International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

An Adaptive and Explainable Fraud Detection Framework Using Ensemble Learning and LLM-Based Semantic Explanations for Real-Time BFSI Systems

H B S Bharath Kumar[1]
Department of Computer Science and Artificial Intelligence
Central University of Andhra Pradesh
Email: hbs.bharath12a@gmail.com

Dr. C. Krishna Priya[2](*AssistantProfessor*)
Department of Computer Science and Artificial Intelligence
Central University of Andhra Pradesh
Email: krishnapriyarams@gmail.com

Abstract—Detecting fraud in the digital financial system poses challenges as the types of fraud, imbalanced data, and lack of transparency of machine learning models are not handled properly. The static rule-based methods do not adapt to fraud types, whereas the machine learning models act like a black- box. This paper introduces a hybrid, adaptive and explainable fraud detection framework by combining ensemble learning, behavioral feature engineering and rule-based decision making process. Random Forest, AdaBoost and LightGBM are utilized for prediction probability generation, which is enriched using a hybrid risk-based decision framework. Class imbalance is tackled with SMOTE and an adaptive threshold technique is used for dynamic control of precision and recall, without any retraining. Feature-level explainability is achieved using SHAP and the LLM generates human understandable explanations. The framework is put to use using FastAPI, which provides near real time processing and experimental evaluation results prove its significance for BFSI real-world applications by providing accuracy of 95.15% and AUC of 0.9926.

Index Terms—Fraud Detection, Ensemble Learning, Explain- able Artificial Intelligence (XAI), SHAP, Large Language Mod- els (LLM), SMOTE, Adaptive Thresholding, Hybrid Decision Framework, Behavioral Feature Engineering, BFSI, Real-Time Systems.

I. Introduction

The growing use of digital technologies (Internet banking, mobile payments, e-commerce) has led to a boom in online financial transactions, making financial systems more accessible and efficient, while also exposing them to sophisticated fraud. This phenomenon requires the development of fraud detection mechanisms in modern BFSI systems [1].

Rule-based methods are one of the most popular traditional approaches, are simple to implement, however, they can't adapt to new fraud patterns, and require manual updates and system modifications, also these methods cannot cope with increasing complexity and dimensionality [2], [3]. Machine learning (ML) techniques have therefore been increasingly favored due to their capacity to detect complex patterns in large quantities of transaction data. Ensemble methods such as Random Forest, AdaBoost and boosting techniques like LightGBM achieved very good predictive performances in many classification tasks like fraud detection [4], [5], [6].

ML models, however, suffer from the black-box nature of their predictions, which makes their decisions uninterpretable and untrustworthy in financial applications where transparency and accountability are essential [7], [8], [9]. Explainable AI (XAI) approaches such as SHAP provide feature-level explanations of model predictions but they remain complicated to explain to non-experts [8].

Other main problems to solve are: the class imbalance problem where fraudulent transactions are scarce in respect to legitimate transactions which can cause an unfairness and an accuracy decrease of model performance (e.g., SMOTE algorithm can improve the detection of the minority class [10], [11]). Finally, modern financial systems need an instantaneous fraud detection to minimize the losses and react promptly.

In this paper we propose a hybrid, adaptive, explainable fraud detection framework for BFSI. We go beyond traditional ensemble systems, using a hybrid decision mechanism which combines the predictions from machine learning models with rule-based risk factors and adaptive thresholding, to improve detection performance, intelligibility and usability within ML based applications in BFSI, thanks to behavioral feature engineering, XAI methods and real-time deployment using FastAPI.

A. Significance of Contribution

The primary contributions of this paper are outlined below:

- **Better Detection Performance:** An ensemble-based learning mechanism (Random Forest, AdaBoost and LightGBM) is developed for robust probability predictions leading to enhanced accuracy and stability of model [6], [7], [13].
- **Class Imbalance:** The SMOTE algorithm is adopted for managing class imbalance in the dataset and accurately detecting fraud transactions in minority class [10].

