International Journal of Communication Networks and Information Security
2024,16(4)
ISSN: 2073-607X,2076-0930
https://https://ijcnis.org/

Research Article

XAI - Credit Risk Analysis

Nilesh Patil [1*], Sridhar Iyer [1*], Chaitya Lakhani [2], Param Shah [2], Ansh Bhatt [2],

Harsh Patel [2], Dev Patel [2]

[1] Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

[2] Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

***Corresponding Author:** nilesh.p@djsce.ac.in,

ARTICLE INFO

Received: 10 Aug 2024
Accepted: 14 Sep 2024

ABSTRACT

This paper delves into the integration of Explainable AI (XAI) techniques with machine learning models for credit risk classification, addressing the critical issue of model transparency in financial services. We experimented with various models, including Logistic Regression, Random Forest, XGBoost, LightGBM, and Artificial Neural Networks (ANN), on real-world credit datasets to predict borrower risk levels. Our results show that while ANN achieved the highest accuracy at 95.3%, Random Forest followed closely with 95.23%. Logistic Regression also performed strongly with an accuracy of 94.68%, while XGBoost and LightGBM delivered slightly lower accuracies of 94.4% and 94.37%, respectively. However, the superior accuracy of these complex models, particularly ANN, comes with a trade-off: reduced transparency, making it difficult for stakeholders to understand the decision-making process. To address this, we applied XAI techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) to provide clear and understandable explanations for the predictions made by these models. This integration not only enhanced model interpretability but also built trust among stakeholders and ensured compliance with regulatory standards. This study illustrates how XAI serves as an effective mediator between the precision of sophisticated machine learning algorithms and the demand for clarity in evaluating credit risk. XAI offers a well-balanced method for managing risk in finance, harmonizing the need for both accuracy and interpretability.

Keywords: Explainable artificial intelligence (XAI), Interpretable machine learning, Credit Risk, Financial Technology (FinTech)

INTRODUCTION

In the fast-evolving landscape of data-driven decision-making, the use of machine learning models across various fields is on the rise. These ”black-box” models are known for their strong predictive capabilities, yet they often lack clear explanations for their predictions. Although they can deliver impressive accuracy, their lack of transparency creates challenges, particularly in sectors where understanding the reasoning behind a model’s predictions is crucial for building trust, maintaining governance, and adhering to regulatory standards.

This research paper examines how black-box models can be interpreted using various X-AI techniques across different applications, particularly in the area of credit risk assessment. It categorizes credit risk and determines eligibility for loan approval. The dataset includes a wide range of factors, such as demographic information, financial metrics, and credit-related variables, all of which play a role in the decision-making process. In this study, several machine learning models are employed to classify credit risk: Logistic Regression, XG-Boost, Support Vector Machine (SVM), Random Forest, and Neural Networks. Following the classification, Explainable AI techniques like LIME and SHAP are applied to these models. This approach allows for an understanding of how each feature contributes to the final decisions made by the models, highlighting both the positive and negative impacts of the most significant features.

The aim of this paper is to show how important Explainable AI is in improving transparency and interpretability in machine learning applications. It applies LIME and SHAP to multiple black-box models and identifies the most important features driving predictions, thus providing valuable insights for analysts and decision-makers in various domains. More specifically, it tries to close the gap between model accuracy and interpretability so that machine learning models can be effective as well as understandable in high-stakes applications where transparency is vital.

LITERATURE REVIEW

[1] examines how machine learning can enhance credit risk analysis by improving both prediction accuracy and model explainability. It identifies Gradient Boosting as more effective than traditional methods like logistic regression, especially in handling imbalanced financial data. Key findings show younger males and inconsistent payments as strong indicators of default risk. The research emphasizes the need for explainable AI to build trust and promote wider adoption in the finance industry.

[2] explores the expanding role of machine learning in finance, highlighting its ability to address challenges that traditional econometric methods struggle with. The authors categorize ML applications into three areas: creating new financial measures, reducing prediction errors, and expanding econometric tools. They emphasize ML's potential to improve financial research through more accurate predictions and novel 429 Nilesh Patil /IJCNIS,16(4),428-442 insights, and they foresee promising applications in areas like corporate finance and behavioral finance, driving a shift toward data-driven research.

[3] examines the role of explainable AI (XAI) in credit risk management, focusing on how techniques like LIME and SHAP improve transparency in machine learning models. Using data from the Lending Club platform, the authors find that while these methods provide useful explanations, challenges like SHAP's computational demands and LIME's model limitations persist. The research highlights the need for transparency to build trust and drive broader AI adoption in finance, advocating for further development of XAI to overcome these barriers and enhance credit risk management.

[4] analyzes 60 studies on Explainable AI (XAI) in finance, highlighting its role in promoting transparency and regulatory compliance in AI-driven decisions. The authors find strong research in areas like risk management and portfolio optimization but note a gap in anti-money laundering applications. They argue that while XAI methods are gaining traction, research is fragmented and lacks a cohesive view of XAI's overall impact. The review offers a comprehensive overview, identifies research gaps, and suggests future directions to advance XAI in finance.

