International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

Ethical Credit Risk Assessment: Leveraging Machine Learning and Ethical Knowledge Acquisition for Responsible Lending

Aryan Guleria[1], Dr.Davinder Goriwal[2]
[1]*Department of Mathematics, Chandigarh University,Gharuan-Mohali, Punjab,India*
[2]*Department of Mathematics, Chandigarh University,Gharuan-Mohali, Punjab,India*

Abstract
Responsible lending decisions involve numerous considerations, and recent research has focused on applying machine learning techniques to credit risk assessment. This advancement holds the potential to promote lending diversity without compromising credit quality, utilizing historical lending data and outcomes. However, highly accurate machine learning models often lack transparency, posing a challenge for human domain experts seeking to ensure fairness and ethics. Consequently, jurisdictions worldwide are introducing regulations mandating explainability for automated decisions. Traditionally, lending decisions relied on human expertise, allowing for assessable and ethical judgment. In this study, we propose leveraging human expertise to address the limitations caused by inadequate data. By combining machine learning on a small training set with the correction of errors through Ripple-Down Rules, we aim to enhance lending decision-making. Through experimentation with benchmark data, we discovered that the resulting combined model not only achieved performance equivalent to models trained on larger datasets but also benefited from the incorporation of human expert rules. This approach has broader applicability, extending beyond lending decisions to domains where limited data quantity or quality hinders responsible decision-making. By embracing explainable machine learning techniques and integrating human expertise, we can not only improve the accuracy and appropriateness of lending decisions but also foster responsible decision-making across various domains. This research contributes to the growing body of knowledge on responsible credit risk assessment, addressing the challenges associated with limited data availability.
Keywords: Explainable AI, Machine learning, Credit risk assessment, Fairness and ethics

1. INTRODUCTION

As banks and lenders strive to automate financial decision-making processes, regulatory authorities are increasingly emphasizing responsibility and accountability in these decisions. One area of concern for banks is credit risk management, where machine learning (ML) holds the potential to enhance existing techniques, particularly in credit scoring, by leveraging the growing availability of data and computing power [1]. Credit scoring traditionally involves models designed to classify credit applicants as either "good" or "bad" risks [2]. While machine learning applications have demonstrated improved accuracy in credit scoring, the most effective algorithms, such as ensemble classifiers and neural networks, often require large datasets that may be challenging to acquire. Additionally, these models lack interpretability, raising concerns regarding fairness, compliance, and risk management [3, 4]. Consequently, in practice, less accurate yet more interpretable models continue to be employed [5]. However, these interpretable models may also exhibit accuracy deficiencies that intersect with issues of fairness.

In this paper, we propose a novel approach to address the aforementioned challenges by combining machine learning (ML) classification models trained on limited data with a well-established "human-in-the-loop" knowledge acquisition technique called Ripple-Down Rules (RDR) [6]. Our framework, referred to as ML+RDR, aims to construct fair, compliant rules while improving overall performance.

DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

The ML+RDR framework empowers a human domain expert to iteratively refine and enhance a machine learning model by incorporating rules based on domain knowledge. These rules are specifically applied to cases where the machine learning model fails to predict outcomes accurately and fairly. By monitoring and utilizing domain knowledge for all processed cases, it becomes possible to correct any errors made by the machine learning model. Additionally, during the development of a machine learning model, there are typically training cases where the model fails to assign the correct label provided in the training data. These misclassified cases serve as a pre-selected set for the domain expert to add rules..

In this study, we focus on applying RDR to rectify these "data errors" without specifically considering fairness issues. Our objective is to evaluate the efficacy of addressing classification errors made by the ML model on the training data. However, we discuss the potential for extending our approach to rectify "fairness errors" as we delve deeper into the error-correcting process of RDR later in the paper. The implications of our approach extend beyond simple error correction, encompassing the judgment of the human domain expert, which can be based on various considerations of what constitutes responsible decision-making by a financial institution in a specific case. These considerations may include regulatory requirements for credit rating criteria, such as those outlined by the Australian Prudential Regulation Authority (APRA), which emphasize fairness, plausibility, intuitiveness, and meaningful differentiation of risk [7]. The rules created by a domain expert are expected to adhere to such criteria, ensuring fairness and interpretability while being easily verifiable against these requirements.

