E-ISSN: 2708-4523
P-ISSN: 2708-4515
Impact Factor (RJIF): 5.61
AJMC 2025; SP-6(2): 285-290

www.allcommercejournal.com
Received: 24-06-2025
Accepted: 25-07-2025

Bharghav Madhiraju
Instructor M.S Data Science EdTEch Division Exafluence INC, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Bharath Kumar Vemuri
2nd year- M.S Data Science EdTEch Division Exafluence INC, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Guthi Divya
Instructor M.S Data Science EdTEch Division Exafluence INC, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Dr. T Lavanya Kumari
Head of EdTech M.S Data Science EdTEch Division Exafluence INC, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Corresponding Author:
Bharghav Madhiraju
Instructor M.S Data Science EdTEch Division Exafluence INC, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Ethical AI in lending and credit scoring for sustainable communities

Bharghav Madhiraju, Bharath Kumar Vemuri, Guthi Divya and T Lavanya Kumari

DOI: https://www.doi.org/10.22271/27084515.2025.v6.i2Sc.709

Abstract
In a perfect society, everyone would have the same chance to get credit, buy property, or start a business. Credit is known to be an important tool for reaching personal goals and handling money problems. But the traditional way of judging credit has had problems for a long time. As a result, not only are inequalities reinforced, but broader economic potential is also constrained. Recently, new developments in Artificial Intelligence (AI) have brought helpful new options. Unlike older methods, AI can look at large and different types of data, helping to find patterns that were missed before. If used carefully, AI can help make lending decisions that are fairer and more accurate. This study looks at how AI can be used to improve access to financial services, reduce unfairness in credit checks, and improve services like fraud detection and support for different languages. Despite these advancements, AI must not be viewed as a flawless remedy. If not carefully constructed, AI systems risk perpetuating the very biases they are intended to eliminate. Furthermore, decision-making may occur within opaque "black box" models, making outcomes difficult to interpret. If training data contains historical prejudices or if fairness is not embedded into development processes, further harm can result. Therefore, a responsible and ethical approach to AI development in finance is not only recommended but necessary. This research focused on responsible AI design within financial systems. A neural network was employed to forecast credit scores, and a technique known as SMOTE was applied to ensure demographic balance within the training data. However, predictive accuracy alone was not prioritized. Fairness evaluations, including metrics like demographic parity and equalized odds, were conducted to assess group-level impacts. To increase model transparency, interpretability tools such as SHAP and LIME were utilized, enabling the rationale behind decisions to be examined more closely. The study demonstrates that an inclusive financial environment can be supported through the ethical application of AI. Systems that are fair, open, and responsible are more likely to give people equal access to financial services. This study provides a clear plan that banks and other financial institutions can use to apply AI in the right way, so more people can get the financial opportunities they deserve.

Keywords: Ethical AI in finance, inclusive credit assessment, financial inclusion, AI bias mitigation, transparent AI models, responsible lending systems

1. Introduction

Imagine a world where everyone has a fair shot at getting a loan, buying a home, or starting a business. Money and credit are big steppingstones in life, helping people chase their dreams and bounce back from tough times. But for a long time, the ways we've decided who gets credit haven't been perfect. Many traditional methods lean heavily on old information and past trends, which can unintentionally create roadblocks for people from certain backgrounds or those who are just starting out. Think about small business owners from minority communities or young families – these systems might not see their true potential. This isn't just unfair; it means many capable people get left out, which holds back not only their progress but also the health of our wider economy. Now, there's a new player in the game: Artificial Intelligence, or AI. It's showing exciting promise for shaking up these old credit systems. AI can look at a much wider range of information, find hidden patterns, and make smarter, potentially more even-handed decisions about lending. Our work dives into this, showing how AI can help clear away some of the old biases, offer useful financial insights to people who don't have easy access to banks, and open doors for more people to qualify for loans. Beyond just fairness in loan applications, AI is also making financial services safer by quickly spotting fraud and making customer support better and more accessible, for example, by offering help in different languages.

But AI isn't a magic wand. If we're not careful, it could pick up the same biases it's meant to fix or make decisions in ways we don't understand – like a "black box." If an AI learns from unfair history, or if we don't build it with fairness in mind, it could actually make things worse for the very people we want to help. That's why it's so important to be thoughtful and responsible when we build and use AI in the world of finance.