[5] examines various machine learning classifiers for credit risk analysis, comparing models like Decision Trees, SVM, Naive Bayes, KNN, and ANN. It emphasizes the importance of selecting suitable classifiers to enhance prediction accuracy and highlights how different models manage data complexity. While SVM and ANN capture complex patterns, simpler models like Naive Bayes and Decision Trees offer better transparency. The authors stress balancing predictive performance and interpretability in credit risk models, contributing to the ongoing research on optimizing machine learning in finance.

[6] explores the use of Artificial Neural Networks (ANN) for evaluating credit risk in commercial banks, focusing on accurately predicting borrowers' ability to repay loans. Using 14 financial indicators and techniques like cluster and factor analysis, the authors develop a credit rating system for their models. Comparing traditional methods with three ANN models, they find that ANN provides superior prediction accuracy. The research underscores the potential of AI-driven techniques in improving credit risk management in banking.

[7] examines the use of explainable machine learning in credit risk management, addressing the need for transparency in financial decision-making. While ML models improve credit risk predictions, their "black-box" nature limits adoption in regulated environments. The research explores explainable methods that enhance both model performance and interpretability, making them understandable for regulators and financial professionals. By integrating these techniques, institutions can maintain predictive accuracy while meeting regulatory standards, offering a practical framework for balancing transparency and performance in credit risk models.

[8] proposes an Explainable AI (XAI) framework for credit evaluation, addressing the transparency issues in loan approval decisions, especially heightened during the COVID-19 pandemic. The authors introduce a Random Forest-based model that predicts loan outcomes with high accuracy and provides clear explanations using techniques like LIME and SHAP. The model offers both local and global interpretability, 430 Nilesh Patil /IJCNIS,16(4),428-442 helping banks justify credit decisions while ensuring compliance. With near 99% performance metrics, this research aligns with Industry 5.0 trends, contributing to more transparent and

trustworthy AI-driven financial decisions.

[9] examines the use of Explainable AI (XAI) in predicting credit risk, addressing the need for greater transparency in financial decisions. As machine learning models become central to credit assessments, concerns about their opacity and fairness have grown. The authors integrate XAI techniques like SHAP and LIME to make these models both accurate and interpretable, allowing financial institutions to stay compliant and build customer trust. The study shows that incorporating XAI ensures clear, understandable decisions while maintaining high prediction accuracy, bridging the gap between complex algorithms and practical financial applications.

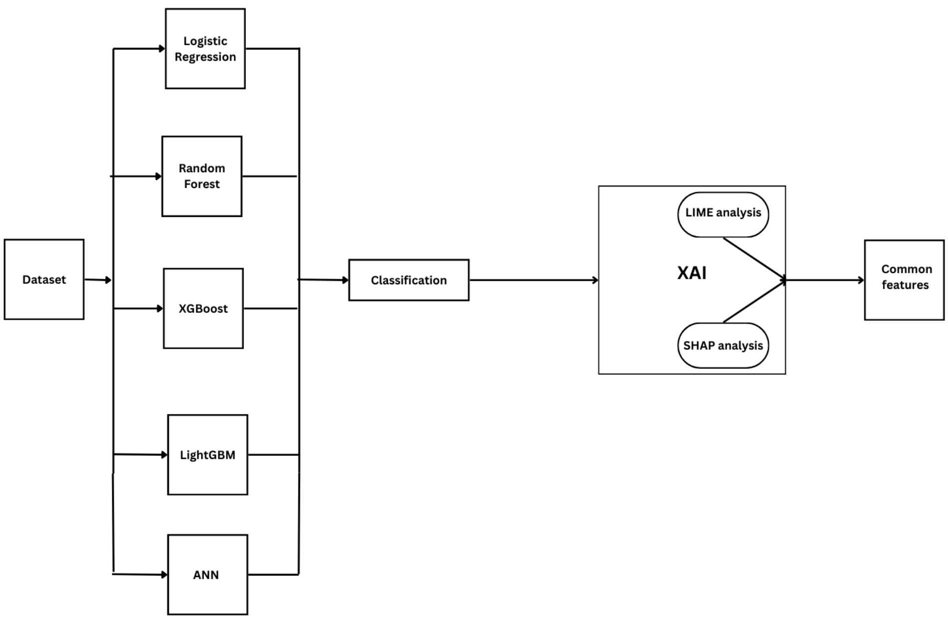
[10] examines how integrating Explainable AI (XAI) techniques with various machine learning algorithms can enhance loan prediction accuracy. By comparing models like decision trees, support vector machines, and ensemble methods, the authors find that XAI improves both the transparency and interpretability of predictions. This integration not only boosts decision-making accuracy but also helps users understand and trust the factors influencing loan approvals.

[11] explores how Explainable AI (XAI) can improve transparency in credit assessment for banks. Traditional models frequently function as black boxes, which complicates the understanding of their decision-making processes. By incorporating XAI techniques, this study enhances the interpretability of machine learning models used for credit scoring, providing clear insights into the factors influencing decisions. This method enables banks to maintain high accuracy while promoting fairness and adhering to regulatory standards, effectively connecting complex algorithms with transparent and trustworthy financial decisions.

