Transparency and Efficiency in Credit Risk Assessment of Alternative Financing: A Green AI Approach

KATSIARYNA RADZKOVA, University of Twente, The Netherlands

Peer-to-Peer (P2P) lending platforms utilize Machine Learning (ML) to predict loan defaults effectively. Nevertheless, the implementation and training of these models require significant computational resources, which raises environmental issues. In addition to analyzing the application of ML models for default risk prediction, the research proposes ways to minimize the environmental impact by following Green AI principles. The goal of the study is to discover algorithms that maintain a balance between predicted accuracy and sustainability by examining four ML models: XGBoost, Random Forests, Decision Trees, and Logistic Regression. Additionally, the research explores strategies of applying Explainable AI techniques to increase transparency of the models. The results of the research demonstrate that although Random Forest model have a high default prediction accuracy, its shows large environmental impact in terms of CO2 emissions. However, despite lower performance, Logistic Regression offers a reasonable balance between accuracy and sustainability of the model.

Additional Key Words and Phrases: Peer-to-Peer Lending, Machine Learning, Default Prediction, Green AI, Model Transparency

1 INTRODUCTION

Over the past years, Peer-to-Peer (P2P) lending platforms transformed the financial industry by giving an alternative to traditional banking. By directly linking investors and borrowers via online platforms, P2P platforms facilitate and simplify the loan application process [22, 25, 26, 31]. Although P2P lending offers simple and rapid services, predicting default risk is a challenge. To avoid possible losses for investors and guarantee the credibility of the lending platform, accurate risk prediction is essential [1, 7, 22, 35]. The introduction of Machine Learning (ML) models has emerged as a strategy for predicting default risks in alternative finance methods. ML models have the ability to accurately analyze huge amounts of data, and more importantly, can identify non-linear relations from borrower data that are not feasible via traditional statistical methods [14, 16, 36]. However, complex models often reach outstanding performance by utilizing advanced computational technologies; therefore, producing significant carbon emissions as model training requires a high power consumption [33, 34].

1.1 Objectives and Research Questions

The primary objective of this research is to investigate potential applications of ML models to decrease the risks associated with P2P lending. Furthermore, it aims to discover methods to guarantee the transparency of ML models. In addition, the study seeks to define "green AI" and understand how to integrate its concepts into P2P lending platforms in order address environmental issues.

These goals could be achieved by answering the following research question: *"How can machine learning techniques be applied to accurately predict risk of default in P2P lending platforms, while ensuring that these models are both environmentally and ethically responsible in terms of processing time and algorithmic transparency?"* In order to support proposed research question, we suggest three sub-questions:

1. What are the environmental issues associated with implementing different ML models for default risk prediction on P2P lending platforms?
2. Which ML models or algorithms offer the best balance between accuracy in predicting loan defaults on P2P platforms and minimal sustainability impacts?
3. How can Explainable AI (XAI) techniques enhance the transparency of ML models in P2P lending?

1.2 Structure and Overview

The study is split into four sections that address the research question outlined in the previous section. First, an analysis of the literature will be performed to collect and analyze previous works on ML models for default loan predictions, as well as Green AI and model transparency. The study's methodology, which is discussed next, covers the study's design and the approaches taken to address the research question. The steps performed to get prepared for the realization of the experiment, which will further support the study, are described in the experimental setup section. Additionally, the results of the study and the research's conclusions are discussed in the corresponding sections.

2 LITERATURE REVIEW

2.1 Methodology

The systematic review will investigate how machine learning techniques could be applied for default predictions in P2P lending platforms, as well as provide insight into how to evaluate transparency and sustainability of these models. In order to establish kowledge about these concepts the University of Twente scientific database[1] would be used. The list of papers for this study would be based on the publications of Scopus[2], IEEE Xplore[3],and ACM Digital Library[4] digital libraries.

To select the initial set of papers the keywords "Peer-to-Peer Lending", "Default Risk", "Machine Learning", "Transparency", "Green AI", and their succeeding synonyms were used as query components. Due to the absence of research addressing the application of ML models for default risk assessment in P2P lending while evaluating the environmental sustainability or "greenness" of such AI models, it

[1] https://www.utwente.nl/en/service-portal/university-library
[2] https://www.scopus.com/home.uri
[3] https://ieeexplore.ieee.org/Xplore/home.jsp
[4] https://dl.acm.org/

TScIT 41, July 5, 2024, Enschede, The Netherlands

became essential to split the originally constructed query into three parts.

(1) **Potential Application of ML to address default risks:**
(("peer-to-peer lending" OR "p2p lending") AND ("machine learning" OR "ML" OR "deep learning") AND ("default risk" OR "credit risk") AND ("problem" OR "issue"))
(2) **Ensuring transparency of ML model that predicts default risks:**
((("machine learning" OR "ML" OR "artificial intelligence" OR "deep learning") AND ("default risk" OR "credit risk") AND ("peer-to-peer lending" OR "P2P lending")) AND (("explainable AI" OR "XAI") AND ("transparen*" OR "fairness" OR "bias*")))
(3) **Assessing sustainability of ML model following the principles of Green AI:**
(("green AI" OR "sustainable AI") AND ("machine learning" OR "ML" OR "deep learning") AND ("processing time" OR "energy consumption") AND ("carbon emission*" OR "carbon footprint"))

The queries were applied to the selected set of digital libraries in order to define relevant articles. The search resulted in 473 articles through the first phase of the search.Further, the resulting set of articles was filtered based on inclusion and exclusion criteria discussed in the Table 1.

As a result of article selection based on described steps, we obtained total of 66 publications, as specified in Appendix A. These articles were used to understand main concepts related to the topic and support the experiment conducted in this study.