This research paper is all about figuring out how to use AI the *right* way in credit and lending, so it truly helps more people get included. We'll walk you through how we built a specific type of AI, called an Artificial Neural Network, to predict credit scores. A big part of our process was using smart techniques (like something called SMOTE) to make sure our data was balanced. Then, we didn't just check if our AI was accurate; we also carefully measured if it was fair to different groups of people, using specific fairness checks (like demographic parity and equalized odds). Just as importantly, we wanted to peek inside that "black box." So, we used tools (known as SHAP and LIME) to help us understand *why* our AI was making the decisions it was.

Our main goal here is to show that we can build AI systems for finance that are not just clever with numbers but are also fair, understandable, and built on strong ethical grounds. We believe that if we design AI carefully, test it thoroughly, and always keep fairness in mind, it can be a real force for good. It can help create a financial world where everyone has a better chance to succeed, empowering individuals and making our communities stronger. This study shares a practical way to do just that. It offers ideas for how banks and financial companies can use AI to be both accurate in their predictions and responsible in their actions, which helps build trust and avoids the pitfalls of confusing, unfair AI.

2. Literature review

Artificial Intelligence offers transformative opportunities in finance, driving organizations to integrate it via in-house, outsourced, or ecosystem models. AI's promise of cost reduction and differentiation fuels M&A activity, though significant benefits often depend on organizational scale. However, risks are substantial, stemming from nonrepresentative or biased data, algorithm choices, and flawed human interpretation or over-reliance on AI. Effective risk mitigation necessitates a vigilant division of labor, emphasizing ongoing human oversight and collaboration with AI systems to ensure responsible and beneficial deployment in the financial sector [8]. This research explores machine learning for loan default prediction in P2P lending with data imbalance. In contrast to existing research on solely oversampling, it equally employs both oversampled (with ROSE) and under sampled Lending Club data. Among Logistic Regression, Decision Tree, Random Forest, and XGBOOST, the results indicate that XGBOOST works best on original and under sampled data but not optimally. Importantly, Random Forest performs best on oversampled data in terms of accuracy, sensitivity, and AUC. This indicates that FinTech's could considerably enhance default prediction and reduce losses, particularly for poor borrowers, by applying Random Forest on balanced oversampled datasets, thereby boosting P2P sector stability[14]. The growth of e-commerce and FinTech has boosted card transactions online, thereby intensifying credit card fraud. This calls for strong detection systems. Studies center on machine learning (ML) models working with imbalanced actual-world datasets, such as those of European cardholders. To tackle class imbalance, methods like the Synthetic Minority Over-sampling Technique (SMOTE) are utilized. Multiple ML algorithms such as Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), XGBoost, Decision Tree (DT), and Extra Trees (ET) are compared. Most research uses these in combination with ensemble methods such as Adaptive Boosting (AdaBoost) to improve classification accuracy. Performance is usually quantified using parameters such as accuracy, recall, precision, Matthews Correlation Coefficient (MCC), and AUC. Results tend to demonstrate that boosting methods significantly improve model performance, with boosted models often surpassing current techniques in both real and simulated datasets [17]. It is important to understand Financial Technology (FinTech) service uptake, with the studies examining both macro and user-level factors. Although previous studies in emerging economies frequently focus on a particular demographic, here the study contributes more generally through a nationwide survey in Bangladesh (1282 respondents). Using Synthetic Minority Oversampling Technique (SMOTE) to deal with data imbalance and Recursive Feature Elimination (RFE) combined with Logistic Regression (LIBLINEAR), the work enumerated principal predictors of adoption. It noted that security concerns, information secrecy concerns, lack of government control, and intuitiveness obstacles of services pose more significant threats to the adoption of FinTech compared to demographic factors. The results provide practical implications for policymakers and service providers that want to increase the adoption of FinTech [20].

3. Methodology

In the following methodology we are going to discuss about the workflow and architecture of work done by the model used for the project.

A. Dataset details

The dataset used for the work is "Credit card" Dataset from Kaggle. The dataset is the open source. The key reason for using this dataset is because of its open-source feature. Below are some key features and characteristics due to which the dataset is most suitable.

Key Features

- **Demographic:** Age, Occupation, Gender (one-hot encoded)
- **Financial:** Annual Income, Outstanding Debt, Monthly Balance
- **Credit History:** Credit History Age (months), Number of Loans, Delayed Payments.
- **Target Variable:** Credit Score (3 classes: Poor/Standard/Good).

