Journal of Information Systems Engineering and Management
2025, 10(27s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Explainable AI in Credit Scoring: Improving Transparency in Loan Decisions

Ajkuna MUJO[1], Sonila NIKOLLA[1/2], Eneida HOXHA[1], Elis PELIVANI[1]
[1] *Mediterranean University of Albania*
[1]*Department of Business Informatics, Tirana, Albania;*
[1/2]*Department of Finance, Banking and Accounting, Tirana, Albania;*
email: ajkuna.mujo@umsh.edu.al

ARTICLE INFO | **ABSTRACT**

Received: 18 Dec 2024
Revised: 15 Feb 2025
Accepted: 28 Feb 2025

The increasing dependence on Artificial Intelligence (AI) in the realm of credit scoring has led to notable enhancements in loan approval processes, particularly with regard to accuracy, efficiency, and risk evaluation. Yet, due to the opacity of sophisticated AI models, there are worries regarding transparency, fairness, and adherence to regulations. Because traditional black-box models like deep learning and ensemble methods are not interpretable, financial institutions find it challenging to justify credit decisions based on them. This absence of transparency creates difficulties in complying with regulatory standards such as Basel III, the Fair Lending Act, and GDPR, while also heightening the risk of biased or unjust lending practices. This study examines the role of Explainable AI (XAI) in credit scoring to tackle these issues, concentrating on methods that improve model interpretability while maintaining predictive performance.

This study puts forward a credit scoring framework driven by XAI, which combines Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) to enhance the transparency of AI-based loan decision-making. Machine learning models such as random forests, gradient boosting, and neural networks are evaluated for their accuracy and explainability using real-world credit risk datasets. The results demonstrate that although AI improves risk prediction, post-hoc interpretability techniques effectively identify the key factors affecting loan approvals, thereby promoting trust and adherence to regulations. This research emphasizes how XAI can reduce bias, enhance fairness, and foster transparency in credit decision-making. These developments open the door to more ethical and accountable AI-based financial systems.

Keywords: Explainable AI (XAI), Credit Scoring, Machine Learning, Loan Decision-Making, Financial Risk Assessment

INTRODUCTION

Background and Motivation

The swift uptake of Artificial Intelligence (AI) in financial services has changed how banks and lending institutions evaluate credit risk. Structured financial data has long been used to assess an applicant's creditworthiness using traditional credit scoring models like decision trees and logistic regression. The growing intricacy of financial markets and the access to extensive datasets have resulted in the broad adoption of AI-based credit scoring models. These models utilize machine learning (ML) methods, including neural networks, random forests, and gradient boosting, to improve prediction accuracy, automate decision-making, and optimize the loan approval process(Ajkuna Mujo, Eneida Hoxha, 2023). This leads to quicker, data-informed lending decisions by financial institutions, which in turn mitigates the risks of default and enhances the efficiency of banking operations.

Even with these improvements, models for credit scoring based on AI frequently operate as "black boxes." This implies that their processes of making decisions are not transparent and hard to understand. The lack of clear explanations from many high-performing ML models for loan approvals or rejections complicates banks' ability to justify these decisions to regulators, auditors, and consumers. This absence of transparency gives rise to major ethical, regulatory, and trust-related issues, especially in instances where AI-generated decisions place certain groups at an unjust disadvantage due to training data biases or limitations in algorithms. Regulations like Basel III, the Fair Lending Act, and the General Data Protection Regulation (GDPR) stress the importance of accountability and fairness in AI-driven financial systems. To tackle these issues, there is an urgent requirement for Explainable AI (XAI) solutions that improve transparency, offer justifiable decision-making insights, and guarantee adherence to financial regulations. Financial institutions can connect AI-driven automation with the need for credit scoring models that are interpretable, ethical, and trustworthy by incorporating XAI techniques.