Table 1. Inclusion and Exclusion criteria

Criteria	Decision
Keywords are present in title, abstract, or keyword list	Inclusion
Publication in a scientific journal/conference	Inclusion
English language	Inclusion
Published before 2019	Exclusion
Duplicate articles	Exclusion
Article which do not have open access	Exclusion

2.2 Findings

2.2.1 Decomposition. This section examines the selected literature by analyzing the trends and distribution of publications related to the study of default risk prediction in P2P lending platforms. First, the timeline analysis will be discussed, and further we will evaluate journal distribution of selected articles.

- **Timeline Analysis:** The temporal analysis of the publications allow to get insight into the patterns related to default risk prediction in alternative financing methods. The distribution of papers for the period from 2019 to 2024 among the three subtopics defined in the scope of this study is shown in Figure 6 of Appendix B.1. Despite Green AI and the sustainability of machine learning models (Query 3) were not commonly studied in 2020, there was a sharp rise in the number of publications on the subject starting in 2023. In addition, papers related to the development of explainable AI for default forecasts (Query 2) have been growing in popularity, aligning the interest in application of ML in P2P lending over the past years.
- **Journal Distribution:** The following analysis presents the journal distribution of the selected literature. Such evaluation allows to understand the interest of the researching community in the given subjects. Figure 7 of Appendix B.2 indicates the amount of publications per journal that are related to certain subtopic. According to the presented distribution, IEEE Access journal contains the most publications across all the subtopics. Nevertheless, IEEE Access is an open access, multidisciplinary journal that features articles on a wide range of fields. The remaining articles, however, are almost exclusively published throughout various scientific journals. These trends demonstrate the topic's widespread relevance and significance in today's society.

2.2.2 Themes. This section provides insights obtained from reviewing selected papers regarding the key problems raised by this research. It investigates applications of ML models for default risk predictions, as well as integration of Green AI and XAI in financial industry.

- **Use of Machine Learning for Default Prediction:** In recent years alternative financing methods, such as P2P lending, significantly changed the financial field [22, 23, 26, 31]. Integration of machine learning models to P2P lending platforms allowed to dramatically improve accuracy and efficiency of default risk predictions, in comparison to traditional banking methods [25, 38]. By identifying complex, non-linear patterns in the data that are not always visible to human investors, ML algorithms can improve accuracy of default assessment and hence increase reliability of the lending platforms [14, 16, 26, 37, 38]. Moreover, automated default risk prediction has the potential to significantly reduce time and error-prone manual analysis of borrower data [16, 25, 36]. In addition, previous researches suggest that integration of ML into P2P lending can boost security of the platform by recognizing potential fraudulent loans [16, 19].
- **Understanding Concepts and Integration Approaches of Green AI:** Although there are many benefits to integrating ML into the finance sector, there are also environmental concerns. Green AI focuses on reducing environmental impact of AI [3, 34]. For instance, training a deep learning model could require large amount of electricity and can lead to high carbon emissions [5, 20]. Therefore, the concept of Green AI focuses on the creation of sustainable AI through optimization of energy resource and carbon footprint minimization, while also maintaining model performance [21, 34]. Previous researches

outline a number of strategies to reduce environmental impact of ML models. These include utilizing energy-efficient hardware [4, 21, 39], decreasing the size and computational requirements for the models [4, 11, 17, 27, 30], and optimizing data storage [5]. Furthermore, a number of frameworks are available which allow the integration of sustainability concepts into ML models. For example, Clover framework uses GPU partitioning and mixed-quality models to lower the model's carbon footprint [18]. Moreover, there are tools such as carbon-Aware Federated Learning (CAFE) [5], environmentally sustainable computing (ESC) [29], eco2AI [6] that enable monitoring and optimizing the carbon emissions of machine learning models without sacrificing model performance. Another strategy to address the issue of inefficient energy consumption of ML models is to define guidelines for various phases of the model life cycle (algorithm design, model optimization, model training, etc.) in order to build energy-efficient models [13].

- **Transparency of ML Models for Default Prediction in P2P lending:** The concept of explainable AI (XAI) refers to the approaches that allow humans to understand choices made by models to reach particular outcome [10, 32, 35]. Ensuring the transparency of the machine learning model is crucial to achieving fair outcomes and system reliability [7, 9, 22, 35]. ML models are sometimes referred to as "black boxes" due to the complexity of advanced algorithms [7, 32]. Hence, it is essential to assess their transparency in order to guarantee that the provided reasoning and results are not only accurate but also fair [7, 9, 12]. The reviewed literature provides numerous recommendations for improving the transparency of ML models that are utilized in default prediction. As an example, combining several models can not only increase the performance but also improve its overall interpretability [24, 28]. Additionally, utilizing (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) one could get a deeper understanding of how individual features of a single instance contributed to the default risk of a particular application [8, 10].

3 METHODOLOGY

3.1 Predicting Default Risks in P2P Lending Platforms

The effectiveness of several ML models will be explored in order to analyze default prediction performance in P2P lending platforms. Selected for the experiment ML models include Decision Tree, Random Forest, Logistic Regression, and XGBoost. These four algorithms had been chosen as they are commonly applied in the literature in default prediction tasks as well as have proven to have high accuracy. GridSearchCV[5] will be used to perform hyperparameter tuning in order to prevent overfitting . By exploring various parameter sets, this approach allows to optimize model's settings and improve overall performance. We will also use k-fold cross-validation to verify the results and guarantee consistency of the model in making predictions based on data that has not been observed. To evaluate the accuracy of each of selected models, F1 score and AUC metrics would be applied. The F1 Score (1) aggregates the results of precision and recall, making a good accuracy indicator for imbalanced datasets.

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{1}$$

AUC (Area under the Receiver Operating Characteristic curve) metric represents the ability of models to distinguish between target classes.