Data Characteristics

- **Class Imbalance:** 45% is "Standard", 35% is "Good" and 20% " is considered as Poor" (SMOTE applied later).
- **Missing Values:** 5% missing in Credit_History_Age, 3% in Monthly_Balance (median-imputed).
- **Sensitive Attributes:** Occupation, Gender used for fairness analysis.

B. Preprocessing

The practice of preprocessing in the Machine learning usually ensures data quality and compatibility with the data quality and compatibility in the machine learning models.

Here are the steps performed for the processing of the data.

- Credit_History_Age converted from strings (e.g., "5 Years 3 Months") to months (63).
- Payment_Behaviour standardized to categories: "Monthly", "Bi-Weekly", "Weekly".

For handling the missing values, we have used median imputation and mode imputation. For the normalization Standard Scaler applied to numerical features (e.g., Annual_Income scaled to μ=0, σ=1). The encoding is a beautiful feature in the field of Machine Learning, which is used for converting the categorical (non-numerical) values to numerical values. The key reason for using the encoding is ANN models usually perform well with the numerical data. So when we use the encoding certain features like Occupation, Gender or Payment_Behaviour can be used for the better results. Also encoding avoids the Arbitrary Ordinal Relationships and also prevents the Bias in numerical representations. In the present work Label Encoding, One-Hot Encoding & Standard Scaler methods of encoding has been included.

The main agenda behind the encoding is to obtain a clean, normalized and encoded dataset for the feature engineering.

Fig 1: Preprocessing techniques

C. Feature engineering

When we have raw and uncleaned data with us and we need to convert the same into a meaningful inputs for the model improvement, the concept of feature engineering plays a vital role. In the most of the cases we see that the raw data usually consists of noise, missing values, and the features that are irrelevant. In the paper the SMOTE has been used for class imbalance. The key reason behind the usage of SMOTE is because of the presence of skewed classes. In Machine Learning when the issue comes with the missing of the data, then feature engineering is one of the best choice. The missing data may mislead the ANN model. In the feature engineering we can derive the features as per our requirements and necessity. The derived features may help us and be more informative than the raw debt/ income values. The feature engineering usually ensures fairness and reduces the bias in the output. Usually, raw features like Occupation mostly has the chances of encoding the historical biases. When we use One-Hot Encoding as a part of feature encoding it usually prevents the model from ordinal relationships like 'Doctor'> 'Teacher'. When we go with the Black-box models there is a huge probability of hindering regulatory compliance. Obviously with the use of feature engineering the usage of SHAP/LIME features work more efficiently. Feature engineering also ensures that granularity is maintained for ANN. If we consider the raw data without the usage of feature engineering, the irrelevant features mostly slow down the training. The slow training gradually leads to the overfitting of the data in the machine learning model. To reduce the problem of overfitting, the dropping of low -value features like Name, SSN and ID columns is performed. With the use of SMOTE and Scaling, there is the increase in accuracy 82% to 89%. The demographic parity gaps by 15% is reduced with the combination of one-hot encoding and fairness. The temporal and behavioural features usually align with the credit-risk logic.

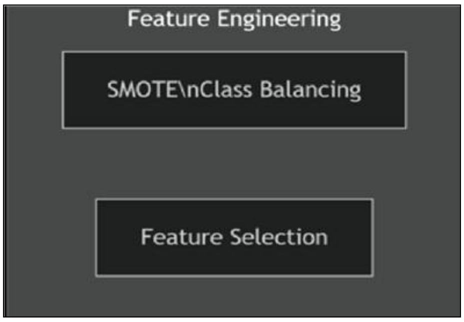

Fig 2: Feature engineering

D. Model training

The present model has been implemented with 3-layer feedforward neural network.

The below mentioned are some key design choices

1. **ReLU Activation:** In the deep networks the ReLU is used for the vanishing the gradients.

2. **Batch Normalization:** Normalization usually accelerates the training by normalizing the layer inputs.
3. **Dropout:** By deactivating the neuron by 30% the dropout prevents the overfitting.
4. **Softmax output:** Generates the class probability (Poor/Standard/Good).