[12] introduces a model for predicting loan approvals using Explainable AI (XAI) techniques to make machine learning decisions more understandable. By combining traditional algorithms like Random Forest and Gradient Boosting with XAI methods such as SHAP and LIME, the model offers clear insights into loan approval factors. The research highlights that while accuracy is important, integrating XAI ensures transparency for both financial institutions and applicants, balancing robust performance with ethical decision-making.

[13] explores credit risk analysis using machine learning algorithms, focusing on how methods like Random Forests can improve accuracy and efficiency in risk management. Presented at the 29th Signal Processing and Communications Applications Conference, this study reveals that these advanced techniques surpass traditional methods in predicting default risk. By showcasing the practical applications of these models using Python, the research emphasizes their potential to improve credit risk evaluation and advance financial risk assessment tools.

METHODOLOGY



Different models

Logistic regression for binary classification

The logistic regression technique estimates the likelihood of an input being classified into a particular group or category. The formula is given below:

$$P(y = 1 \mid A) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 A_1 + \beta_2 A_2 + \cdots + \beta_n A_n)}} \quad (1)$$

Where:

• P(y = 1 | A) = probability of the target variable y being 1 given the input features A = (A1, A2, . . . , An). • β0 = intercept term.

• β1, β2, . . . , βn are coefficients for input features A1, A2, . . . , An.

• e = base of the natural logarithm.

This formula essentially transforms the relationship between the features and the outcome into a format for binary classification, producing results between 0 and 1.

To evaluate the logistic regression model, the data is split into two sets: one for training the model (80%) and the other for testing its accuracy (the remaining 20%). A fixed random state is applied to ensure that the results are consistent each time the model is executed. The model's parameters are set with penalty='l2', C=1, solver='lbfgs', and max-iter=100. To prevent data type issues, the feature names in both the training and testing datasets are converted to strings. Furthermore, any categorical features are transformed into numeric values before training the model, which aids the algorithm in accurately interpreting these variables.

Random Forest Classifier

The Random Forest classifier is an ensemble learning technique that merges several decision trees to achieve more reliable and accurate predictions. In the context of random forests, each tree is trained on a random subset of the data. The final prediction is then determined through the majority vote or average of the individual trees. This technique helps reduce overfitting and improves generalization, making Random Forest a dependable model for various classification tasks. In credit risk analysis, the Random Forest technique is highly beneficial due to its ability to handle large datasets with intricate patterns and interactions. It effectively captures non-linear relationships between borrower characteristics and default risk and provides feature importance scores that are instrumental in identifying critical risk factors. Furthermore, its resilience against overfitting renders it especially well-suited for financial settings where data may be noisy or imbalanced, consequently ensuring accurate and consistent credit risk predictions.

$$\mathrm{Gini}(G) = 1 - \sum_{i=1}^{N} x_i^2 \quad (2)$$

• G = dataset at the node.

• N = number of classes.

• xi = probability of class i in dataset D.

$$\hat{y} = \mathrm{mode}\{h_1(k), h_2(k), \ldots, h_m(k)\} \quad (3)$$

• y^ = predicted class.

• hi(k) = prediction from the i-th decision tree for input x.

• m = number of trees in the forest.

XGBoost(Extreme Gradient Boosting)

It is a powerful version of gradient boosting, a machine learning technique where models are built one after another, each improving on the mistakes of the last. XGBoost stands out by adding smart features like regularization, parallel processing, and the ability to handle missing data. It’s fast and efficient, especially with large datasets, which is why it's a go-to tool for many data scientists in competitive settings.

$$F(x) = \sum_{m=1}^{M} \gamma \times f_m(x_i) \tag{4}$$

LightGBM(Light Gradient Boosting Machine)

LightGBM is a popular framework for gradient boosting that is used for efficiently handling large datasets and complex, high-dimensional data. Unlike traditional methods that build trees level by level, LightGBM uses a leaf-wise strategy. This approach enables it to concentrate on the most significant splits, resulting in faster training and enhanced accuracy. One of its strengths is the ability to process categorical features directly, without requiring one-hot encoding, and it also deals with missing data effectively. Additionally, LightGBM's use of histogram-based decision tree learning makes it faster and more memory-efficient. This algorithm functions by employing gradient descent to decrease the discrepancy between forecasted and true values. Its goal is to minimize errors and enhance overall effectiveness.

$$L = \sum_{i=1}^{n} \mathrm{Loss}(y_i, \hat{y}_i) + \lambda \sum_{j=1}^{p} w_j^2 \tag{5}$$

Where:

- yi and ˆyi represent the actual and predicted values, respectively.
- λ = regularization parameter.
- wj = weights of the model.

Artificial Neural Networks (ANN) for Classification

Artificial Neural Networks (ANNs) are inspired by the human brain and are composed of interconnected neurons organized in layers. The architecture includes an input layer, multiple hidden layers with ReLU activation:

$$\mathrm{ReLU}(y) = \max(0, y) \tag{6}$$

and an output layer with a sigmoid function:

$$\sigma(y) = \frac{1}{1 + e^{-y}} \tag{7}$$

Training is performed using backpropagation with weight updates defined as:

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial \mathcal{L}}{\partial w_{ij}} \tag{8}$$

The model employs the Adam optimizer, sets a learning rate of 0.001, and trains for 200 epochs while incorporating early stopping. Input data is standardized, and categorical features are one-hot encoded. This setup effectively captures non-linear relationships for binary classification tasks.