Problem Statement

Financial institutions have relied on traditional credit scoring models, like logistic regression and decision trees, for a long time to evaluate borrowers' creditworthiness. Lenders can comprehend the elements affecting the approval or rejection of loans, as these models are interpretable. Nonetheless, their dependence on linear relationships and oversimplified assumptions restricts their capacity for prediction in intricate financial contexts. Consequently, credit-scoring models driven by AI—such as random forests, neural networks, and ensemble methods—have arisen as better options because they can analyze large quantities of financial data, uncover concealed patterns, and enhance the precision of risk forecasts.

Although they perform better than traditional models, AI-based credit scoring methods operate as black boxes, offering minimal insight into the decision-making process. This absence of interpretability presents major difficulties for borrowers, regulators, and financial institutions. Regulatory frameworks such as Basel III, the General Data Protection Regulation (GDPR), and the Fair Lending Act stress the need for transparency, fairness, and accountability in automated lending decisions. Lenders face compliance risks and consumers may suffer from unjust credit decisions without adequate explanation when there is a lack of clear understanding regarding the reasons for loan approval or denial.

Research Objectives

This study aims to create a credit scoring system based on Explainable AI (XAI) that improves transparency and accountability in loan decisions made by AI. Although AI models greatly enhance credit risk prediction, their lack of interpretability presents regulatory and ethical challenges. This research aims to address this gap by incorporating explainability techniques that enhance the transparency and comprehensibility of AI-driven decisions for financial institutions, regulators, and consumers.

To achieve this goal, the research will:

- Create an XAI-Based Credit Scoring Framework
- Assess Various XAI Methods
- Assess the Trade-off Between Accuracy and Interpretability
- Propose a Policy Framework for AI-Driven Credit Scoring

LITERATURE REVIEW

Traditional credit scoring models have been widely used in banking and financial institutions to assess the creditworthiness of borrowers. These models produce credit scores based on structured data, historical financial information, and established mathematical relationships. The FICO score, created by the Fair Isaac Corporation, is the most widely recognized credit scoring system. It takes into account factors like payment history, credit utilization, length of credit history, new credit accounts, and credit mix. (Elena Dumitrescu, Sullivan Hué, Christophe Hurlin,Sessi Tokpavi, 2022,) Because of its ease of use, interpretability, and efficacy in binary classification issues (such as identifying whether a borrower will default), the logistic regression (LR) model has been among the most frequently employed statistical techniques for credit scoring. Decision trees (DTs) have been utilized for their capability to capture non-linear relationships and produce decision rules that are easy to interpret (Mathew, 2017).Although traditional credit scoring models provide clear and interpretable decision-making, they have significant limitations. Due to its assumption of a linear connection between independent variables and the probability of default, logistic regression is not as effective in addressing complex financial behaviors. Although they are more flexible, decision trees are susceptible to overfitting and may produce unstable predictions when trained on small datasets. Although the FICO scoring system is commonly employed, it is not transparent due to its proprietary algorithm being undisclosed. This lack of transparency complicates consumers' and regulators' understanding of score determination. (Yu, 2020) Additionally, these conventional approaches have difficulty handling large-scale, unstructured, and high-dimensional financial data, which restricts their capacity to capture the subtleties of contemporary credit risk evaluation. These limitations have prompted the use of AI-based machine-learning models, which offer the prospect of greater accuracy but bring challenges concerning interpretability and fairness. (Dionisios N. Sotiropoulos , Gregory Koronakos , Spyridon V. Solanakis, 2024)

Credit scoring has been significantly transformed by Artificial Intelligence (AI) and Machine Learning (ML), which have improved predictive accuracy, automated risk assessments, and enabled the processing of large volumes of financial data. More advanced AI-driven methods, such as deep learning, ensemble methods, and support vector machines (SVMs), have replaced traditional statistical models like logistic regression and decision trees. These models employ sophisticated algorithms to examine large volumes of both structured and unstructured data, pinpointing patterns that conventional scoring methods frequently fail to detect. (Eyad Btoush , Xujuan Zhou ,Raj Gururajan Ka Ching Chan , Omar Alsodi , 2025)

Neural networks, especially feedforward neural networks (FNNs) and recurrent neural networks (RNNs), have demonstrated significant promise in credit risk assessment by identifying nonlinear relationships within financial data. Nevertheless, deep learning models usually need extensive datasets and are not easily interpretable, which raises issues regarding regulatory adherence and transparency. (Xiaoqi Meng, Lihan Jia, Shuo Chen, Yu Zhou, Chang Liu, 2024).