3.2 Evaluating Model Sustainability

Evaluation of sustainability of ML models requires to analyze energy consumption and carbon emission produces during the training phase. Increased carbon emissions have a negative impact on the environment and, in terms of AI, are directly correlated with increasing energy consumption. Therefore, this study's main objective is to evaluate the energy consumption of the selected models, which include Decision Tree, Random Forest, Logistic Regression, and XGBoost. The sustainability of the models will be evaluated with the use of `codecarbon`[6] library. Codecarbon calculates environmental impact by converting the hardware energy (CPU,GPU and RAM combined) consumed during model training to an estimate of CO2 emissions.

$$\text{CO2 Emissions} = \gamma \times \text{Energy Consumption} \tag{2}$$

In this context, γ(kg/(kWh) refers to region emission intensity coefficient. Finding the models that maintain a balance between high accuracy and little environmental damage is the aim of the experiment. Through a comparative analysis of each model's energy consumption and carbon emissions, the study will determine the most sustainable models to proceed with transparency evaluation.

3.3 Enhancing Model Transparency

Transparency of ML models plays a critical role in ensuring reliability of the system and building trust with stakeholders. XAI techniques, specifically SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), are utilized in this study to interpret the decision-making process of the chosen models: Decision Tree, Random Forest, Logistic Regression, and XGBoost.

- **SHAP:** SHAP values (3) give insight into the features' overall performance over the entire dataset. The values present the contribution of each feature to the model outcome, which helps to identify the most significant features in the default risk prediction [2].

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \tag{3}$$

[5] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

[6] https://github.com/mlco2/codecarbon

In this context, ϕ_i is the contribution of feature i to the final prediction, N is the set of all features, and S is a subset of N not including feature i.

- **LIME:** LIME analysis offers thorough explanation for particular outcome instances. By tracking the changes in predictions as the model moves from the input data to the result, LIME approximates the behavior of the model. The formula for LIME explanation is given as:

$$\text{LIME Explanation} = \arg\min_{g \in G} \mathcal{L}(f, g, \pi_x) + \Omega(g), \quad (4)$$

where the simple interpretable model g locally approximates the set of potentially interpretable complex models G; $\mathcal{L}$ indicates the precision of g with respect to the original model f, and $\Omega(g)$ represents the complexity of explanation g.

Implementing SHAP and LIME methods improves reliability and fairness in models, which is linked to the purpose of integrating XAI into P2P lending default predictions.

4 EXPERIMENTAL SETUP

4.1 Data Source

This project will be utilizing the data set of a European P2P lending platform Bondora.com[7]. Since the loans description was extracted from existing lending platform, the chosen dataset appears to be highly relevant to the performed experiment. Presented dataset was imported from Kaggle[8] platform. The retrieved data contains information about defaulted and non-defaulted loans from the time period between February 2009 and July 2021, including the details about demographic and financial status of the borrowers as well as their loan transactions. The dataset contains various borrower's characteristics, including their age, work and marital status, annual income, and loan purpose. Additionally, details regarding the loan amount, interest rate, and payback schedule are provided in order to understand the borrower's financial situation. Table 2 represents distribution of loans in dataset per status.

Table 2. Distribution of loans in dataset based on "LoanStatus"

Loan Status	**Number of Loans**
Late	68,574
Repaid	52,887
Current	57,774

4.2 Data Preprocessing

The raw dataset of Bondora Lending Platform has 179235 records and 112 features. In order to use this dataset for ML model training, it is necessary to perform data preprocessing. The following steps were taken to ensure that retrieved data is suitable for the performed study.

(1) **Handling missing values:** the columns containing more than 40% missing values were removed from the dataset [14, 26, 31]. As a result, 37 features were eliminated.
(2) **Removing unrelated features:** features that were intuitively unrelated to default prediction, such as "LoanId", "LoanNumber", "UserName", etc., were removed from the dataset[26, 31]. Also, all the features that are related to the dates were excluded from the dataset, since the study does not use time series analysis for default prediction. Furthermore, records with status of loan marked as "Current" were removed as well, since these loans did not mature to defaulted or repaid.
(3) **Creation of target variable:** The dataset do not explicitly specify the defaulted and non-defaulted loans [31]. However, there were two features that might be related to the default of the loan: "Status", and "DefaultDate". These features were combined to define a new feature "Defaulted". If the loan status set as "Late" but have no default date, this loan will be considered as non-defaulted. Also, if loan has a status of "Repaid", it will be considered non-defaulted independent from the "DefaultDate" value. The final dimension of the dataset was reduced to 121461 records and 51 attributes.
(4) **Re-coding categorical variables that have integer values:** The dataset contained a number of features that are marked as numerical, however are categorical but coded with numbers. These features were re-coded into categorical according to data description[26, 31].
(5) **Handling null values:** All the null values for categorical variables were renamed as "Unknown". Null values of numerical attributes were replaced to the median value of the variables within the column since distribution of most of the columns is skewed [31].
(6) **Label Encoder:** categorical variables were encoded via Label Encoder method of `scikit-learn`[9] library for Python.

4.3 Models Implementation

Once data preprocessing was performed for the Bondora dataset the feature selection [26] is carried out to identify features that are most relevant to a target variable, and reduce the dimension of the dataset. For this purpose the method of Mutual Information(MI) [15]. Mutual Information is represented as non-negative value describing the statistical dependency of different variables. High MI indices show high correlation between the variables of the dataset. Choosing dependant variables allows to predict target variable more accurately. As illustrated on Figure 1, 15 features with largest statistical dependency towards target variable were selected to proceed with training ML models.