With these objectives in mind, our approach is designed to align with common requirements from lender institutions or banks. Firstly, we maintain the flexibility to customize the conversion from raw scores to calibrated probabilities and ultimately to binary outcomes, aligning with the specific business use case's cost/reward structure. Secondly, we ensure the stability of the decision strategy as the raw scores are updated after the ML model improves through rule acquisition, taking into account the calibration of scores. Finally, we provide transparency in the process of converting raw scores to calibrated scores and the decision strategy through a declarative specification. The subsequent sections of the paper are organized as follows: Section 2 presents related work and establishes the context for the framework proposed in this paper. Section 3 outlines the experimental setup and introduces the framework for domain knowledge-driven credit risk assessment. Section 4 presents the experimental results, while Section 5 demonstrates how domain knowledge can enhance model performance in credit risk assessment using the knowledge acquisition framework (RDR) that enables a domain expert to construct new features and rules as needed. Finally, Section 6 concludes the paper, providing closing remarks and discussing potential future directions..

1.1. RELATED WORK

This paper tackles the challenge of developing responsible AI systems, with a specific focus on the field of financial technology (fintech). The concept of responsibility, as defined in dictionaries, refers to the obligation to adhere to rules and laws, as well as the ability to control or influence outcomes and actions. In essence, being responsible entails being accountable and capable of providing a comprehensive account or explanation for events, which aligns with the characteristic of human expertise [8].

In the context of AI, the term "responsible" primarily pertains to algorithmic fairness and the avoidance of prohibited or undesirable biases in AI-driven decision-making [9]. To ensure fairness, AI systems need to be accountable according to legal and ethical requirements, which, in turn, necessitates their ability to explain their decisions to domain experts [10]. While responsibility is often viewed as the intersection of fairness and explainability, both concepts pose significant challenges. Evaluating fairness and defining it precisely remain complex tasks, with multiple definitions of fairness emerging in the literature [12].

In the pursuit of fairness through explainability, there has been a notable increase in research over the past decade [13] aiming to enhance the intelligibility of AI systems, the data they are trained on, and

the outputs they produce [14, 15]. This research encompasses techniques commonly referred to as explainable AI (XAI) [16] or interpretable ML (IML) [17]. Various terms, such as transparent or intelligible, are often used interchangeably with these concepts. Notably, it is commonly assumed that there is a trade-off between the predictive accuracy and interpretability of ML models. However, by applying additional explainability methods to high-performing black box models, it becomes possible to analyze and potentially correct their errors [10].

Currently, commercially available or open-source software tools are employed for explaining ML models. Some commonly used techniques can be summarized as follows:

1. Variable or feature importance: Techniques like Shapley values [18] can assess whether the output predicted by a black box model aligns with human understanding in a particular domain.
2. Surrogate models: Inherently interpretable models such as linear regression or decision trees can be used as surrogate models (e.g., in LIME [19]) to explain the predictions of black box models. These surrogate models are trained on synthetic data that closely resembles the instance being explained.
3. Quantitative and qualitative standardized documentation: "Model cards" [20] provide structured documentation that evaluates ML models and highlights potential impacts on different groups based on characteristics of the data used in black box models.
4. Fairness metrics: ML models can be trained under the constraint that classification performance should be similar for individuals, regardless of whether they belong to sensitive groups. This approach involves indicating membership in known protected or sensitive groups through a variable in the data [21].
5. Visualizations: Various visualization techniques, such as plots or dashboards, can be employed to present the results generated by the aforementioned methods [22].

In our previous work [23, 24], we have acknowledged the usefulness of feature importance measures in evaluating the performance of credit risk assessment models. However, we have also recognized that relying solely on these measures is insufficient to address prediction errors. Surrogate models, although helpful for explanation purposes, present challenges concerning stability (where different runs of an explainability method can yield varying explanations for the same instance [25]) and faithfulness (where the relationship between the explanation and the model's prediction is unclear [26]).
Model cards are expected to play an increasingly important role in the future, possibly as part of upcoming legislation regarding AI systems. While fairness metrics, particularly those implementable in ML models [12], can be explored in our future research, our current experiments in this paper lack data containing information on protected attributes. We have applied a new implementation of Rule-Based Decision (RDR) with a GUI that allows human experts to inspect data, identify ML errors, and add or update rules (Section 3.2.2). Although the addition of visualizations to assist human experts in the RDR process could enhance the GUI, this aspect is left for future work.