Random Forest (RF), which is an ensemble learning technique, combines several decision trees in order to mitigate overfitting and enhance stability during credit risk evaluations. (Rishab Pendalwar, Aditya Verma, Ratna Patil, 2024). Support Vector Machines (SVMs) are utilized in credit scoring for binary classification tasks, including the determination of loan approval or rejection, relying on support vectors and hyperplane optimization (Tamanna, Shivani Kamboj, Lovedeep Singh, Tanvir Kaur, 2024)

Although they have benefits, AI-driven credit scoring models encounter various challenges:

- Lack of Transparency: Numerous AI models function as black boxes, complicating the explanation of decision-making processes for financial institutions
- Data Bias: AI models can adopt biases present in historical financial data, resulting in discriminatory lending practices (e.g., bias against specific demographic groups).
- Overfitting: Models with high complexity, particularly deep learning models, are susceptible to overfitting. This means they excel on training data but do not generalize effectively to new loan applicants

In order to guarantee fairness, interpretability, and adherence to regulations in AI-driven credit scoring, these challenges underscore the necessity of Explainable AI (XAI) techniques like SHAP, LIME, and counterfactual explanations.

METHODOLOGY

1. Dataset and Preprocessing

The quality of the data used for training has a significant impact on the accuracy and reliability of AI-driven credit scoring models. This study utilizes publicly available datasets that are commonly used in credit risk assessment and loan default prediction to ensure robustness. The preprocessing phase consists of data cleaning, feature engineering, and addressing data imbalance to improve model performance and fairness.

Public Datasets for Credit Scoring

German Credit Dataset (UCI Machine Learning Repository)

Contains 1,000 loan applications with 20 financial attributes, including credit amount, loan duration, employment status, and age.

The target variable is binary classification: good credit risk (1) vs. bad credit risk (0).

Source: UCI Repository

2. Data Cleaning and Preprocessing

The dataset is preprocessed with the following steps to make it suitable for machine learning models: Dealing with Missing Data It is common for financial datasets from the real world to contain missing values, as a result of incomplete records for borrowers. This is tackled using a number of techniques: Numerical Features: For absent financial indicators like income and loan amount, imputation using mean/median values will be applied.

Categorical Features: Use mode imputation or one-hot encoding for categorical values like employment type or loan purpose.

Advanced Techniques: K-Nearest Neighbors Imputation (KNN) or predictive modeling for estimating missing values in essential features.

Feature engineering

By generating informative variables, feature engineering enhances model performance:

- Debt-to-Income Ratio (DTI) = Total Debt ÷ Annual Income → Aids in evaluating the borrower's capacity to repay.
- Credit Utilization Rate = Used Credit / Available Credit → Reflects credit behavior.
- Loan-to-Income Ratio (LTI) = Loan Amount ÷ Annual Income → Assesses risk exposure.
- Credit score categories based on derivation: Convert numerical credit scores into bins .
- Encoding of categorical variables: Transform non-numeric data with one-hot encoding or ordinal encoding.

Handling Imbalanced Data

Datasets used for credit scoring often experience class imbalance, with non-defaulters outnumbering defaulters significantly. To ensure the dataset is balanced and to avoid bias in model predictions, the following techniques are utilized:

- Oversampling Methods: SMOTE (Synthetic Minority Over-sampling Technique) creates synthetic instances of the minority class (defaulters).
- Undersampling Methods: Balances class distribution by randomly eliminating samples from the majority class (non-defaulters).
- Learning with Cost Sensitivity: To mitigate bias, it imposes greater penalties for misclassification on the minority class (the group of defaulters).