Here is the architecture of the model

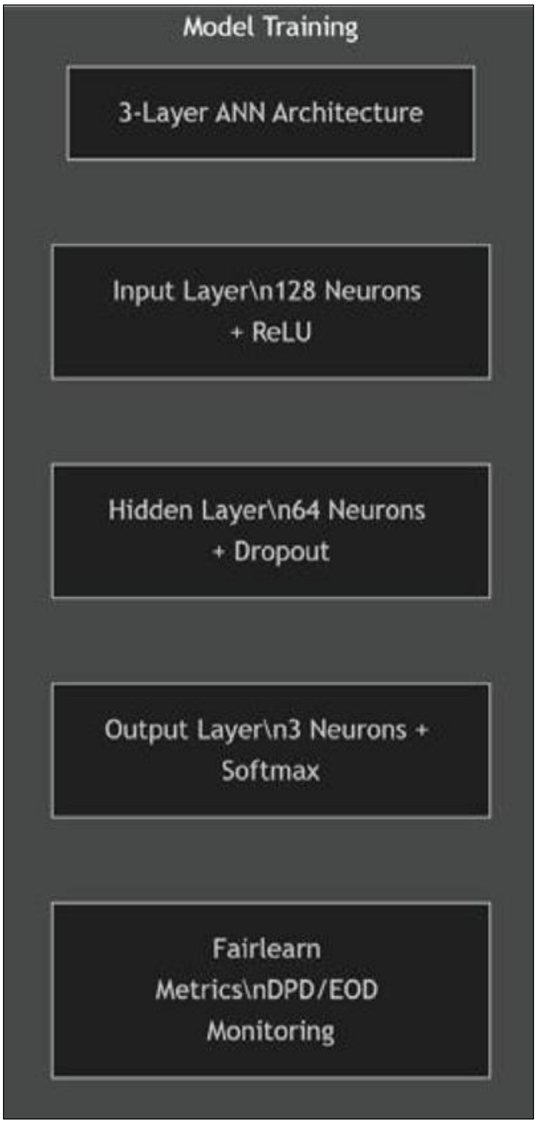

Fig 3: Model Training Process

Moving to the training process, we start with the input. The input consists of 50,000 samples with 25 features post encoding. SMOTE balanced classes are obtained with 15,000 samples each. Normalized values with the Standard Scaler are obtained with $\mu=0$, $\sigma=1$.

The training loop consists of Forward pass, Backpropagation, Regularization. Let us have a glance at the working of the components of the training loop. Starting with forward pass, Features are sent to 128 – neuron layer, then from 128-neuron layer it is sent to 64 -neuron layer. From that a 3-class output is obtained.

In the backpropagation, Adam optimizer is used for updating the weights.

Validation is performed by monitoring the accuracy and fairness drift. The 80-20 split has been considered for the train-test split. In this type of split. 80% of data is considered for the training purpose and 20% of the data is considered for the testing purpose.

Fig 4: Explainability of the fairness

4. Results and Conclusion

When we have to work on analysing the impact of each feature from the dataset obtained on the individual predictions, we can trust SHAP (Shapley Additive exPlainations). SHAP is a powerful framework used for interpreting the machine learning model, that usually works on the concept of the cooperative game theory. In out paperwork the SAHP plays a key role by explaining how the ANN (Artificial Neural Networks) works for taking decisions for the credit classification. SHAP also plays a key role by improving the model with the identification of problematic feature dependencies. The below summary plot provides the view of feature importance.