EXPERIMENTATION AND RESULTS

Data Preprocessing head

The dataset utilized in this study [14] required several preprocessing steps. To deal with missing values, Columns with more than 35% missing values were dropped. For the remaining columns with missing values, different strategies were used:

Imputation with Mode: For categorical columns like ’OCCUPATION-TYPE’, ’NAME-TYPE-SUITE’. Imputation with Median For numerical columns. Imputation with Zero: For columns related to credit bureau inquiries. Handling Clerical Errors: In ’CODE-GENDER’, ’XNA’ values were replaced with ’F’ (Female). In ’ORGANIZATION-TYPE’, ’XNA’ values were replaced with NaN. Outlier Handling: Outliers in ’AMT-INCOME-TOTAL’ and ’AMT-ANNUITY’ were removed based on visual inspection of box plots.

Label Encoding: Categorical columns in a subset of the data were label encoded to convert them into numerical values. Principal Component Analysis: PCA reduces the number of features by transforming the original variables into a smaller set of uncorrelated components, which retain most of the variance in the data. Additionally, it also helps in addressing multicollinearity by combining correlated features into principal components, reducing redundancy in the dataset.

XAI Results and Inference

The study employed LIME and SHAP techniques on the dataset to enhance understanding of feature significance. These methods revealed which attributes were crucial in the model’s decision-making process for classifying entries as either ”LOAN” or ”NO LOAN”. By identifying key features, researchers gained insight into the primary factors influencing loan approval determinations.

Table 1: Features Selected from [14]

Features	Description	Data Type
TARGET	shows whether the applicant is qualified for loan approval (1 = not qualified for loan, 0 = qualified for loan).	Numeric
NAME-CONTRACT-TYPE	The nature of the loan deal.	Categorical
CODE-GENDER	The client’s gender	Categorical
FLAG-OWN-CAR	States if the client has a car (1 = Yes, 0 = No)	Categorical
FLAG-OWN-REALTY	Denotes whether the client is a property owner (1 = Yes, 0 = No).	Categorical
CNT-CHILDREN	The client’s total number of children.	Numeric
AMT-INCOME-TOTAL	The client’s overall income.	Numeric
AMT-CREDIT	The client’s requested credit amount.	Numeric
AMT-ANNUITY	The loan’s annuity amount, or the amount of periodical payments.	Numeric
AMT-GOODS-PRICE	The cost of the items that are being lent money for	Numeric
NAME-INCOME-TYPE	The nature of the client’s revenue source.	Categorical
NAME-EDUCATION-TYPE	The client’s education level.	Categorical
NAME-FAMILY-STATUS	The client’s family situation.	Categorical
NAME-HOUSING-TYPE	The kind of dwelling the customer resides in.	Categorical
OCCUPATION-TYPE	The client’s line of work.	Categorical
REGION-RATING-CLIENT-W-CITY	Based on the degree of development, a rating of the client’s city and region.	Numeric

ORGANIZATION-TYPE	The kind of the company the client works for.	Categorical
OBS-60-CNT-SOCIAL-CIRCLE	The count of the client's social environments that were observed in the last sixty days.	Numeric
DEF-60-CNT-SOCIAL-CIRCLE	The count of defaulters inside the client's social circle that was noticed in the last sixty days.	Numeric
AMT-REQ-CREDIT-BUREAU-QRT	The quantity of inquiries about the client submitted to the credit bureau during the most recent quarter.	Numeric

Table 2: Model Accuracy Comparison

Model Used	**Accuracy**
Logistic Regression	94.68%
Random Forest	95.23%
XGBoost	94.4%
LightGBM	94.37%
ANN	95.3%