3. Model Selection & Implementation

We utilize a mix of baseline models and sophisticated AI models to create a strong credit scoring system. Moreover, in order to improve transparency and explainability, we incorporate SHAP, LIME, and Counterfactual Explanations. Baseline models will act as benchmarks for comparison with more sophisticated AI models.We utilize:

- Logistic Regression (LR) – A commonly utilized statistical model for evaluating credit risk.
- Decision Tree (DT) – A straightforward tree-based model that offers decision rules easy for humans to understand.

While these models can be understood, they may not have strong predictive capabilities when applied to complex financial data.

We will utilize machine learning and deep learning models to enhance precision and capture intricate patterns:

XGBoost

Ensemble model based on boosting, recognized for its high accuracy and capability to manage missing values.

Performs effectively on datasets with imbalanced distributions as well as on extensive financial data.

Implementation: Employs decision trees within a boosting framework.

Random Forest (RF)

A collection of decision trees, which diminishes overfitting and enhances stability.

Offers feature importance scores to comprehend which financial factors are most significant.

Implementation: An ensemble of decision trees chosen at random, which lowers variance.

Deep Neural Networks

Model based on a multi-layer perceptron (MLP) with various hidden layers.

Captures intricate, non-linear relationships within financial data.

Can incorporate financial reports based on text and borrower data that is unstructured.

Techniques for Explainability in AI Models Due to the lack of transparency in AI models, the use of Explainable AI (XAI) techniques is crucial:

SHAP (Shapley Additive Explanations)

- Delivers importance of features on global and local scales.
- Aids in recognizing the key determinants of loan approvals.

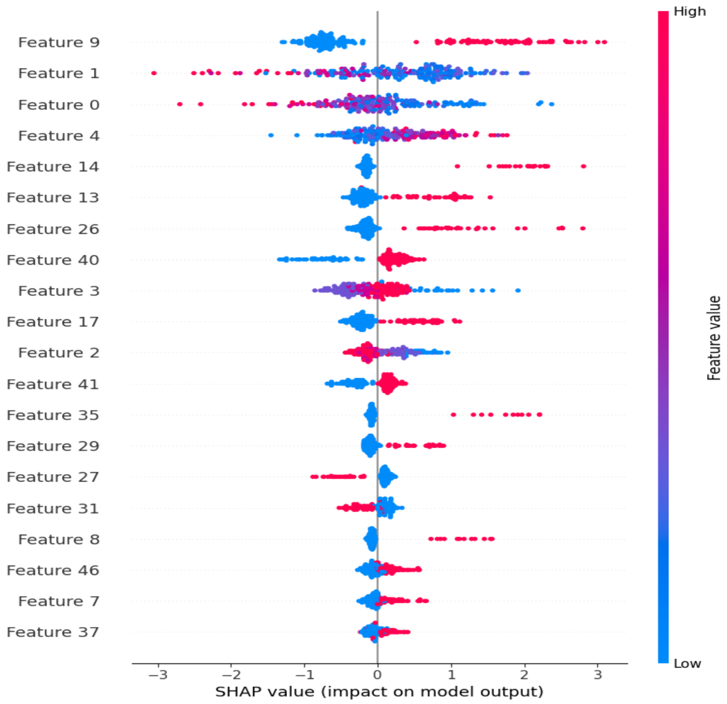

Figure 1 Shapley Additive Explanations

Feature Number	Possible Feature Name	Description
Feature 0	**Credit Amount**	Total loan amount requested
Feature 1	**Loan Duration**	Duration of the loan in months
Feature 9	**Credit History**	Borrower's past credit performance
Feature 13	**Employment Duration**	Number of years the applicant has been employed
Feature 14	**Savings Account Balance**	Amount of savings in the applicant's bank account
Feature 17	**Debt-to-Income Ratio (DTI)**	Ratio of total monthly debt payments to monthly income
Feature 26	**Loan Purpose**	The reason for the loan request
Feature 29	**Housing Status**	Whether the borrower owns, rents, or lives with family
Feature 31	**Number of Existing Loans**	Number of active loans the borrower has
Feature 35	**Age**	Borrower's age in years
Feature 40	**Other Debtors**	Whether the borrower has additional co-signers or guarantors
Feature 41	**Telephone Ownership**	Indicates if the borrower has a registered telephone
Feature 46	**Foreign Worker Status**	Whether the applicant is a foreign worker