[7]https://www.bondora.com.
[8]https://www.kaggle.com/datasets/sid321axn/bondora-peer-to-peer-lending-loan-data/data
[9]https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.LabelEncoder.html

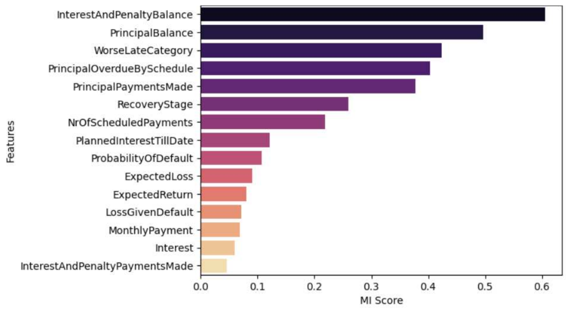

Fig. 1. Mutual Information Scores

In order to evaluate dafault risks of P2P lending the ML models with focus on prediction and classification were utilized. Hence, the selected models include Decision Tree, Random Forest, Logistic Regression, and XGBoost.

- **Decision Tree:** Decision Tree is commonly used in machine learning for classification and regression tasks. The algorithm learns from rules derived from a features set. Hence, decision trees consist of nodes representing attributes, and edges representing a rule the decision is based on. The ML algorithm recursively splits the training data into smaller components based on the feature predictions, until the algorithm decides whether presented data described "Defaulted" or "Non-Defaulted" loan [22].
- **Random Forest:** This tree-based algorithm generates numerous decision trees based on random sets of features [22]. The final prediction, whether specified loan is at risk of default, is determined from the decisions of all the trees using some voting algorithm.
- **Logistic Regression:** Logistic Regression, an extension of the Linear Regression algorithm, is used to predict the probability of an outcome. It is a classification algorithm that calculates the likelihood of an instance belonging to one of the specified classes. The model uses a sigmoid function, which converts a combination of input variables into a probability value between 0 and 1. In this study, we utilize binary Logistic Regression, meaning the classification will occur between two classes: "Defaulted" and "Non-Defaulted". Logistic Regression, an extension of the Linear Regression algorithm, is used to predict the probability of an outcome. It is a classification algorithm that calculates the likelihood of an instance belonging to one of the specified classes. The model uses a sigmoid function, which converts a combination of input variables into a probability value between 0 and 1 [22]. The sigmoid function is defined as:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

where z is the linear combination of features:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$$

In this study, we utilize binary Logistic Regression, meaning the classification will occur between two classes: "Defaulted" and "Non-Defaulted".

- **XGBoost:** XGBoost, or Extreme Gradient Boosting, is well-known for its predictive performance and computational efficiency. The algorithm is based on gradient-boosted decision trees, where each new tree attempts to correct the errors of the previous trees in each iteration. This process involves optimizing a cost function by adjusting the model's parameters based on the gradient of the loss function. Additionally, the algorithm runs multiple iterations concurrently, ensuring its computational efficiency.

GridSearchCV[5] was implemented to tune the hyperparameters of the models in order to maximize performance. This strategy allows to search through sets of predefined hyperparameters and identify configuration that presents best performance metrics on training data. In order to guarantee the models' generalizability and good performance on the unknown data, k-fold cross validation[10] was applied with k=5 (one test set and four training sets at each fold).

5 RESULTS AND DISCUSSIONS

This section discusses the results obtained during the models training phase. First the performance of the model and accuracy of default predictions will be explained. Further, the article evaluates sustainability and transparency reports in order to determine which models not only accurately predict the risk of default in peer-to-peer lending but also follow the concepts of Green AI.

5.1 Accuracy Evaluation

The study included training and evaluation of four ML models: Decision Tree, Random Forest, Logistic Regression, and XGBoost. In order to evaluate performance of these models F1 and AUC scores were used. From the perspective of indicated metrics all the models show high performance and overcame the threshold of 95%, as described in Table 3.

Table 3. Performance metrics of selected models

Model	**F1 Score**	**AUC**
Decision Tree	0.997028	0.999154
Random Forest	0.997729	0.999812
Logistic Regression	0.966812	0.991722
XGBoost	0.997976	0.999917

However, while Logistic Regression algorithm shows high results, its F1 score appears to be the lowest (96,68%). Since F1 score is based on the precision and recall metrics, the value related to LR algorithm indicates that this model has the largest number of misclassified loans. XGBoost and Random Forest algorithms showed the highest results classifying default and non-default loans of P2P lending platform.

5.2 Sustainability evaluation

As was outlined earlier, Green AI emphasized the implementation of ML models that would not only provide accurate

[10] https://scikit-learn.org/stable/modules/generated/sklearn.model_-selection.KFold.html

results but will be computationally efficient [13]. The objective of the conducted experiment was to evaluate the environmental impact of each model individually and provide insight into environmental sustainability, mainly carbon footprint and energy consumption, of chosen algorithms. Due to limited time and resources of the experiment, we used Python library `CodeCarbon`[11] to receive necessary measurements.

The result presented in Figure 2 shows that Logistic Regression required the least amount of resources (Energy, CPU, and Memory usage) as well as produced the least amount of carbon emissions (8.735222 × 10-7 kg CO2). On the contrary, Random Forest was the most computationally inefficient, producing 1.753160 × 10-4 kg CO2.

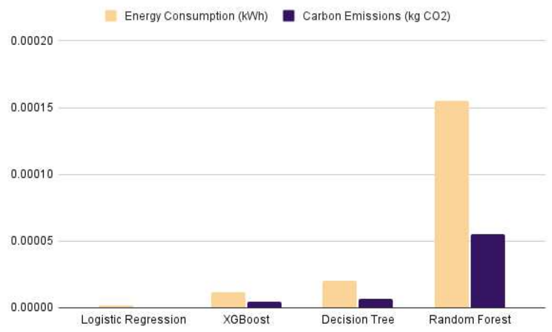

Fig. 2. Environmental footprint of ML models

Overall, the evaluation of the sustainability of the chosen ML models indicates a considerable trade-off between environmental impact and prediction accuracy. Despite having the best default prediction accuracy, Random Forest appears to be the least sustainable model producing the highest rates of carbon emissions. On the contrary, despite the fact that Logistic regression achieved the lowest performance, it appears to be the most environmentally friendly model as it produces the least amount of carbon emissions.