In the finance industry, adhering to regulatory and ethical standards is of utmost importance when applying AI systems, particularly in credit decision-making [27]. A recent survey [3] reviewed 136 papers on ML applications for credit risk evaluation, focusing on credit scoring, prediction of non-performing assets, and fraud detection. Support vector machines, neural networks, and ensemble methods were the most commonly employed ML techniques. However, only one early work on explainability [28] was cited. Another study examined 76 research papers on statistical and ML approaches to credit scoring, revealing that ensembles and neural networks outperformed traditional logistic regression. Nevertheless, a significant limitation was
the inability of certain ML models to explain their predictions [29].

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

Current research on credit decisioning tasks involving explainable ML primarily revolves around assessing variable importance (e.g., [30–32]). Instead of employing complex ML algorithms to enhance accuracy in credit risk prediction and then applying explainability techniques, a contrasting approach was investigated in [33]. This approach utilized a linear method like logistic regression, which allows for assessing the impact of each variable on the output within current regulatory requirements. However, it required manual engineering of complex features to achieve accuracies comparable to more sophisticated algorithms. Nonetheless, this reliance on human feature engineering may pose scalability challenges, particularly when utilizing alternative data sources [34]. The approach we propose in this paper differs from previous work on explainable AI. Prior work in XAI primarily focuses on providing explanations for an existing machine learning model. In contrast, our proposal involves integrating additional knowledge in the form of rules to rectify errors in an existing model. These data-driven rules are inherently understandable to domain experts. However, it should be noted that only the newly added knowledge is intrinsically explainable, and additional explainability techniques would be required for the machine learning component of the model that performs well on the training data. In our research, we have not needed to employ such techniques as our main focus has been to demonstrate that human knowledge can be easily incorporated into a machine learning model. This additional human knowledge not only improves the overall system's performance but also explicitly This additional human knowledge not only improves the overall system's performance but also explicitly addresses issues of fairness and other concerns that a human user may identify in the output of an ML-based system.

1.2. Methodology

In credit risk assessment, binary prediction alone is insufficient because what is required is a model score that represents the Probability of Default (PD). To leverage machine learning, it is necessary to train a model with a numeric output, specifically a regression-type model. In this paper, we utilized XGBoost [35], a popular algorithm that trains an ensemble of regression trees using gradient boosting [36]. Although XGBoost is extensively used, this paper does not provide any argument about its applicability to credit assessment. Instead, it serves as an illustration of how a machine learning method can be enhanced by incorporating domain knowledge from human experts.

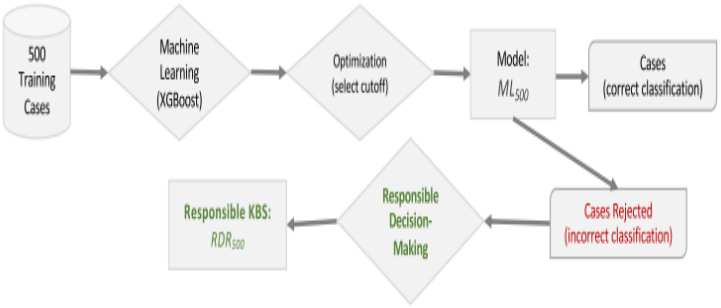

Figure1: Framework to build ML_{500} and responsible knowledgebase RDR_{500}

We have developed a novel implementation of Rule-Based Decision (RDR) based on the structure of the code released in [6], which is freely available. Our implementation includes a new graphical user interface (GUI) specifically designed for the credit risk assessment domain. This GUI allows domain experts to efficiently view data, add rules, and evaluate their impact. To construct a rule that covers a misclassified ML training case, we incrementally add conditions until the evolving rule accurately classifies all cases that the ML knowledge base misclassified. The selection of features to be added as rule conditions is determined by the domain expert's expertise. Our RDR implementation is designed to seamlessly integrate with other components, resembling a data and decisioning pipeline commonly used in the finance industry. This implementation was utilized in all the experiments presented in this paper.

To determine the cutoff point for binary prediction, the problem should be approached as a binary

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

prediction task. This means that there are four possible outcomes based on the labeled data. To ensure consistency with real-world applications, a realistic cost structure is adopted, although other cost structures are also possible. In this case, the following cost assignments were used:

- A prediction of default results in no loan being granted, resulting in zero costs or profits.
- Predictions of loans being repaid have non-zero values assigned. A "good" loan (repaid) is assigned a profit value of $300, while a "bad" loan (default) is assigned a cost of $1,000.