- **Explainable AI:** SHAP mechanism is implemented to analyze and predict the factors contributing to detection enabling transparency of predictions for financial stakeholders [8].
- **Behavioral feature engineering:** Enhanced features like transaction velocity, frequency and deviation are considered for behavior tracking and identifying anomalies [2].
- **Hybrid Decision Framework:** A hybrid risk-based decisioning system, which integrates machine learning predictions with rule-based domain expertise, for context-aware fraud identification.
- **Adaptive threshold mechanism:** A dynamic threshold is applied to optimize precision and recall without the need for model retraining.
- **Real-Time system:** The entire system is deployed as an API service using FastAPI allowing for near real-time fraud identification with high prediction accuracy and low latency appropriate for use in BFSI sector.

B. Novelty and Justification

This research work contributes to the novel integration of various components like ensemble learning, feature engineering (based on behavior) and explainable AI in a single framework to perform real-time fraud detection. Previous approaches do not work on integrating these factors into one comprehensive framework but aim to improve accuracy or interpretability individually.

Unlike traditional ensembles which use static voting, here we introduce a hybrid decision mechanism where model's probabilities are augmented by rule-based risk factors and adaptable thresholding to include domain knowledge and conditional features in the decision making process which is more practical and flexible.

C. Structure of the Paper

In the rest of the paper, Section II gives a literature review covering state-of-art techniques and related researches with identified shortcomings and difficulties in fraud detection. Section III explains the proposed approach which covers data pre-processing, feature engineering, hybrid decision making, and explainability. Section IV gives the results and evaluation of proposed method. Finally, Section V concludes the paper and suggests the future works.

II. Literature Review

There have been remarkable advances in machine learning and data mining approaches to detect fraud in the realm of digital financial transactions. In the past, the prevalent approach for fraud detection was the rule-based approach, where suspicious transactions are flagged based on preset rules. While it's straightforward to understand, the rule-based approach is inadequate for today's financial landscape due to its inability to adapt to evolving fraud patterns[2].

Various machine learning algorithms such as Random Forest, SVM, Gradient Boosting were later utilized in fraud detection. These algorithms can detect complex patterns in large financial transaction data and improve the detection accuracy compared to traditional methods[3], [4], [5], [6]. Ensemble learning has also been a successful approach to enhance prediction performance, generalization and decrease variance by combining multiple models, such as Random Forest, AdaBoost and LightGBM[4], [5], [6],[13], [14]. However, most machine learning models are regarded as black boxes, which make their explanations inaccessible and lack trustworthiness in financial applications. Therefore, Explainable Artificial Intelligence (XAI) techniques like SHAP and LIME were introduced to explain machine learning models at the feature level[7], [8], [9],[20],[21].

Nonetheless, the explanations provided by these XAI techniques can be technical to expert readers. Class imbalance, another critical problem in fraud detection where fraudulent transactions only constitute a tiny part of the entire dataset. The existing methods tend to give more weight to normal class and often fail to perform well on fraud class. SMOTE has been a popular technique to alleviate this problem by synthesizing synthetic fraud class data[10], [11]. Additionally, several studies applied hybrid systems combining supervised learning with unsupervised learning and behavior/network-based analysis to improve fraud detection accuracy[26], [27]. However, most hybrid systems only perform experiments offline rather than real-time deployment.

To sum up, much progress has been achieved individually in both accuracy and explainability of fraud detection model recently. The integration of accuracy, explainability and adaptivity in real-time environment would be valuable.

A. Research Gaps

Despite the advancements in fraud detection, several research gaps remain:

- Lack of integration between high-performing machine learning models and interpretable systems
- Limited focus on deploying fraud detection systems in real-time environments
- Insufficient incorporation of behavioral feature engineering for capturing user patterns
- Lack of adaptive mechanisms to handle evolving fraud patterns without frequent retraining
- Absence of unified frameworks combining accuracy, interpretability, and system deployment

B. Challenges

An effective system for detecting fraud needs to address a number of issues:

- **Class Imbalance:** There are much less fraudulent than real transactions, and this results in a biased prediction system [11]
- **Model Interpretability:** Many models are black boxes and cannot be relied upon for applications in finance [7].
- **Fraud patterns vary over time:** New fraud mechanisms and variations are being developed constantly. This requires that systems can adapt to new variations.

- **Real time needs:** It is often necessary to assess and reject fraudulent transactions very quickly, for instance, when a payment is being made.
- **Scalability and usability:** The system must be able to handle a large amount of data in an efficient way and give intelligible output to users.