5.3 Transparency Evaluation

Since the goal of the experiment is to determine models that would both be accurate in predictions and have the least environmental impact, only several models were selected for transparency analysis. The decision was based on the previously assessed performance of the models. Hence, for explainability evaluation Decision Tree, Logistic Regression, and XGBoost models were handpicked.

In order to assess the transparency of the models, the methods of XAI were utilized. This experiment focuses on SHAP [13, 22] and LIME [22] methods to understand how certain features impact the prediction of default in P2P lending platforms.

- **Decision Tree:** Using the bar plot of `shap`[12] library, we could get insight into the feature contribution into the final result of the model . Feature importance distribution of Decision Tree model show "PrincipalBalance", "RecoveryStage", and "InterestAndPenaltyBalance" as the attributes that impact prediction the most, with 0.26, 0.12, 0.1 mean absolute SHAP values respectively (see Figure 8). Moreover, when evaluating a single defaulted loan, LIME analysis highlighted the same features as top three variables that helped to classify the selected loan with 100% confidence, as specified in Figure 11.
- **Logistic Regression:** Feature importance distribution for Logistic Regression model was similar to the one obtained for DT algorithm, with "InterestAndPenaltyBalance" and "RecoveryStage" as mainly contributing features. However, while mean absolute of "InterestAndPenaltyBalance" for DT algorithm was 0.26, the mean SHAP value for LR model reaches 11.11 (see Figure 9). Additionally, LIME analysis outline "InterestAndPenaltyBalance", "RecoveryStage", and "PrincipalBalance" as core attributes in classifying the single loan as "Defaulted" with 100% certainty (see Figure 12)
- **XGBoost:** Bar plot for XGBoost model revealed "PrincipalBalance", "InterestAndPenaltyBalance", and "RecoveryStage" to be the most important features in default prediction, with mean absolute SHAP values 4.95, 3.8, 2.04 respectively (see Figure 10). This result is identical to the Decision Tree bar plot with the only difference in magnitude of mean values. LIME analysis confirmed "RecoveryStage", "PrincipalBalance", and "InterestAndPenaltyBalance" to be the main attributes in single instance classification, as described in Figure 13.

To summarize the transparency report, the features "PrincipalBalance", "RecoveryStage", and "InterestAndPenaltyBalance" appeared the most important features positively contributing to the prediction of defaulted loans across all the models. According to both SHAP and LIME evaluations, these characteristics are crucial in default prediction in P2P lending platforms based on the Bondora dataset. "PrincipalBalance" value indicates the borrower's remaining loan debt. "RecoveryStage" denotes the stage at which the recovery process is in progress; in other words, the feature tracks the progression of debt payback. "InterestAndPenaltyBalance" indicates the interest and penalties associated with the loan, which shows the financial stress of the borrower. By identifying the most significant features, the study enhances the transparency of ML models used for default prediction in P2P lending platforms. The results provided via SHAP and LIME analyses allow not only improve reliability of the system but to provide valuable information about decision-making process of the models to stakeholders.

6 CONCLUSION

The research focused on the application of ML models for default prediction on P2P lending platforms, with an emphasis on model transparency and environmental sustainability. The results demonstrate that while all four models — Decision Tree, Random Forest, Logistic Regression, and XGBoost

[11] https://codecarbon.io/
[12] https://shap.readthedocs.io/en/latest/index.html

— achieve high accuracy, their environmental impacts differ significantly. As was reported during experiment, despite its high performance, Random Forest model requires a lot of energy resources and emits large amounts CO2. Controversially, Logistic Regression model, although less accurate, showed a smaller environmental impact, highlighting the trade-off between sustainability and model performance. Furthermore, integration of XAI techniques such as SHAP and LIME improved interpretability and reliability of the models, allowing stakeholders to gain insight into the decision making process of the default prediction models.

6.1 Limitations

This research has a number of limitations. Firstly, the Bondora dataset was obtained from a European P2P lending company. Therefore, the study may not accurately reflect the worldwide financial situation within P2P lending and may rather restrict its findings to a specific geographic region. Furthermore, depending too much on a single platform may result in system-specific biases and limit the relevance of the findings to other lending platforms. Furthermore, there could possibly exist alternative algorithms for default prediction with better performance or carbon efficiency than the four ML models selected for the experiment.

6.2 Future Work and Recommendations

Further studies could focus on the development of ML models that offer a good compromise between high accuracy in predictions and minimal environmental impact. Exploration and development of new algorithms that maintain high accuracy without utilizing more computer power should become a primary objective for future works. By utilizing such techniques as pruning, quantization, and considering use of renewable energy sources, new ML algorithms would not only become computationally efficient but also align with concepts of Green AI and global sustainability goals.

REFERENCES

[1] Alessandra Amato, Joerg R. Osterrieder, and Marcos R. Machado. 2024. How can artificial intelligence help customer intelligence for credit portfolio management? A systematic literature review. *International Journal of Information Management Data Insights* 4, 2 (Nov. 2024), 100234. https://doi.org/10.1016/j.jjimei.2024.100234

[2] G. Babaei and P. Giudici. 2024. InstanceSHAP: an instance-based estimation approach for Shapley values. *Behaviormetrika* 51, 1 (2024), 425–439. https://doi.org/10.1007/s41237-023-00208-z

[3] E. Barbierato and A. Gatti. 2024. Toward Green AI: A Methodological Survey of the Scientific Literature. *IEEE Access* 12 (2024), 23989–24013. https://doi.org/10.1109/ACCESS.2024.3360705

[4] Brian R. Bartoldson, Bhavya Kailkhura, and Davis Blalock. 2024. Compute-efficient deep learning: algorithmic trends and opportunities. *J. Mach. Learn. Res.* 24, 1 (March 2024). Publisher: JMLR.org.