With these cost assignments, the cutoff point is determined as follows using the training data consisting of n cases (n = 500):

1. The predicted scores and actual binary outcomes are extracted into a table with two columns and n rows.
2. This table is sorted in descending order based on the predicted scores.
3. n-1 cutoff points are calculated, where the ith cutoff is the mean of the ith and (i + 1)th scores.
4. The cutoff points are used to convert the scores into binary predictions. If a score is greater than the cutoff, the prediction is 1; otherwise, it is 0.
5. Predictions are made for each of the n cases using each cutoff point and compared to the actual outcomes.
6. For each cutoff point, the total monetary value is calculated as the sum of (TN × 300) - (FN × 1000), where TN represents the number of true negatives (cases predicted to be repaid and actually repaid) and FN represents the number of false negatives (cases predicted to be repaid but actually defaulted).
7. The selected cutoff point is the one that maximizes the monetary value.

The optimal cutoff point determined by this method is used to convert the scores obtained from the machine learning model into the binary outcomes used in the decision-making strategy for all experiments.

Figure 2: Testing Framework

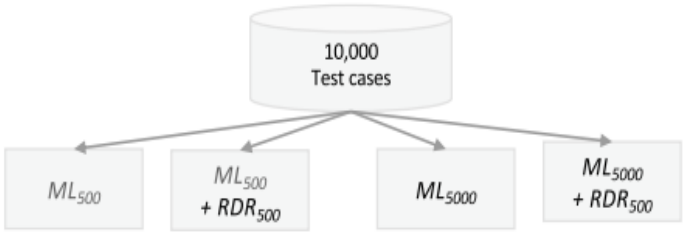

Figure 3: Performance Evaluation Framework

2. RESULTS AND DISCUSSION

To evaluate the performance of ML500 and ML5000, we selected 10,000 stratified random samples from credit card and debt consolidation data specifically from the year 2018. These samples were used to construct a test dataset. In addition, we augmented each machine learning model with RDR500 to assess the effectiveness of RDR500, which is a knowledge base acquired using only the error cases from the machine learning process with 500 training cases.

The testing framework for ML500 + RDR500 is illustrated in Figure 2. In these experiments, the

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

conclusion reached by the machine learning model for a particular case was accepted, unless an RDR rule fired and overrode the machine learning conclusion. In such cases, the conclusion provided by RDR was accepted. The same comparison and evaluation were conducted for ML5000 and ML5000 enhanced with RDR500. The four experiments conducted are depicted in Figure 3.

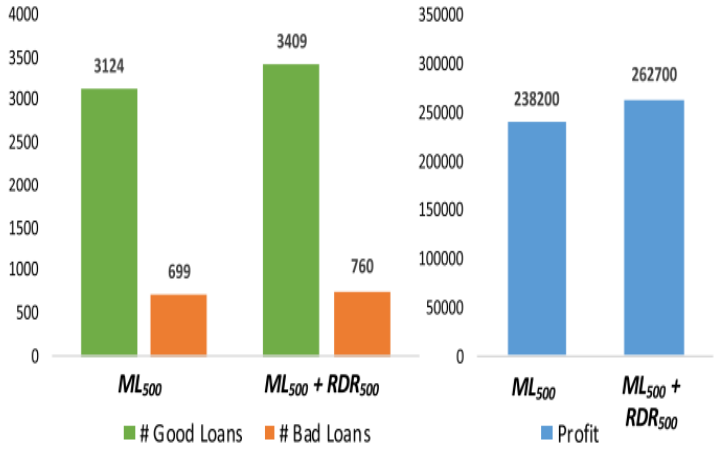

Figure 4: Output

The results indicate that the RDR rules acquired based on past experience with risk assessment relax the approval requirements in a way that ultimately increases overall profit. When developing the ML model, the cutoff between good and bad loans was determined based on a loss of -$1000 for a bad loan and a profit of $300 from a good loan. We also experimented with other values that altered the cutoff threshold, but the overall outcome remained consistent: augmenting ML500 with RDR500 increased profit. It is important to note that adjusting the unit benefit for good loans or the unit cost for bad loans will impact overall profit. However, this impact will affect both ML500 and ML500 + RDR500. Our results clearly demonstrate an improvement in the number of good loans versus bad loans for ML500 + RDR500 compared to ML500, supporting our claim that the technique is valuable. The relative cost of good loans versus bad loans is a business decision that determines the usefulness of predictive models in practical applications. For instance, increasing the profit to $400 while maintaining a cost of -$1000 leads to an overall profit improvement of 9.63%. Similarly, keeping the profit at $300 and increasing the cost to -$1100 results in a profit improvement of 10.93%. In this paper, we have presented indicative costs based on typical real-world experience, and slight variations that are reasonable in business applications have also demonstrated profit improvement, showcasing the robustness of the approach. It is worth noting that larger variations in relative costs will either increase or decrease the overall profit, and this aspect could be further explored in future work.
results rather than a repetition of the Results. The Discussion should be an interpretation of the results rather than a repetition of the Results. The Discussion should be an interpretation of the results rather than a repetition of the Results.