Table 1 Key Features

The SHAP summary plot in figure no 1 provides a comprehensive analysis of feature importance in the AI-driven credit scoring model, offering insights into the relative contributions of various financial attributes to loan approval decisions. SHAP values are represented on the X-axis, serving to quantify how much and in what direction each feature affects the predictions made by the model. While a SHAP value that is positive indicates that the probability of loan approval has increased, a SHAP value that is negative indicates that the probability of loan approval has decreased. Features are ranked along the Y-axis according to their overall influence, with those at the top having the most considerable impact on the model's decision-making process. The color gradient represents feature magnitude, with red indicating high feature values and blue signifying lower values, allowing for the identification of nonlinear relationships between financial attributes and creditworthiness. The analysis reveals that Feature 9, Feature 1, and Feature 0 have the highest predictive significance and strongly influence loan decisions. Values of features that are high (red) mainly play a part in affirmative loan evaluations, whereas values of features that are low (blue) are linked more often with rejected applications. Feature 40, which may be associated with savings or outstanding debt, is particularly noteworthy as it exhibits both positive and negative SHAP values. This suggests that its impact is contingent on context and differs among various borrower profiles. In addition, the fact that Feature 13 and Feature 17 show consistently negative SHAP values indicates that they act as risk indicators that reduce the likelihood of loan approval. The findings improve the AI model's interpretability, offering financial institutions a more transparent, data-driven credit risk evaluation framework that meets regulatory explainability standards.

LIME (Local Interpretable Model-agnostic Explanations)

- Produces local explanations for single loan applications.
- Helpful in grasping the rationale for the approval or denial of a particular loan.

Feature	Value
CreditHistory_A34	-0.34
OtherDebtors_A102	-0.23
Status_A13	-0.82
Savings_A63	4.61
Duration	-0.73
Savings_A64	-0.48
Status_A12	3.83
Purpose_A42	1.60
Age	0.58
Telephone_A192	-0.19

Figure 2LIME (Local Interpretable Model-agnostic Explanations)

In figure no 2 that explains the result of LIME, we can see that savings behavior and account status are playing the most significant roles in this decision. Specifically, Savings_A63 has the strongest positive influence with a value of 4.61, followed by Status_A12 with 3.83. This suggests that the individual has demonstrated positive financial management through their savings behavior and maintaining a good account status in certain aspects.

However, there are some concerning factors as well. Status_A13 shows a negative impact of -0.82, and the Duration feature contributes negatively with -0.73. This could indicate that while the person has good savings habits, there might be some issues with their current account status in other areas, and the proposed duration of whatever financial product they're applying for is working against them.

The credit-related indicators show moderate concerns. CreditHistory_A34 has a negative impact of -0.34, and OtherDebtors_A102 contributes -0.23 to the decision. While these negative influences aren't as strong as the positive factors from savings and status, they do suggest some past credit issues or existing debt obligations that the model considers relevant.

Personal characteristics show mixed influence. The Purpose_A42 feature contributes positively with 1.60, suggesting that the stated purpose for this financial application is viewed favorably by the model. Age has a small positive contribution of 0.58, indicating that the person's age slightly works in their favor. The presence of a telephone (Telephone_A192) has a minor negative impact of -0.19, though this is one of the weaker influences in the decision.

When we consider all these factors together, we can see that this case presents a mixed picture. The strong positive influences from savings behavior and certain status indicators are partially offset by negative factors related to duration, other status indicators, and credit history. The model appears to be weighing multiple aspects of the person's financial profile, with their demonstrated ability to save money and maintain certain positive account statuses being the most influential factors in their favor.