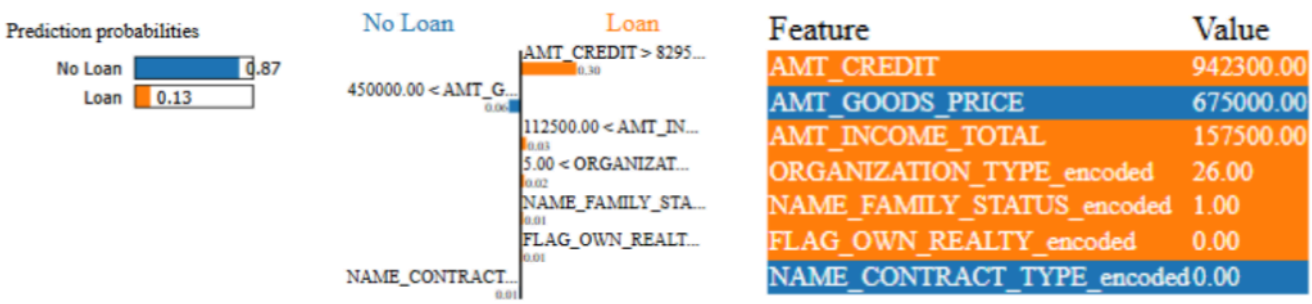


Fig 2: LIME graph for Logistic Regression

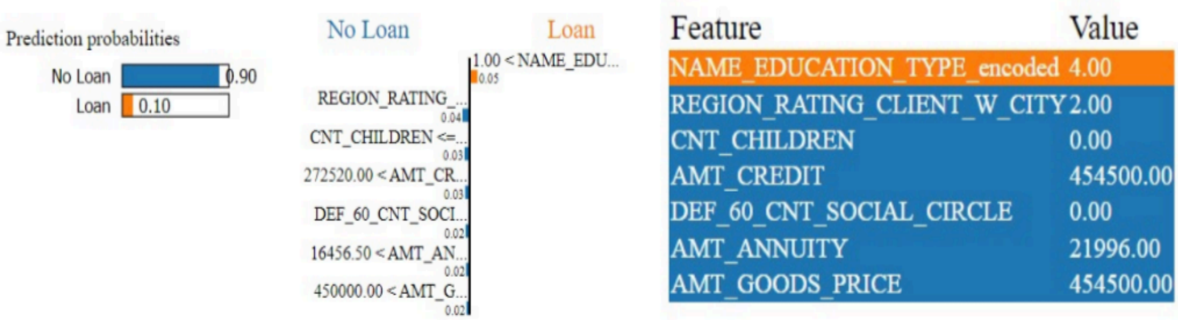


Fig 3: LIME graph for Random Forest

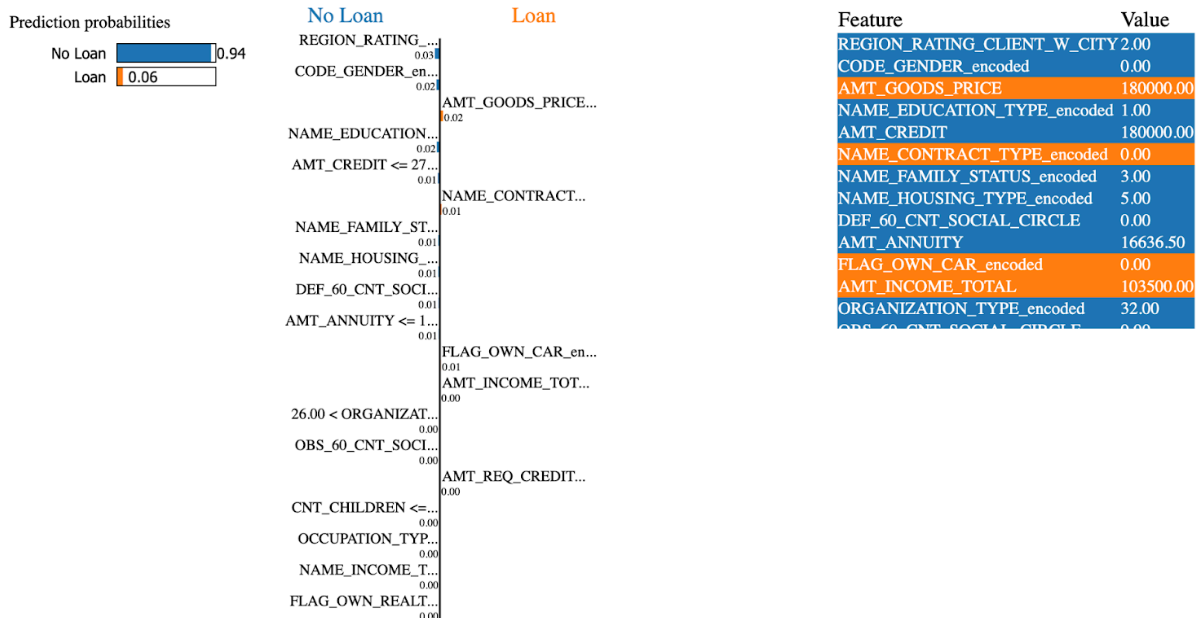


Fig 4: LIME graph for XGBoost

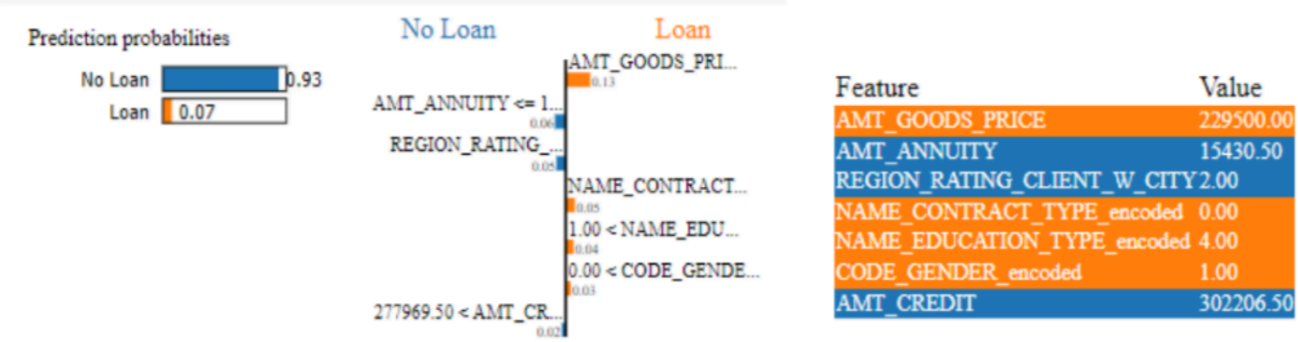


Fig 5: LIME graph for LightGBM

Fig 6: LIME graph for ANN

Figures [2],[3],[4],[5] and [6] show the LIME graphs for the data point using Logistic Regression, Random Forest, XGBoost, LightGBM and ANN models that belongs to the NO LOAN category. Features like AMT-GOODS-PRICE, REGIONRATING-CLIENT-W-CITY, CODE-GENDER, AMT-CREDIT, AMT-ANNUITY, AMT-INCOME-TOTAL play an important role to classify that datapoint as NO LOAN.