CONCLUSION

One of the most significant challenges in today's financial technology (fintech) landscape, with the increasing use of AI (ML), is to ensure that lending criteria are ethically sound and align with regulatory requirements. The experimental results presented in this paper demonstrate that it is possible to effectively incorporate credit risk assessment insights from human experts to augment machine learning models and enhance their performance, which can include rectifying unfair decisions. While this study does not directly address unfair decisions, it showcases the potential to capture and leverage human insights to prevent or correct unethical lending practices by machine learning models, thereby ensuring better adherence to guidelines.
In this study, ML models were constructed using XGBoost. In future research, our approach will be

evaluated with other machine learning methods and training data, particularly simulated training data. The aim of these further studies is not to prove that the approach will work with other methods or data, as it is inherently expected, but rather to explore how its performance relates to the size of the training dataset and the presence of anomalous patterns in the data, as opposed to merely rare patterns. By incorporating simulated data, we can introduce varying numbers of rare patterns into the training data. It is anticipated that including more rare patterns may potentially degrade the performance of the machine learning model, while simultaneously enhancing the overall performance improvement achieved by adding RDR rules. Moreover, we will evaluate different proportions of purely anomalous data versus data that represents rare, but genuine patterns. The machine learning model may attempt to learn from anomalous data if its methods for avoiding overfitting do not effectively work, whereas an expert should be able to recognize that the data instance does not make sense.

Additionally, we will explore the construction of models using RDR without relying on machine learning. Although this will require more effort from the domain expert, it may facilitate the development of more ethically sound models. This is because the model captures and expresses an expert's intuitions and knowledge, which are presumably more consistent with ethical lending practices and guidelines than the knowledge derived solely from the essential statistical analysis of data upon which machine learning relies.

REFERENCES

1. Leo M, Sharma S, Maddulety K. Machine learning in banking risk management: a literature review. Risks. 2019;7:29. https://doi.org/ 10.3390/risks7010029.
2. Hand DJ, Henley WE. Statistical classification methods in consumer credit scoring: a review. J R Stat Soc Ser A. 1997;160(3):523–41.
3. Bhatore S, Mohan L, Reddy YR. Machine learning techniques for credit risk evaluation: a systematic literature review. J Bank Financ Technol. 2020;4:111–38. https://doi.org/10.1007/ s42786-020-00020-3.
4. Kleinberg J, Mullainathan S, Raghavan M. Inherent trade-ofs in the fair determination of risk scores; 2016. arXiv preprint arXiv: 1609.05807v2
5. Dumitrescu E, Hué S, Hurlin C, Tokpavi S. Machine learning for credit scoring: improving logistic regression with non-linear decision-tree efects. Eur J Oper Res. 2022;297(3):1178–92.
6. Compton P, Kang BH. Ripple-down rules: the alternative to machine learning. CRC Press; 2021.
7. Australian Prudential Regulation Authority: Capital Adequacy: Internal ratings-based approach to credit risk. Prudential standard APS 113 ; 2020
8. Feltovich P, Prietula M, Ericsson K. Studies of expertise from psychological perspectives: historical foundations and recurrent themes. In: The cambridge handbook of expertise and expert performance. Cambridge: CUP; 2018. p. 59–83.
9. Dignum V. Responsible artifcial intelligence: how to develop and use AI in a responsible way. Springer; 2019. https://doi.org/10. 1007/978-3-030-30371-6.
10. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrer F. Explainable artifcial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inform Fusion. 2020;58:82–115.
11. Dai J, Upadhyay S, Aïvodji U, Bach S, Lakkaraju H. 2022 Fairness via Explanation Quality: Evaluating Disparities in the Quality of Post hoc Explanations. In: AIES ’22, AAAI/ACM Conference on Artifcial Intelligence, Ethics, and Society, pp. 203–214.
12. Toreini E, Aitken M, Coopamootoo K, Elliott K, Zelaya V, Missier P, Ng M, van Moorsel A. Technologies for trustworthy machine Human-Centric Intelligent Systems (2023) 3:232–243 243 1 3 learning: A survey in a socio-technical context; 2022. arXiv preprint arXiv:2007.08911
13. Graziani M, Dutkiewicz L, Calvaresi D, Pereira Amorim J, Yordanova K, Vered M, Nair R, Henriques Abreu P, Blanke T, Pulignano V, Prior J, Lauwaert L, Reijers W, Depeursinge A,