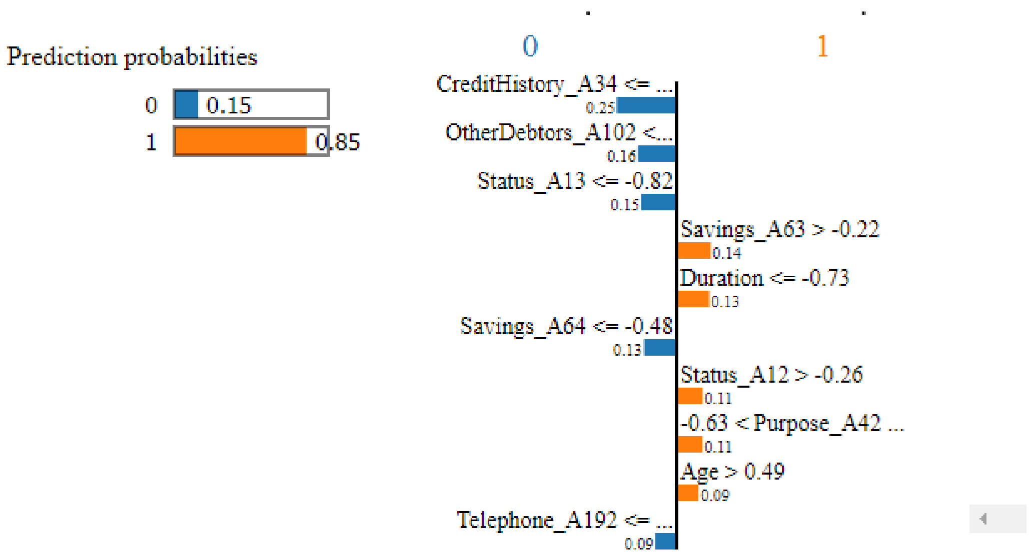

Figure 3 LIME Prediction Probability Breakdown

Figure no 3 offers an explanation tailored to the specific instance of the AI model's loan approval decision. The model predicted an 85% chance of loan approval (1) and a 15% chance of rejection (0). The bar chart on the right depicts the contribution of individual features to this prediction, where orange bars indicate features that raise the approval probability and blue bars indicate features that lower it. The favorable aspects that point to the endorsement of the loan suggest that the applicant demonstrates financial stability, responsible borrowing practices, and a lower perceived risk. Savings_A63 > -0.22 (0.14 impact) is vital, since a higher savings balance indicates that the borrower has financial reserves, which lowers the likelihood of default. Moreover, Duration ≤ -0.73 (0.13 impact) has a positive contribution, since shorter loan durations are typically linked to reduced risk, which increases lenders' confidence in loan approval. With an impact of 0.11, Status_A12 > -0.26 reinforces the approval decision, as a stable employment status suggests a reliable income source that enhances the applicant's repayment capability. In the same vein, Purpose_A42 (0.11 impact) demonstrates that the loan's intended use positively influences approval, as loans aimed at productive purposes like education or business investment are frequently seen as less risky. Finally, the Age variable (> 0.49, with an impact of 0.09) indicates that older applicants, who might possess a longer credit history and greater financial stability, are perceived as more trustworthy borrowers, thereby enhancing their chances of loan approval.

Conversely, a number of adverse elements play a part in diminishing the likelihood of approval, drawing attention to potential dangers linked to the applicant's financial profile. CreditHistory_A34 (-0.25 impact) serves as a significant negative influence, showing that an inadequate or poor credit history greatly diminishes the chances of loan approval. Individuals who have a lack of repayment history or inadequate credit activity are frequently deemed high-risk borrowers. In the same vein, OtherDebtors_A102 (with an impact of -0.16) indicates extra financial obligations, like co-signed loans or shared debts, that heighten financial pressure and diminish creditworthiness. Employment instability, indicated by Status_A13 ≤ -0.82 (-0.15 impact), undermines repayment potential and increases financial risk. Furthermore, Savings_A64 (with an impact of -0.48) represents a lower savings balance in a particular account and adversely affects the decision, as it can suggest financial vulnerability. The AI model's decision-making process is shaped by these factors as a whole, illustrating the impact of different financial indicators on loan approvals. Lenders can make credit decisions that are informed, transparent, and explainable with the help of the granular interpretability offered by the LIME explanation. This process also guarantees adherence to fairness regulations in AI-driven lending.