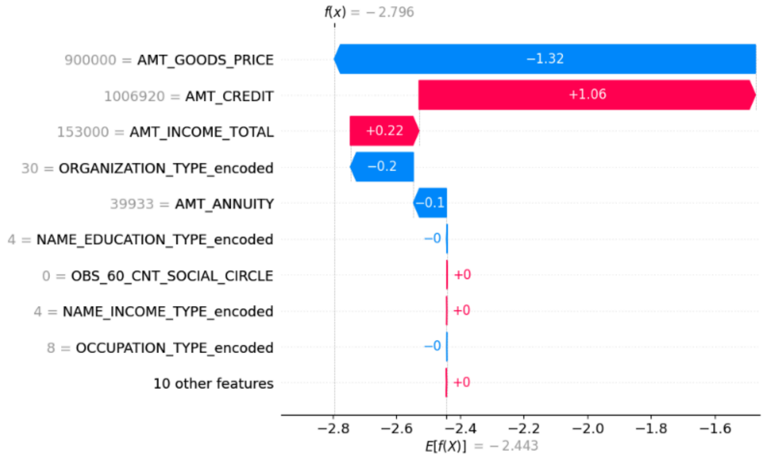


Fig 7: Waterfall plot for SHAP values graph for Logistic Regression

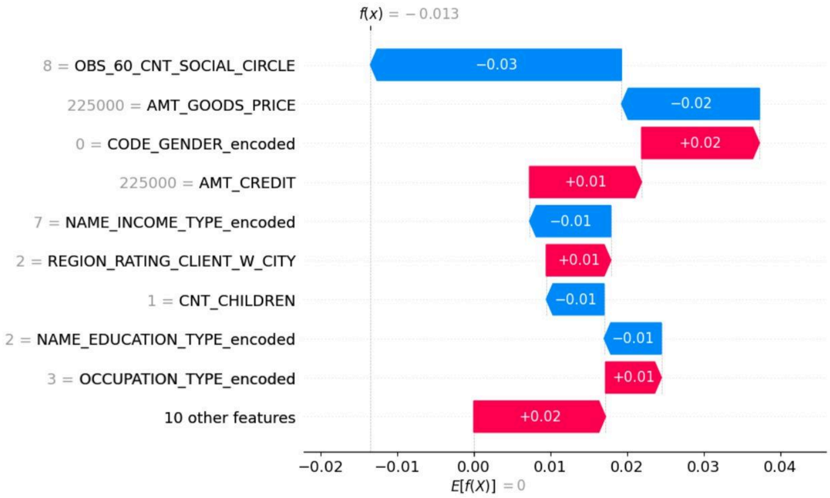


Fig 8: Waterfall plot for SHAP values graph for Random Forest

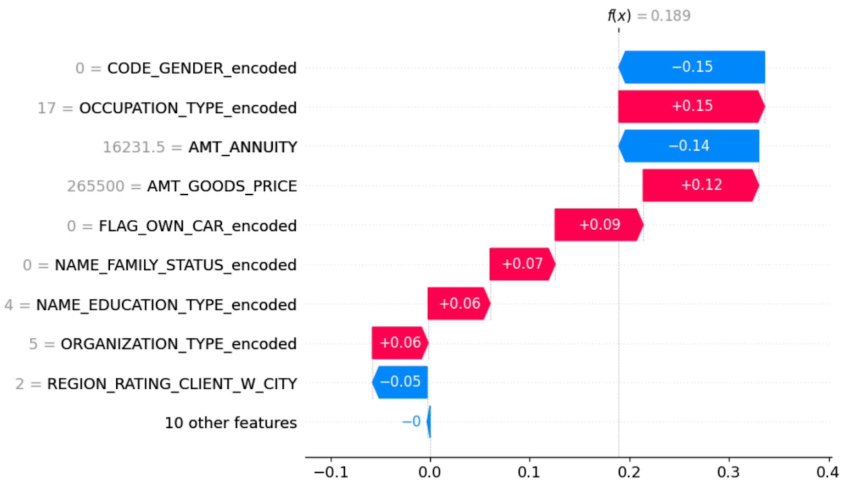


Fig 9: Waterfall plot for SHAP values graph for XGBoost

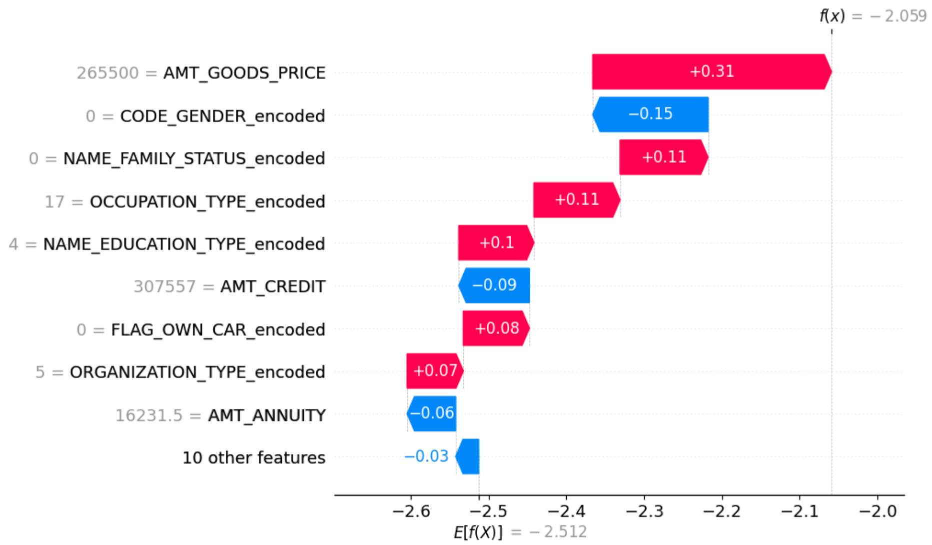


Fig 10: Waterfall plot for SHAP values graph for LightGBM

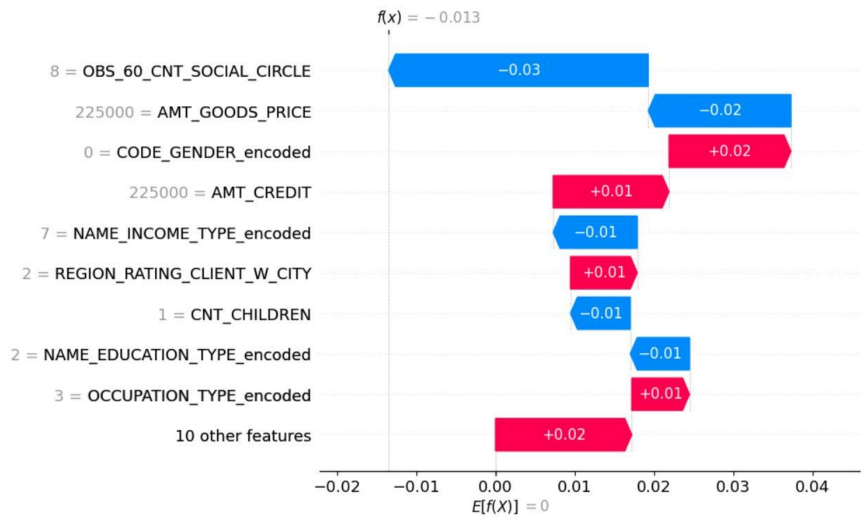


Fig 11: Waterfall plot for SHAP values graph for ANN

Figures [7],[8],[9],[10] and [11] show the waterfall plots for SHAP values generated by Logistic Regression, Random forest, XGBoost, LightGBM and ANN model respectively.

The value of the items being purchased with the loan is indicated by the AMTGOODS-PRICE. This is a variable that lenders look at to make sure the loan amount and the items' value match. A loan denial may occur if the loan amount is greater than a fair portion of the items' value, which could indicate increased risk.

The client's gender is indicated by CODE-GENDER. Gender may be used in combination with other risk assessment variables, but it shouldn't be the only aspect taken into account when approving a loan. The lender's decision-making process may occasionally be influenced by statistical differences in default rates between genders, but ethical lending practices demand that this usage be fair.

NAME-EDUCATION-TYPE reflects the client's education level, which can be an indicator of their earning potential and financial stability. Lenders may consider higher education levels as a sign of lower risk, as clients with more education might have better job prospects, increasing the likelihood of timely loan repayment.

REGION-RATING-CLIENT-W-CITY measures the development level of the client's region and city. Lenders may use this rating to assess the economic stability of the client's location. Clients in highly developed regions might be viewed as lower risk, potentially increasing their chances of loan approval due to better local economic conditions.

FLAG-OWN-CAR indicates whether the client owns a car, which can be seen as a sign of financial stability and asset ownership. Lenders may consider car ownership as a positive factor, suggesting the client has a reliable means of transportation and potentially more disposable income, which could increase the probability of loan approval.