Andrearczyk V, Müller H. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Artif Intell Rev. 2022. https://doi.org/10. 1007/978-3-030-30371-6.
14. Gunning D, Aha D. DARPA's explainable artifcial intelligence program. AI Mag. 2019;40(2):44–58.
15. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv. 2019;51(5):1–42. https://doi.org/10.1145/ 3236009.
16. Ding W, Abdel-Basset M, Hawash H, Ali A. Explainability of artifcial intelligence methods, applications and challenges: a comprehensive survey. Inf Sci. 2022;615:238–92. https://doi.org/10. 1016/j.ins.2022.10.013.
17. Chen V, Li J, Kim JS, Plumb G, Talwalkar A. Interpretable machine learning: moving from mythos to diagnostics. Commun ACM. 2022;65(8):43–50. https://doi.org/10.1145/3546036.
18. Lundberg S, Lee SI. 2017 A unifed approach to interpreting model predictions. Adv Neural Inform Process Syst .
19. Ribeiro M, Singh S, Guestrin C. 2016 Why Should I Trust You?: Explaining the Predictions of Any Classifer. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining . p. 1135–1144.
20. Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji I, Gebru T. Model Cards for Model Reporting. In: FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019. p. 220–229.
21. Vogel R, Bellet A, Clémençon S. Learning fair scoring functions: bipartite ranking under ROC-based fairness constraints. Proceed Int Conf Artif Intell Stat (AISTATS). 2019;130:784.
22. Baniecki H, Kretowicz W, Piatyszek P, Wisniewski J, Dalex Biecek P. Responsible machine learning with interactive explainability and fairness in python. J Mach Learn Res. 2022;22:1–7.
23. Suryanto H, Guan C, Voumard A, Beydoun G.:Transfer learning in credit risk. In: ECML-PKDD 2019: Proceedings of European Conference on machine learning and knowledge discovery in databases, Part III; 2020. p. 483–498.
24. Suryanto H, Mahidadia A, Bain M, Guan C, Guan A. Credit risk modelling using transfer learning and domain adaptation. Front Artif Intell. 2022;70:1.
25. Alvarez-Melis D, Jaakkola T. On the robustness of interpretability methods; 2018. arXiv preprint arXiv:1806.08049.
26. Dasgupta S, Frost N, Moshkovitz M. Framework for Evaluating Faithfulness of Local Explanations. In: ICML 2022: Proceedings of the 39th International Conference on Machine Learning. PMLR 162; 2022.
27. Chen J. Fair lending needs explainable models for responsible recommendation; 2018. arXiv preprint arXiv:1809.04684.
28. Baesens B, Setiono R, Mues C, Vanthienen J. Using neural network rule extraction and decision tables for credit-risk evaluation. Manage Sci. 2003;49(3):312–29.
29. Dastile X, Celik T, Potsane M. Statistical and machine learning models in credit scoring: a systematic literature survey. Appl Soft Comput J. 2020;106263:1–21.
30. Bracke P, Datta A, Jung C, Sen S. Machine learning explainability in fnance: an application to default risk analysis. Bank of England Staf Working Paper no. 816 (2019)
31. Bussmann N, Giudici P, Marinelli D, Papenbrock J. Explainable machine learning in credit risk management. Comput Econ. 2021;57:203–16.
32. Jammalamadaka K, Itapu S. Responsible AI in automated credit scoring systems. AI Ethics. 2022. https://doi.org/10.1007/ s43681-022-00175-3.
33. Bücker M, Szepannek G, Gosiewska A, Biecek P. Transparency, auditability, and explainability of machine learning models in credit scoring. J Oper Res Soc. 2022;73(1):70–90.
34. Djeundje V, Crook J, Calabrese R, Hamid M. Enhancing credit scoring with alternative data.

DOI:10.46647/ijetms.2024.v08i03.063 ISSN: 2581-4621

Expert Syst Appl. 2021;163:113766
35. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785–794.
36. Friedman J. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29(5):1189–232.