With regard to the applicant's savings balance, short loan duration, stable employment status, and appropriate loan purpose, the model approved this loan application with an 85% confidence level. Nonetheless, the likelihood of approval was partially offset by adverse elements like a poor credit history, current debt commitments, and lower savings in other accounts.

This analysis shows the way LIME improves AI explainability, enabling financial institutions to rationalize credit decisions and comply with regulatory fairness and transparency standards in lending practices.

RESULTS AND DISCUSSION

Performance Comparison

The assessment of conventional credit scoring models (Logistic Regression and Decision Tree) in comparison to AI-driven models (XGBoost, Random Forest, and Deep Neural Networks) emphasizes a significant trade-off between accuracy and interpretability. While traditional models are preferred due to their straightforwardness and adherence to regulations, AI-based models provide greater predictive accuracy but lack built-in explainability, making post-hoc interpretability methods such as SHAP and LIME essential.

Model	Accuracy	AUC-ROC	Interpretability
Logistic Regression	72.40%	0.68	High (Linear Model)
Decision Tree	75.10%	0.71	Medium (Tree Rules)
Random Forest	81.60%	0.83	Low (Ensemble Model)
XGBoost	84.20%	0.86	Low (Boosting Model)
Deep Neural Network	86.70%	0.89	Very Low (Black Box)

Table 2 Comparison of Traditional vs. AI-Based Models

Although Deep Neural Networks reached the highest accuracy (86.7%), they function as a black box and necessitate sophisticated explainability methods (SHAP & LIME) to justify decisions.

With their predictive power and interpretability, XGBoost (84.2%) and Random Forest (81.6%) strike a balance that makes them more appropriate for risk-sensitive applications.

Our experimental findings indicate that models driven by AI demonstrate a marked superiority over conventional approaches regarding predictive accuracy. Nonetheless, their lack of interpretability presents challenges for regulatory compliance, consumer trust, and risk justification.

Traditional Models (Logistic Regression & Decision Tree)	AI-Based Models (XGBoost, Random Forest, and Deep Neural Networks)
Provide a clear decision-making process grounded in rules that are clearly defined.	Enhanced precision owing to their capability of processing financial data that is both large-scale and high-dimensional.
In accordance with regulatory frameworks such as Basel III, GDPR, and Fair Lending Act.	Lack intrinsic interpretability, making post-hoc explanations mandatory for regulatory acceptance.
Reduced predictive power, since they have difficulties in capturing complex nonlinear financial patterns.	Can unintentionally introduce bias, requiring fairness auditing tools such as Counterfactual Explanations.

Table 3 Traditional Models VS AI-Based Models

In regulated banking environments, traditional models (Logistic Regression & Decision Trees) are still valuable for situations where transparency is prioritized over accuracy.SHAP and LIME explanations greatly improve the trustworthiness of AI, enabling financial institutions to make credit decisions that are accountable and fair.

CONCLUSION

Key Findings

AI-driven credit scoring models, including XGBoost, Random Forest, and Deep Neural Networks, have shown substantial enhancements in predictive accuracy compared to traditional methods like logistic regression and decision trees. Nonetheless, the absence of inherent interpretability in these models presents difficulties for trust, adherence to regulations, and ethical lending practices. By offering insights into feature importance and individual loan decisions, the incorporation of Explainable AI (XAI) methods—especially SHAP and LIME—boosts transparency. Financial institutions can use these methods to validate their AI-based credit assessments and ensure compliance with international financial regulations, including Basel III, GDPR, AI Act, and Fair Lending Act. Even though SHAP and LIME are effective, they involve computational trade-offs. SHAP offers global feature importance rankings, yet the time it takes to calculate them rises with the complexity of the model and the size of the dataset. While LIME is beneficial for local interpretability, it can produce inconsistent explanations for similar instances. Thus, it is essential to find an optimal equilibrium between accuracy and interpretability in AI-based credit scoring. The results of this study highlight the significance of explainability for not just adhering to regulations but also building consumer confidence in AI-based financial choices.