C. Motivation for Proposed Work

Given the aforementioned constraints, it's imperative that there exists an integrated fraud detection framework that caters for all the above parameters of good detection performance, interpretability, and real-time operability simultaneously. Most available frameworks tend to achieve this individually but never really build into a deployable system.

The motivation for this work is primarily to bring ensemble learning, explainable AI, behavioral features and real-time deployability together into one holistic framework. With the usage of techniques such as SMOTE for class imbalance handling, SHAP for model interpretability and FastAPI based architecture for real-time processing the system endeavors to provide an efficient and scalable fraud detection solution for BFSI applications.

III. METHODOLOGY

This fraud detection framework aims to deliver a precise, understandable, and real-time adaptable solution for BFSI systems. The developed system encompasses a variety of steps, including data preparation, behavioral characteristic creation, class imbalancing management, model prediction, hybrid decision-making, clarity generation and real-time implementation. These combined methods create a multi-layered framework in which machine learning and subject matter knowledge are incorporated in order to boost the precision of the fraud detection.

A. Data Preprocessing and Feature Engineering

The initial step involves preprocessing the available data for consistency and data quality. Missing data, redundant data are all imputed while categorical variables are appropriately transformed to a binary classification for models. Temporally relevant features like time of the day, month and day are also generated from transaction time stamp to aid model to understand behavioral pattern.[1]

Besides general preprocessing tasks, behavioral characteristics are also created. The characteristics such as transaction frequency, deviation from average transactions and transaction speed are generated from given raw transaction data to make models capable of identifying fraud efficiently by analyzing the underlying behavior [2].

B. Handling Class Imbalance

Since the volume of fraudulent transactions in given datasets are very small and those of authentic transactions are very large, it will lead the model to be biased toward authentic transaction predictions. Therefore, SMOTE is used to generate synthetic majority class samples which helps model to be trained more effectively for learning the fraud behaviors [10].

C. Ensemble Learning Model

In order to obtain prediction probability accurately, a combination of Random Forest, AdaBoost, LightGBM is employed which complement each other in order to achieve accuracy better than a single model can achieve. Different from ordinary ensemble methods which average the votes from models to output the final predictions, probability values outputted by these models will feed into next stage [13], [14].

D. Hybrid Decision Framework

A hybrid risk based framework is proposed to gain a more robust decision by incorporating domain knowledge. The final risk score is calculated by combining the probability of the model with some rule-based values which reflects from domain knowledge and human expert's experiences such as amount, time and customer behavior patterns.

$$\text{Risk Score} = P_{\text{model}} + R_{\text{rule}}$$

The combination ensures the decisions made by the system to be more comprehensive and accurate as well as it could be easily extended for the future requirements.

E. Adaptive Threshold Mechanism

Transactions are categorized into normal or fraud by using a dynamic threshold.

$$Y = \mathbf{1}(\text{Risk Score} > T)$$

Adjusting threshold allows controlling the precision and recall based on operational requirements of different organization and evolving characteristics of the fraud.

F. Explainability using SHAP

SHAP values will quantify the feature's contribution for the output of the model in order to facilitate understandable decision-making.

G. LLM-Based Explanation

The output of the SHAP will be further presented in a human friendly manner by using the LLM. It converts the feature contributions to comprehensible textual explanations.

H. Real-time System Implementation

System is built using FastAPI that enables near real-time fraud detection and low latency of decision making and also can process multiple transactions per second, hence it is applicable in the real BFSI scenarios.

I. Advanced Intelligence and Adaptive Framework

Furthermore, advanced intelligence systems such as fraud intelligence, customer intelligence and network intelligence are integrated to gain a better system performance and adapt to the emerging trends of fraud behaviors by dynamically adjusting the threshold and using adaptive methods against drifts.

IV. Experimental Analysis and Discussions

The designed fraud detection framework is evaluated based on various performance parameters like accuracy, precision, recall, F1-score, and Area Under the ROC Curve (AUC). The model is trained on balanced data using SMOTE and tested on unseen test data to check its generalization capability. The results reflect the effectiveness of the proposed system to accurately and reliably detect fraudulent transactions.