[5] Jieming Bian, Lei Wang, Shaolei Ren, and Jie Xu. 2024. CAFE: Carbon-Aware Federated Learning in Geographically Distributed Data Centers. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems (e-Energy '24)*. Association for Computing Machinery, New York, NY, USA, 347–360. https://doi.org/10.1145/3632775.3661970 event-place: <conf-loc>, <city>Singapore</city>, <country>Singapore</country>, </conf-loc>.

[6] S.A. Budennyy, V.D. Lazarev, N.N. Zakharenko, A.N. Korovin, O.A. Plosskaya, D.V. Dimitrov, V.S. Akhripkin, I.V. Pavlov, I.V. Oseledets, I.S. Barsola, I.V. Egorov, A.A. Kosterina, and L.E. Zhukov. 2022. eco2AI: Carbon Emissions Tracking of Machine Learning Models as the First Step Towards Sustainable AI. *Doklady Mathematics* 106 (2022), S118–S128. https://doi.org/10.1134/S1064562422060230

[7] M. Bücker, G. Szepannek, A. Gosiewska, and P. Biecek. 2022. Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society* 73, 1 (2022), 70–90. https://doi.org/10.1080/01605682.2021.1922098

[8] Y. Chen, R. Calabrese, and B. Martin-Barragan. 2024. Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research* 312, 1 (2024), 357–372. https://doi.org/10.1016/j.ejor.2023.06.036

[9] X. Dastile and T. Celik. 2021. Making Deep Learning-Based Predictions for Credit Scoring Explainable. *IEEE Access* 9 (2021), 50426–50440. https://doi.org/10.1109/ACCESS.2021.3068854

[10] P.E. de Lange, B. Melsom, C.B. Vennerød, and S. Westgaard. 2022. Explainable AI for Credit Assessment in Banks. *Journal of Risk and Financial Management* 15, 12 (2022). https://doi.org/10.3390/jrfm15120556

[11] M. Dorrich, M. Fan, and A.M. Kist. 2023. Impact of Mixed Precision Techniques on Training and Inference Efficiency of Deep Neural Networks. *IEEE Access* 11 (2023), 57627–57634. https://doi.org/10.1109/ACCESS.2023.3284388

[12] P. Giudici and E. Raffinetti. 2023. SAFE Artificial Intelligence in finance. *Finance Research Letters* 56 (2023). https://doi.org/10.1016/j.frl.2023.104088

[13] Heli Järvenpää, Patricia Lago, Justus Bogner, Grace Lewis, Henry Muccini, and Ipek Ozkaya. 2024. A Synthesis of Green Architectural Tactics for ML-Enabled Systems. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS'24)*. Association for Computing Machinery, New York, NY, USA, 130–141. https://doi.org/10.1145/3639475.3640111 event-place: <conf-loc>, <city>Lisbon</city>, <country>Portugal</country>, </conf-loc>.

[14] Po-Chang Ko, Ping-Chen Lin, Hoang-Thu Do, and You-Fu Huang. 2022. P2P Lending Default Prediction Based on AI and Statistical Models. *Entropy* 24, 6 (2022). https://doi.org/10.3390/e24060801

[15] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. 2004. Estimating mutual information. *Physical review. E, Statistical, nonlinear, and soft matter physics* 69, 6 Pt 2 (June 2004), 066138. https://doi.org/10.1103/PhysRevE.69.066138 Place: United States.

[16] L. Li, Y. Feng, Y. Lv, X. Cong, X. Fu, and J. Qi. 2019. Automatically Detecting Peer-to-Peer Lending Intermediary Risk—Top Management Team Profile Textual Features Perspective. *IEEE Access* 7 (2019), 72551–72560. https://doi.org/10.1109/ACCESS.2019.2919727

[17] L. Lannelongue, J. Grealey, and M. Inouye. 2021. Green Algorithms: Quantifying the Carbon Footprint of Computation. *Advanced Science* 8, 12 (2021). https://doi.org/10.1002/advs.202100707

[18] Baolin Li, Siddharth Samsi, Vijay Gadepally, and Devesh Tiwari. 2023. Clover: Toward Sustainable AI with Carbon-Aware Machine Learning Inference Service. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3581784.3607034 event-place: <conf-loc>, <city>Denver</city>, <state>CO</state>, <country>USA</country>, </conf-loc>.

[19] Z. Liu, Z. Zhang, H. Yang, G. Wang, and Z. Xu. 2023. An innovative model fusion algorithm to improve the recall rate of peer-to-peer lending default customers. *Intelligent Systems with Applications* 20 (2023). https://doi.org/10.1016/j.iswa.2023.200272

[20] Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat. 2024. Estimating the carbon footprint of BLOOM, a 176B parameter language model. *J. Mach. Learn. Res.* 24, 1 (March 2024). Publisher: JMLR.org.

[21] Tao Luo, Weng-Fai Wong, Rick Siow Mong Goh, Anh Tuan Do, Zhixian Chen, Haizhou Li, Wenyu Jiang, and Weiyun Yau. 2023. Achieving Green AI with Energy-Efficient Deep Learning Using Neuromorphic Computing. *Commun. ACM* 66, 7 (June 2023), 52–57. https://doi.org/10.1145/3588591 Place: New York, NY, USA Publisher: Association for Computing Machinery.