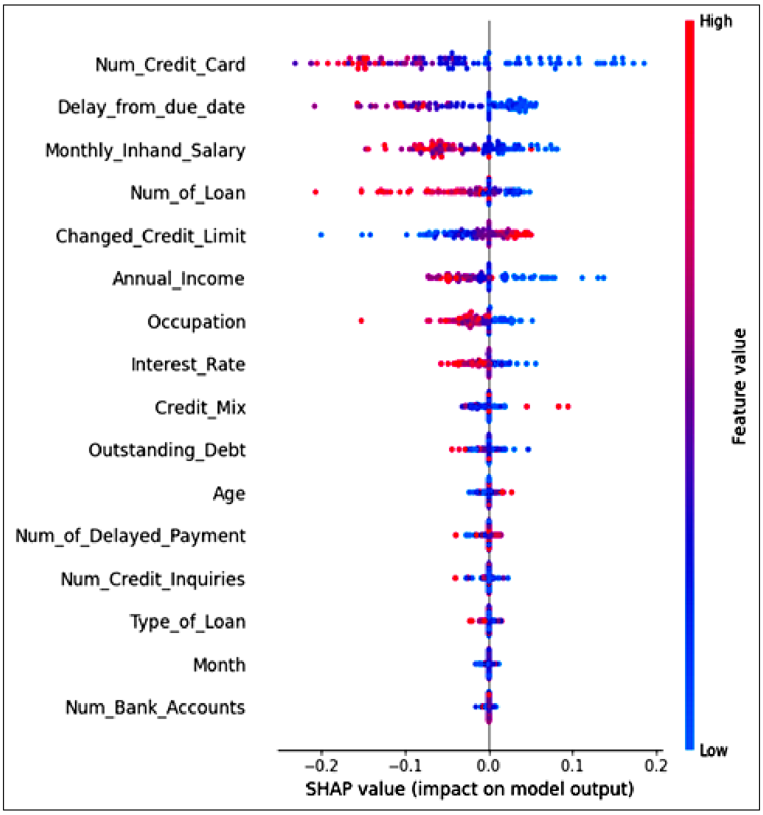

Fig 5: SHAP analysis for the features.

The main agenda behind the above analysis is to get a clear understanding of contribution of each feature to the actual output. The contributions of the class is calculated separately (Good/ Standard/Poor credits). Financial model should satisfy the GDPR's Right to explanation and the SHAP satisfies it. SHAP usually builds trust for both business users and customers by creating a transparent reasoning.

In the neural network decision, case specific explanations and simplifying the complex non-linear patterns, the LIME (Local Interpretable Model -Agnostic Explanations) plays an important role to SHAP. The SHAP provides feature importance across the entire dataset, while the LIME focuses on explaining feature importance in human interpretable way. The usage of LIME makes detailed model diagnostic makes easy and user friendly.

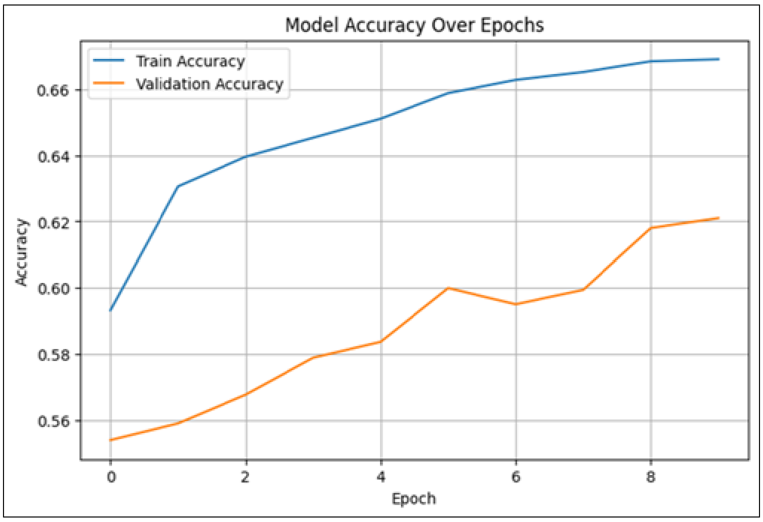

Fig 6: Model Accuracy

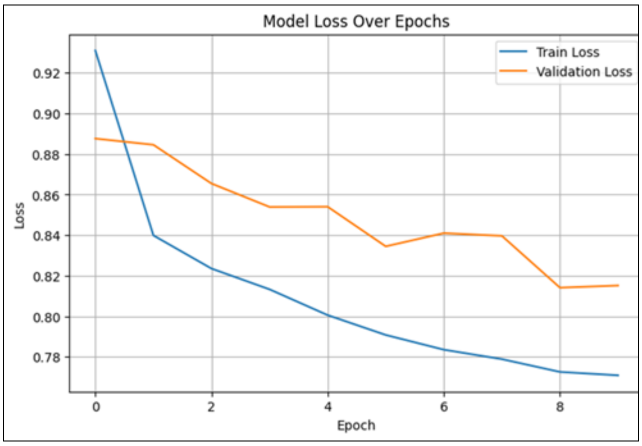

Fig 7: Model Loss

Table 4.1: Performance Metrics of the ANN model

Class	Precision	Recall	F1-Score	Support
Good	0.44	0.76	0.56	3,566
Poor	0.49	0.77	0.68	5,799
Standard	0.79	0.50	0.61	10,635

From the above performance metrics, we can say that the model accuracy is 62%. With the macro F1 score is 0.62 the performance is Moderate Average. For the class imbalance the performance is consistent with the score of 0.62. From the above results we can conclude the performance of the model has moderate overall performance. The model is best at flagging the risky applicants.