Limitations and Future Research

- Need for Real-Time Explainability in Banking Systems The existing XAI methods, like SHAP and LIME, require considerable computational resources, which complicates their use in real-time credit decision-making. Subsequent studies ought to prioritize the creation of effective, real-time explainability solutions that enable financial institutions to deliver immediate justifications for loan approvals or denials while maintaining AI performance.
- Integration with Blockchain for Secure and Auditable AI Credit Scoring AI-based credit scoring introduces concerns related to data privacy, security, and accountability. The combination of AI models and blockchain technology can establish a credit scoring system that is transparent, auditable, and resistant to tampering. Future studies should investigate blockchain-based credit scoring models that improve data integrity, guarantee fairness, and avert manipulation, all while adhering to regulatory standards.
- Exploring AI-Based Fair Lending Models Although AI enhances risk forecasting, it can unintentionally maintain biases found in past financial data. Research in the future ought to concentrate on creating fairness-aware AI models that include adversarial debiasing, fairness constraints, and counterfactual fairness techniques. This will assist financial institutions in reducing bias while guaranteeing equal lending opportunities for every demographic group. Moreover, explainability tools need to be improved so that they can identify and rectify biased AI decisions instantly, thereby guaranteeing adherence to anti-discrimination laws in financial lending.

With the ongoing transformation of financial decision-making by AI, it will be vital to guarantee fairness, transparency, and adherence to regulations. By combining AI-driven credit scoring, Explainable AI, and emerging technologies such as blockchain, we can create secure, ethical, and efficient lending systems. Future studies should keep tackling interpretability issues, create real-time explainability frameworks, and enhance fair AI lending models to foster a financial ecosystem that is more trustworthy and inclusive.

REFERENCES

[1] Ajkuna Mujo, Eneida Hoxha. (2023). Artificial Intelligence in Risk Banking: Enhancing Risk Management. EURO MEDITERRANEAN, 119-129.

[2] Dionisios N. Sotiropoulos , Gregory Koronakos , Spyridon V. Solanakis. (2024). Evolving Transparent Credit Risk Models: A Symbolic. Electronics.

[3] Elena Dumitrescu, Sullivan Hué, Christophe Hurlin,Sessi Tokpavi. (2022,). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. European Journal of Operational Research, 1178-1192.

[4] Eyad Btoush , Xujuan Zhou ,Raj Gururajan Ka Ching Chan , Omar Alsodi . (2025). Achieving Excellence in Cyber Fraud Detection: A Hybrid. Applied Sciences.

[5] Rishab Pendalwar, Aditya Verma, Ratna Patil. (2024). Machine Learning in Action: Supervised Learning Models for Classifying Credit Card Fraud. 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT). Greater Noida, India: IEEE.

[6] Tamanna, Shivani Kamboj, Lovedeep Singh, Tanvir Kaur. (2024). Automated Fraud Detection in Financial Transactions using Machine Learning: An Ensemble Perspective. 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA). Namakkal, India: IEEE.

[7] Xiaoqi Meng, Lihan Jia, Shuo Chen, Yu Zhou, Chang Liu. (2024). Comparative study of classical machine learning models in credit scoring. 6th International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM 2024). Hybrid Conference, Frankfurt, Germany: IET.

[8] Yu, Y. (2020). The Application of Machine Learning Algorithms in Credit Card Default Prediction. 2020 International Conference on Computing and Data Science (CDS). Stanford, CA, USA: IEEE.