[22] M. J. Ariza-Garzón, J. Arroyo, A. Caparrini, and M. -J. Segovia-Vargas. 2020. Explainability of a Machine Learning Granting Scoring Model in Peer-to-Peer Lending. *IEEE Access* 8 (2020), 64873–64890. https://doi.org/10.1109/ACCESS.2020.2984412

[23] Marcos R. Machado and Salma Karray. 2022. Integrating Customer Portfolio Theory and the Multiple Sources of Risk Approaches to Model Risk-Adjusted Revenue. *18th IFAC Workshop on Control Applications of Optimization CAO 2022* 55, 16 (Jan. 2022), 356–363. https://doi.org/10.1016/j.ifacol.2022.09.050

[24] P. Nooji and S. Sugave. 2024. Explainable ensemble technique for enhancing credit risk prediction. *IAES International Journal of Artificial Intelligence* 13, 1 (2024), 917–924. https://doi.org/10.11591/ijai.v13.i1.pp917-924

[25] E. Owusu, R. Quainoo, S. Mensah, and J.K. Appati. 2023. A Deep Learning Approach for Loan Default Prediction Using Imbalanced Dataset. *International Journal of Intelligent Information Technologies* 19, 1 (2023).

https://doi.org/10.4018/IJIIT.318672
[26] P. Mukherjee and Y. Badr. 2022. Detection of Defaulters in P2P Lending Platforms using Unsupervised Learning. In *2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*. 1–5. https://doi.org/10.1109/COINS54846.2022.9854964 Journal Abbreviation: 2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS).
[27] D. Patterson, J.M. Gilbert, M. Gruteser, E. Robles, K. Sekar, Y. Wei, and T. Zhu. 2024. Energy and Emissions of Machine Learning on Smartphones vs. the Cloud. *Commun. ACM* 67, 2 (2024), 86–97. https://doi.org/10.1145/3624719
[28] N. Pavitha and S. Sugave. 2024. Explainable Multistage Ensemble 1D Convolutional Neural Network for Trust Worthy Credit Decision. *International Journal of Advanced Computer Science and Applications* 15, 2 (2024), 351–358. https://doi.org/10.14569/IJACSA.2024.0150237
[29] A. Pazienza, G. Baselli, D.C. Vinci, and M.V. Trussoni. 2024. A holistic approach to environmentally sustainable computing. *Innovations in Systems and Software Engineering* (2024). https://doi.org/10.1007/s11334-023-00548-9
[30] Xinchi Qiu, Titouan Parcollet, Javier Fernandez-Marques, Pedro P. B. Gusmao, Yan Gao, Daniel J. Beutel, Taner Topal, Akhil Mathur, and Nicholas D. Lane. 2024. A first look into the carbon footprint of federated learning. *J. Mach. Learn. Res.* 24, 1 (March 2024). Publisher: JMLR.org.
[31] S. Mondal, S. K. Shah, and V. Kumbhar. 2023. Predicting Credit Risk in European P2P Lending: A Case Study of "Bondora" Using Supervised Machine Learning Techniques. In *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*. 1–6. https://doi.org/10.1109/GCAT59970.2023.10353353 Journal Abbreviation: 2023 4th IEEE Global Conference for Advancement in Technology (GCAT).
[32] M. Sahakyan, Z. Aung, and T. Rahwan. 2021. Explainable Artificial Intelligence for Tabular Data: A Survey. *IEEE Access* 9 (2021), 135392–135422. https://doi.org/10.1109/ACCESS.2021.3116481
[33] Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni. 2020. Green AI. *Commun. ACM* 63, 12 (Nov. 2020), 54–63. https://doi.org/10.1145/3381831 Place: New York, NY, USA Publisher: Association for Computing Machinery.
[34] Aimee van Wynsberghe. 2021. Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 1, 3 (Aug. 2021), 213–218. https://doi.org/10.1007/s43681-021-00043-6
[35] P. Weber, K.V. Carl, and O. Hinz. 2024. Applications of Explainable Artificial Intelligence in Finance—a systematic review of Finance, Information Systems, and Computer Science literature. *Management Review Quarterly* 74, 2 (2024), 867–907. https://doi.org/10.1007/s11301-023-00320-0
[36] J. Xu, Z. Lu, and Y. Xie. 2021. Loan default prediction of Chinese P2P market: a machine learning methodology. *Scientific Reports* 11, 1 (2021). https://doi.org/10.1038/s41598-021-98361-6
[37] Y. Xia. 2019. A Novel Reject Inference Model Using Outlier Detection and Gradient Boosting Technique in Peer-to-Peer Lending. *IEEE Access* 7 (2019), 92893–92907. https://doi.org/10.1109/ACCESS.2019.2927602
[38] Z. Zhang, K. Niu, and Y. Liu. 2020. A Deep Learning Based Online Credit Scoring Model for P2P Lending. *IEEE Access* 8 (2020), 177307–177317. https://doi.org/10.1109/ACCESS.2020.3027337
[39] Dan Zhao, Siddharth Samsi, Joseph McDonald, Baolin Li, David Bestor, Michael Jones, Devesh Tiwari, and Vijay Gadepally. 2023. Sustainable Supercomputing for AI: GPU Power Capping at HPC Scale. In *Proceedings of the 2023 ACM Symposium on Cloud Computing (SoCC '23)*. Association for Computing Machinery, New York, NY, USA, 588–596. https://doi.org/10.1145/3620678.3624793 event-place: <conf-loc>, <city>Santa Cruz</city>, <state>CA</state>, <country>USA</country>, </conf-loc>.