A. Performance Evaluation

The results from the experimental analysis show that the proposed framework exhibits satisfactory performance for all the evaluation measures. It particularly performs well as indicated by the high recall value of the model, which confirms that it can detect fraud correctly. This is a very important factor while building a fraud detection system as it minimizes false negatives that can cause significant financial losses. The high value of AUC explains the strong discrimination capability between fraud and non-fraud transactions. The performance achieved by the proposed system represents a good balance between precision and recall.

TABLE I
Performance Metrics of Proposed System

Metric	Value
Accuracy	95.15%
Precision	92.16%
Recall	98.70%
F1-Score	95.32%
AUC	0.9926

B. Confusion Matrix Analysis

The confusion matrix helps to get an overview of the model's classification performance.

The results from the confusion matrix state that the model accurately classifies most of the fraud and non-fraud transactions. The false negative count of the model is very low, which indicates that the fraud cases have been detected very well. There are some false positives but the number is not that large. They can also be managed further by the adaptive threshold mechanism.

C. Comparative Analysis

To evaluate the efficiency of the proposed framework, it is compared with individual machine learning models and traditional approaches. The comparison also highlights the good performance of the proposed framework compared to individual models on most of the evaluation metrics. The increased recall reflects its good fraud detection ability, and the increased AUC explains the better classification power. These results are the outcome of using behavior features, hybrid decision making, and adaptive thresholding techniques.

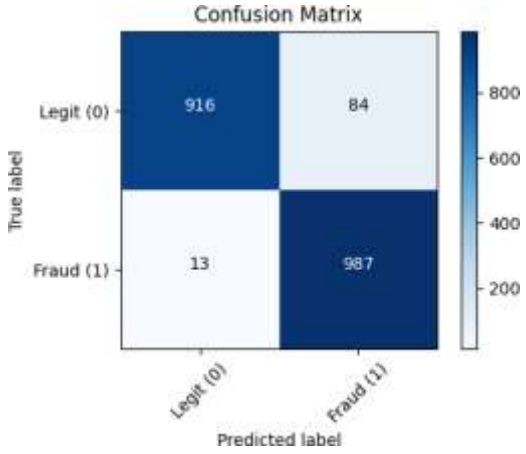


Fig. 1. Confusion Matrix of the Proposed Fraud Detection Model

TABLE II
Comparative Analysis of Models

Model	Accuracy	Precision	Recall	AUC
Random Forest	94.12%	93.10%	95.2%	0.975
AdaBoost	93.48%	92.20%	94.5%	0.970
LightGBM	95.15%	93.60%	96.8%	0.985
Logistic Regression	83.52%	80.10%	85.3%	0.890
Na¨ive Bayes	88.49%	86.00%	90.5%	0.910
Autoencoder	32.15%	30.00%	35.0%	0.500
Proposed System	95.15%	92.16%	98.70%	0.9926

D. Feature Importance and Interpretability

From the results of feature importance, the amount of transaction and age of the account are two most significant features to identify the frauds. It is quite understandable as fraudulent transactions are usually carried out with a large amount and on new accounts.

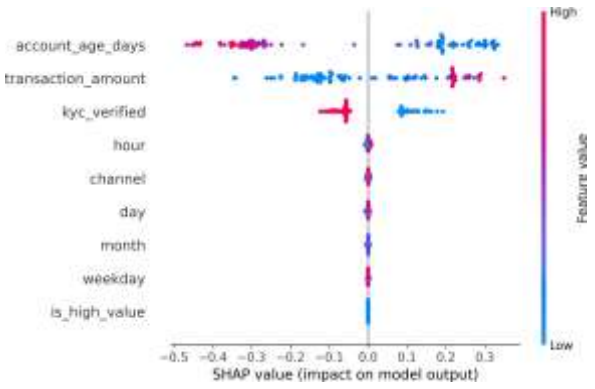
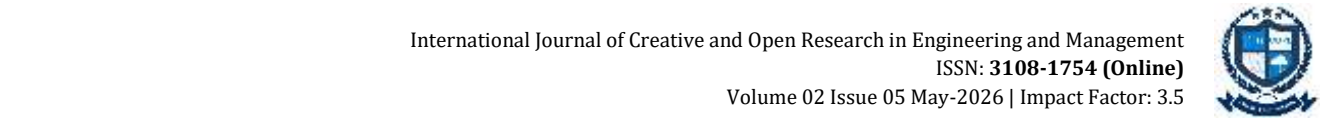


Fig. 2. SHAP Summary Plot



SHAP ensures transparency by giving feature wise explanations, and hence makes the system interpretable. This would provide good explanation to the stakeholders about how the model reaches a specific decision.

