liu, Y. Ye, H. Y. Lee, High-Dimensional Learning Under Approximate Sparsity with Applications to Nonsmooth Estimation and Regularized Neural Networks, Operations Research (2022). doi:10.1287/opre.2021.2217.

[17] P. Borchert, K. Coussement, A. De Caigny, J. De Weerdt, Extending business failure prediction models with textual website content using

deep learning, European Journal of Operational Research 306 (1) (2022) 348–357. doi:10.1016/j.ejor.2022.06.060.

[18] D. Shin, The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI, International Journal of Human Computer Studies 146 (2021) 102551. doi:10.1016/j.ijhcs.2020.102551.

[19] A. Rai, Explainable AI: from black box to glass box, Journal of the Academy of Marketing Science 48 (1) (2020) 137–141. doi:10.1007/s11747-019-00710-5.

[20] B. Kim, J. Park, J. Suh, Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information, Decision Support Systems 134 (July 2019) (2020) 113302. doi:10.1016/j.dss.2020.113302.

[21] K. Topuz, T. L. Urban, M. B. Yildirim, A markovian score model for evaluating provider performance for continuity of care—an explainable analytics approach, European Journal of Operational Research (2023).

[22] D. Zhdanov, S. Bhattacharjee, M. A. Bragin, Incorporating FAT and privacy aware AI modeling approaches into business decision making frameworks, Decision Support Systems 155 (2022) 113715. doi:https://doi.org/10.1016/j.dss.2021.113715.

[23] G. Phillips-Wren, M. Daly, F. Burstein, Reconciling business intelligence, analytics and decision support systems: More data, deeper insight, Decision Support Systems 146 (March) (2021) 113560. doi:10.1016/j.dss.2021.113560.

[24] L. Wang, R. Gopal, R. Shankar, J. Pancras, Forecasting venue popularity on location-based services using interpretable machine learning, Production and Operations Management (2022). doi:https://doi.org/10.1111/poms.13727.

[25] P. Harrigan, T. M. Daly, K. Coussement, J. J. A. J. Lee, G. G. N. Soutar, U. Evers, Identifying influencers on social media, International Journal of Information Management 56 (September 2019) (2021) 102246. doi:10.1016/j.ijinfomgt.2020.102246.

[26] A. De Caigny, K. Coussement, W. Verbeke, K. Idbenjra, M. Phan, Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach, Industrial Marketing Management 99 (September) (2021) 28–39. doi:10.1016/j.indmarman.2021.10.001.

[27] K. Sadeghi R., D. Ojha, P. Kaur, R. V. Mahto, A. Dhir, Explainable artificial intelligence and agile decision-making in supply chain cyber resilience, Decision Support Systems 180 (2024) 114194. doi:https://doi.org/10.1016/j.dss.2024.114194.

[28] D. Pamucar, M. Deveci, I. Gokasar, D. Delen, M. K¨oppen, W. Pedrycz, Evaluation of metaverse integration alternatives of sharing economy in transportation using fuzzy Schweizer-Sklar based ordinal priority approach, Decision Support Systems 171 (2023) 113944. doi:https://doi.org/10.1016/j.dss.2023.113944.

[29] M. Kraus, D. Tschernutter, S. Weinzierl, P. Zschech, Interpretable generalized additive neural networks, European Journal of Operational Research (2023).

[30] A. Hjort, I. Scheel, D. E. Sommervoll, J. Pensar, Locally interpretable tree boosting: An application to house price prediction, Decision Support Systems 178 (2024) 114106. doi:https://doi.org/10.1016/j.dss.2023.114106.

[31] K. W. De Bock, A. De Caigny, Spline-rule ensemble classifiers with structured sparsity regularization for interpretable customer churn modeling, Decision Support Systems 150 (2021) 113523, interpretable Data Science For Decision Making. doi:https://doi.org/10.1016/j.dss.2021.113523.

[32] Z. Yi, Z. Liang, T. Xie, F. Li, Financial risk prediction in supply chain finance based on buyer transaction behavior, Decision Support Systems 170 (2023) 113964. doi:https://doi.org/10.1016/j.dss.2023.113964.

[33] Y. Huang, S. Meng, Automobile insurance classification ratemaking based on telematics driving data, Decision Support Systems 127 (2019) 113156.

[34] S. Lundberg, S.-I. Lee (Eds.), Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, 2017.

[35] A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin, Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation, J. Comput. Graph. Stat. 24 (1) (2015) 44–65.

[36] M. T. Ribeiro, S. Singh, C. Guestrin (Eds.), Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2016.

[37] B. Davazdahemami, H. M. Zolbanin, D. Delen, An explanatory machine learning framework for studying pandemics: The case of covid-19 emergency department readmissions, Decision Support Systems 161 (2022) 113730, data Analytics and Decision-Making Systems: Implications of the Global Outbreaks. doi:https://doi.org/10.1016/j.dss.2022.113730.

[38] E. Stripling, S. vanden Broucke, K. Antonio, B. Baesens, M. Snoeck, Profit maximizing logistic model for customer churn prediction using genetic algorithms, Swarm and Evolutionary Computation 40 (September) (2018) 116–130. doi:10.1016/j.swevo.2017.10.010.

[39] S. Lessmann, J. Haupt, K. Coussement, K. W. De Bock, Targeting customers for profit: An ensemble learning framework to support marketing decision-making, Information Sciences 557 (2021) 286–301.

[40] Z. Wang, J. Xiao, L. Wang, J. Yao, A novel federated learning approach with knowledge transfer for credit scoring, Decision Support Systems (2023) 114084doi:https://doi.org/10.1016/j.dss.2023.114084.