A ARTICLES SELECTION PROCESS

A.1 Query 1: Potential application of ML to address default risks

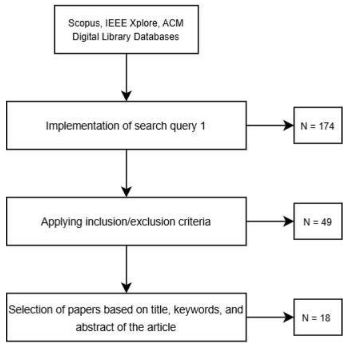

Fig. 3. Illustration of articles selection process for Query 1

A.2 Query 2: Ensuring transparency of ML model that predicts default risks

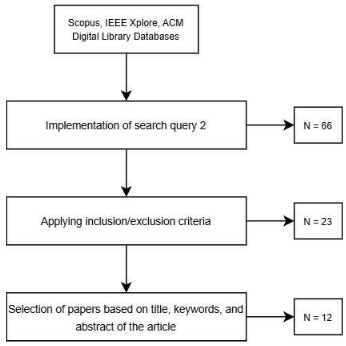

Fig. 4. Illustration of articles selection process for Query 2

A.3 Query 3: Assessing sustainability of ML model following the principles of Green AI

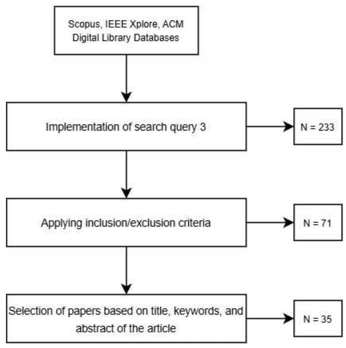

Fig. 5. Illustration of articles selection process for Query 3

B DISTRIBUTION OF ARTICLES

B.1 Time distribution of articles

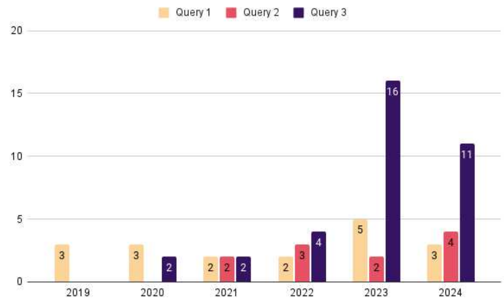

Fig. 6. Trends of publications with related queries

B.2 Journal distribution of articles

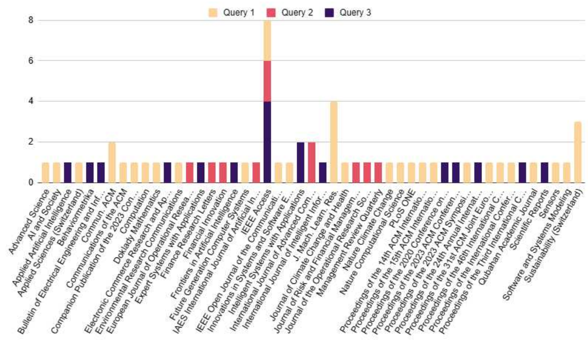

Fig. 7. Journal distribution with related queries

C TRANSPARENCY EVALUATION OF ML MODELS

C.1 SHAP

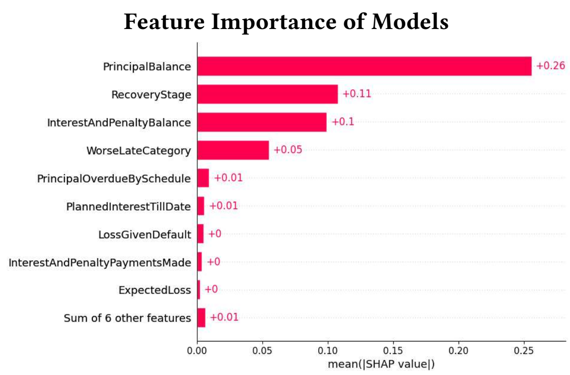

Fig. 8. Feature Importance of Decision Tree Model

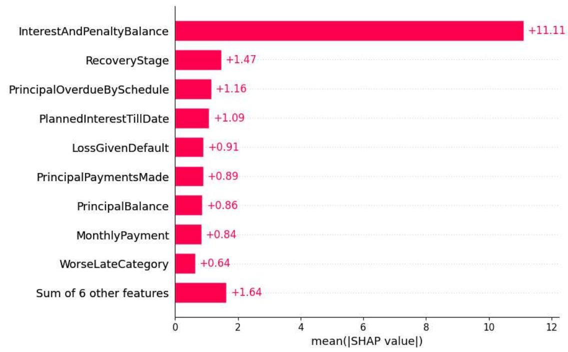

Fig. 9. Feature Importance of Logistic Regression Model

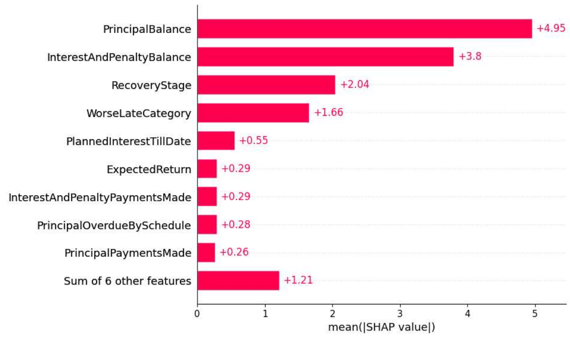

Fig. 10. Feature Importance of XGBoost Model

C.2 LIME

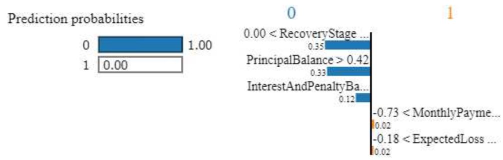

Fig. 11. LIME analysis of data instance prediction marked as "Defaulted" by Decision Tree algorithm

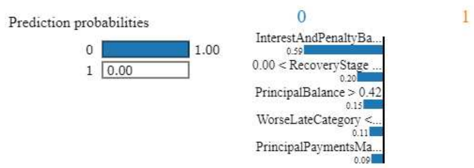

Fig. 12. LIME analysis of data instance prediction marked as "Defaulted" by Logistic Regression algorithm

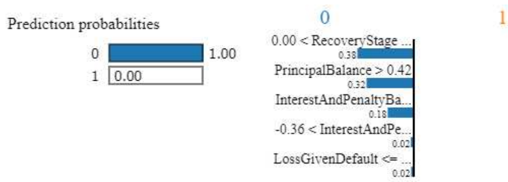

Fig. 13. LIME analysis of data instance prediction marked as "Defaulted" by XGBoost algorithm