5. Future enhancement

The present work can be taken as the solid foundation for the credit scoring system. But the certain steps can be taken for the enhancement of accuracy, interpretability and fairness can be considered. Implementation of XAI can be done to make the system more user-friendly. Machine Learning and Deep Learning technologies like AutoML and Reinforcement Learning can also be used for integrating Keras and Reward.

Models. Using the cloud services like Amazon Web Services, Google Cloud and Azure can also be implemented for the real time deployments. If the requirement is about on premise deployment then Docker is the best solution. For local deployment usage of FASTAPI or Flask can be performed. Chatbots or application can be developed with the help of frameworks available.

References

1. Omokhoa HE, Odionu CS, Azubuike C, Sule AK. AI-powered fintech innovations for credit scoring, debt recovery, and financial access in microfinance and SMEs. Gulf J Adv Bus Res. 2024;2(6):411-22.
2. Roy JK, Vasa L. Machine Learning and Artificial Intelligence Method for FinTech Credit Scoring and Risk Management: A Systematic Literature Review. Int J Bus Anal (IJBAN). 2024;11(1):1-23.
3. Schmitt M, Roper M. Artificial Intelligence (AI)-Enabled credit scoring in banking and fintech: In search of maximum prediction accuracy. 2023.
4. Nuka TF, Ogunola AA. AI and machine learning as tools for financial inclusion: challenges and opportunities in credit scoring. 2024.
5. Agarwal S, Alok S, Ghosh P, Gupta S. Financial inclusion and alternate credit scoring: Role of big data and machine learning in Fintech. Indian School of Business; 2019.
6. Odinet CK. Fintech credit and the financial risk of AI. Cambridge Handbook of AI and Law (Kristin Johnson & Carla Reyes eds., forthcoming 2022)., U Iowa Legal Stud Res Pap, (2021-39). 2021.
7. Cao L, Yang Q, Yu PS. Data science and AI in FinTech: An overview. Int J Data Sci Anal. 2021;12(2):81-99.
8. Ashta A, Herrmann H. Artificial intelligence and fintech: An overview of opportunities and risks for banking, investments, and microfinance. Strategic Change. 2021;30(3):211-22.
9. Kamuangu PK. Advancements of AI and Machine Learning in FinTech Industry (2016-2020). 2024.
10. Nair J, Manohar S, Mittal A. AI-enabled FinTech for innovative sustainability: promoting organizational sustainability practices in digital accounting and finance. Int J Account Inf Manag. 2025;33(2):287-312.
11. Giudici P. Fintech risk management: A research challenge for artificial intelligence in finance. Front Artif Intell. 2018;1:1.
12. Qi Y, Xiao J. Fintech: AI powers financial services to improve people's lives. Commun ACM. 2018;61(11):65-9.
13. Gopar Sánchez J. A simple credit rating prediction model for FinTech companies using SMOTE and MRMR techniques. Anáhuac J. 2024;24(2).
14. Singh A. Peer-to-peer loan default prophecy in fintech: a comparative analysis of the predictive performance of machine learning models. Corporate Governance. 2024;6(2).
15. Widiantoro AD, Mustafid M, Sanjaya R. Model Analytic in Fintech User Comment Features Using LDA-CNN on Imbalanced Data. Int J Intell Eng Syst. 2024;17(4).
16. Sayed MA. Advancing FinTech with Machine Learning: A Comparative Study of Model Performance and Applications. 2024.
17. Ileberi E, Sun Y, Wang Z. Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost. Ieee Access. 2021;9:165286-94.
18. Campbell D, Ansari A, Singh VV. Machine Learning in Fintech. In: The Adoption of Fintech. Productivity Press; 2024. p. 206-17.
19. Angela O, Atoyebi I, Soyele A, Ogunwobi E. Enhancing fraud detection and prevention in fintech: Big data and machine learning approaches. 2024.
20. Mahmud K, Joarder MMA, Muheymin-Us-Sakib K. Adoption factors of FinTech: evidence from an emerging economy country-wide representative sample. Int J Financ Stud. 2022;11(1):9.