CONCLUSION

Based on the comparison of model performance, the Random Forest classifier demonstrated the highest accuracy at 95.23%, followed closely by Model E at 95.3%, while XGBoost, LightGBM, and Logistic Regression achieved accuracies of 94.4%, 94.37%, and 94.68%, respectively. These results indicate that ensemble methods like Random Forest and XGBoost provide superior predictive power in credit risk analysis. To enhance model interpretability, LIME and SHAP were implemented to determine which features significantly influenced the classification decisions, especially in distinguishing between LOAN and NO LOAN categories. The feature importance analysis revealed that variables such as AMT-GOODS-PRICE, REGION-RATINGCLIENT-W-CITY,CODE-GENDER, AMT-CREDIT, AMT-ANNUITY, and AMTINCOME-TOTAL are critical in classifying data points, particularly for the NO LOAN category. Visual representations through LIME graphs for models like Logistic Regression, Random Forest, XGBoost, and LightGBM further validate the contribution of these features. Overall, combining high-performing models with explainability techniques offers a balanced approach to both accuracy and transparency in credit risk decision making.

FUTURE SCOPE

Future research could investigate the incorporation of additional XAI methods beyond LIME and SHAP, concentrating on innovative techniques that improve interpretability while providing strong explanations. Moreover, enhancing the accuracy of credit risk models may involve refining data preprocessing methods, including more effective management of imbalanced datasets, feature engineering, and outlier detection. Another exciting avenue is the development of hybrid models that merge traditional statistical methods with machine learning approaches, capitalizing on the advantages of both while ensuring interpretability through XAI. Broadening the use of XAI to various financial products, such as mortgages, business loans, and personal loans, could further demonstrate the effectiveness of explainable models in different credit risk contexts.

REFERENCES

[1] Chang, V., Xu, Q.A., Akinloye, S.H., et al.: Prediction of bank credit worthiness through credit risk analysis: An explainable machine learning study. Annals of Operations Research (2024) https://doi.org/10.1007/s10479-024-06134-x

[2] Hoang, D., Wiegratz, K.: Machine learning methods in finance: Recent applica- tions and prospects. European Financial Management (2022) https://doi.org/10. 1111/eufm.12408 . First published: 17 December 2022

[3] Misheva, B.H., Hirsa, A., Osterrieder, J., Kulkarni, O., Lin, S.F.: Explainable ai in credit risk management (2021) arXiv:2103.01511 [cs.LG]. A preprin.

[4] Weber, P., Carl, K.V., Hinz, O.: Applications of explainable artificial intelligence in finance—a systematic review of finance, information systems, and computer science literature. Management Review Quarterly 74, 867–907 (2024) https:// doi.org/10.1007/s11301-023-00320-0

[5] Pandey, T.N., Jagadev, A.K., Mohapatra, S.K., Dehuri, S.: Credit risk analysis using machine learning classifiers. In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 1850–1854 (2017). https://doi.org/10.1109/ICECDS.2017.8389769

[6] Hu, Y., Su, J.: Research on credit risk evaluation of commercial banks based on artificial neural network model. Procedia Computer Science 199, 1168–1176 (2022) https://doi.org/10.1016/j.procs.2022.01.148 . The 8th International Con- ference on Information Technology and Quantitative Management (ITQM 2020 2021): Developing Global Digital Economy after COVID-19

[7] Bussmann, N., Giudici, P., Marinelli, D., et al.: Explainable machine learning in credit risk management. Computational Economics 57, 203–216 (2021) https://doi.org/10.1007/s10614-020-10042-0 . Accepted 17 August 2020, Published 25 September 2020, Issue Date January 2021

[8] Nallakaruppan, M.K., Balusamy, B., Shri, M.L., Malathi, V., Bhattacharyya, S.: An explainable ai

framework for credit evaluation and analysis. Applied Soft Computing 153, 111307 (2024) https://doi.org/10.1016/j.asoc.2024.111307

[9]Shiam, S., Hasan, M., Pantho, M., Shochona, S., Nayeem, M.B., Choudhury, M.T.H., Nguyen, T.: Credit risk prediction using explainable ai. Journal of Busi- ness and Management Studies 6, 61–66 (2024) https://doi.org/10.32996/jbms. 2024.6.2.6

[10] Meenakshi, B., Jhansi, A., Sri, M.R., Chandra, K.R.: Enhancing loan predic- tion accuracy: A comparative analysis of machine learning algorithms with xai integration. International Journal of Scientific Research in Engineering and Man- agement (IJSREM) 8(5), 1–7 (2024) https://doi.org/10.55041/IJSREM33859 . SJIF Rating: 8.44

[11] Lange, P.E., Melsom, B., Vennerød, C.B., Westgaard, S.: Explainable ai for credit assessment in banks. Journal of Risk and Financial Management 15(12) (2022) https://doi.org/10.3390/jrfm15120556

[12] Ayad, O.M., Hegazy, A.-E.F., Dahroug, A.: A proposed model for loan approval prediction using xai. International Journal of Scientific Research in Engineer- ing and Management (IJSREM) 8(5), 1–7 (2024) https://doi.org/10.55041/ IJSREM33859 . SJIF Rating: 8.448

[13] Alag¨oz, G., C¸ anako˘glu, E.: Credit risk analysis using machine-learning algorithms. In: 2021 29th Signal Processing and Communications Applications Conference (SIU), pp. 1–4 (2021). https://doi.org/10.1109/SIU53274.2021.9477873

[14] Dutta, G.: Risk Analytics in Banking & Financial Services. Accessed: April 10, 2024 (2024). https://www.kaggle.com/code/gauravduttakiit/ risk-analytics-in-banking-financial-services-1