[41] K. Topuz, B. D. Jones, S. Sahbaz, M. Moqbel, Methodology to combine theoretical knowledge with a data-driven probabilistic graphical model, Journal of Business Analytics 4 (2) (2021) 125−139.

[42] A. Ahmed, K. Topuz, M. Moqbel, I. Abdulrashid, What makes accidents severe! explainable analytics framework with parameter optimization, European Journal of Operational Research (2023).

[43] R. Van Belle, B. Baesens, J. De Weerdt, Catchm: A novel network-based credit card fraud detection method using node representation learning, Decision Support Systems 164 (2023) 113866. doi:https://doi.org/10.1016/j.dss.2022.113866.

[44] B. D. Jones, K. Topuz, S. Sahbaz, Predicting undergraduate student evaluations of teaching using probabilistic machine learning: The importance of motivational climate, Studies in Educational Evaluation 81 (2024) 101353. doi:10.1016/j.stueduc.2024.101353.

[45] N. Lawrance, G. Petrides, M.-A. Guerry, Predicting employee absenteeism for cost effective interventions, Decision Support Systems 147 (2021) 113539. doi:https://doi.org/10.1016/j.dss.2021.113539.

[46] K. Topuz, D. Delen, A probabilistic bayesian inference model to investigate injury severity in automobile crashes, Decision Support Systems 150 (2021) 113557.

[47] K. Topuz, B. Davazdahemami, D. Delen, A bayesian belief networkbased analytics methodology for early-stage risk detection of novel diseases, Annals of Operations Research (2023) 1–25.

[48] K. Zkik, A. Belhadi, S. Kamble, M. Venkatesh, O. Mustapha, A. Sebbar, Cyber resilience framework for online retail using explainable deep learning approaches and blockchainbased consensus protocol, Decision Support Systems (2024)

114253doi:https://doi.org/10.1016/j.dss.2024.114253.

[49] L. Malandri, F. Mercorio, M. Mezzanzanica, A. Seveso, Model-contrastive explanations through symbolic reasoning, Decision Support Systems 176 (2024) 114040. doi:https://doi.org/10.1016/j.dss.2023.114040.

[50] B. Biswas, A. Mukhopadhyay, A. Kumar, D. Delen, A hybrid framework using explainable AI (XAI) in cyber-risk management for defence and recovery against phishing attacks, Decision Support Systems 177 (2024) 114102. doi:https://doi.org/10.1016/j.dss.2023.114102.

[51] J. Wang, C. Jiang, L. Zhou, Z. Wang, Assessing financial distress of SMEs through event propagation: An adaptive interpretable graph contrastive learning model, Decision Support Systems 180 (2024) 114195. doi:https://doi.org/10.1016/j.dss.2024.114195.

[52] D. Wasilefsky, W. N. Caballero, C. Johnstone, N. Gaw, P. R. Jenkins, Responsible machine learning for United States Air Force pilot candidate selection, Decision Support Systems 180 (2024) 114198. doi:https://doi.org/10.1016/j.dss.2024.114198.

[53] M. Rabiee, M. Mirhashemi, M. S. Pangburn, S. Piri, D. Delen, Towards explainable artificial intelligence through expert-augmented supervised feature selection, Decision Support Systems 181 (2024) 114214. doi:https://doi.org/10.1016/j.dss.2024.114214.

[54] S. Chen, S. Ke, S. Han, S. Gupta, U. Sivarajah, Which product description phrases affect sales forecasting? An explainable AI framework by integrating WaveNet neural network models with multiple regression, Decision Support Systems 176 (2024) 114065. doi:https://doi.org/10.1016/j.dss.2023.114065.

[55] L. Sobrie, M. Verschelde, Real-time decision support for human–machine interaction in digital railway control rooms, Decision Support Systems 181 (2024) 114216. doi:https://doi.org/10.1016/j.dss.2024.114216.

[56] J. Michiels, J. Suykens, M. De Vos, Explaining the model and feature dependencies by decomposition of the Shapley value, Decision Support Systems 182 (2024) 114234. doi:https://doi.org/10.1016/j.dss.2024.114234.

[57] M. Boersma, K. Manoorkar, A. Palmigiano, M. Panettiere, A. Tzimoulis, N. Wijnberg, Outlier detection using flexible categorization and interrogative agendas, Decision Support Systems 180 (2024) 114196. doi:https://doi.org/10.1016/j.dss.2024.114196.

[58] A. De Caigny, K. W. De Bock, S. Verboven, Hybrid black-box classification for customer churn prediction with segmented interpretability analysis, Decision Support Systems 181 (2024) 114217. doi:https://doi.org/10.1016/j.dss.2024.114217.

[59] F. Gwinner, C. Tomitza, A. Winkelmann, Comparing expert systems and their explainability through similarity, Decision Support Systems (2024) 114248doi:https://doi.org/10.1016/j.dss.2024.114248.

[60] A. Artelt, A. Gregoriades, Supporting organizational decisions on How to improve customer repurchase using multi-instance counterfactual explanations, Decision Support Systems (2024) 114249doi:https://doi.org/10.1016/j.dss.2024.114249.

[61] S. Niu, Q. Yin, J. Ma, Y. Song, Y. Xu, L. Bai, W. Pan, X. Yang, Enhancing healthcare decision support through explainable AI models for risk prediction, Decision Support Systems 181 (2024) 114228. doi:https://doi.org/10.1016/j.dss.2024.114228.

[62] N. Rodis, C. Sardianos, G. T. Papadopoulos, P. Radoglou-Grammatikis, P. Sarigiannidis, I. Varlamis, Multimodal explainable artificial intelligence: A comprehensive review of methodological advances and future research directions, arXiv preprint arXiv:2306.05731 (2023).