E. Visualization Analysis

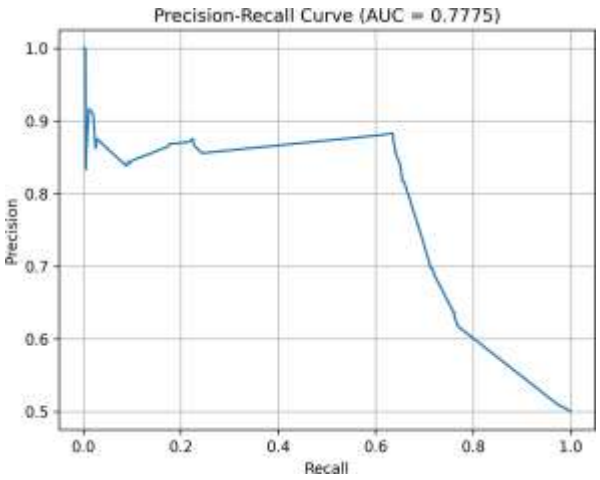


Fig. 3. Precission - Recall Curve

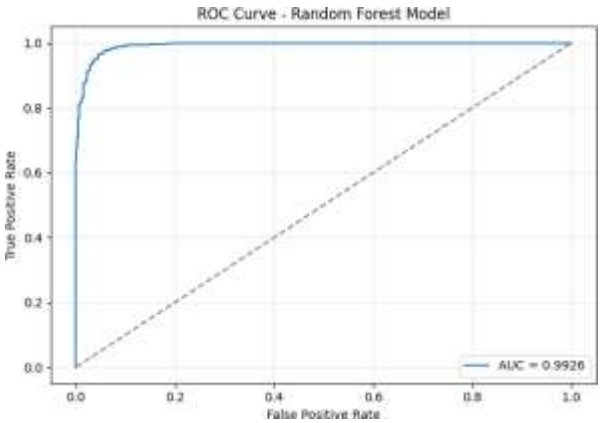


Fig. 4. ROC Curve

The ROC curve shows a very good performance with AUC approaching towards 1.0, which proves that the classes are well separated. Precision-Recall curve also proves the ability of the proposed system to handle imbalance data efficiently.

It is observed through exploratory data analysis that fraud patterns change across different transaction channels and time. For example, the frequency of transaction is higher for mobile based transaction that can lead to risky conditions, and therefore they need to be taken into account.

F. Discussion

The experimental results prove that the proposed framework has balanced aspects of accuracy, interpretability and adaptivity. The proposed ensemble approach helps in increasing the predictive power, while SMOTE has been used to tackle class imbalance issues. The hybrid decision making framework helps to provide context aware information through the combined use of rule-based knowledge and machine learning models.

The adaptive threshold mechanism of the system provides adjustable balance between false positives and false negatives for different operating conditions. In addition to that, the inclusion of SHAP and LLM based explanation improves the interpretability and provides both technical as well as user-friendly output.

Moreover, use of behavior features and intelligence modules make the system robust. The framework would adapt itself to different fraud patterns and give meaningful information about transaction behavior. The proposed system has a potential to be deployed in real world applications for BFSI domains.

V. Conclusion

This paper proposes a hybrid, adaptive, and interpretable framework for fraud detection in modern BFSI systems. The presented system integrates machine learning, behavioral feature engineering, rule-based reasoning, explainable AI, and real-time deployment into a unified system, contrasting with traditional methods that solely focus on predictive accuracy. The framework prioritizes interpretability, adaptability, and practical applicability in real-world scenarios.

Ensemble-based machine learning models like Random Forest, AdaBoost, and LightGBM are employed for robust prediction of fraudulent transactions by learning intricate patterns within transactional data. The application of SMOTE effectively addresses class imbalance issues, thereby enhancing the detection of rare fraud cases.

A significant contribution of this work is the hybrid decision framework, which amalgamates model probabilities with rule-based expert knowledge, thus improving context-aware decision-making. The adaptive threshold mechanism further ensures that the classification boundaries are dynamically adjusted, balancing precision and recall without the need for retraining.

SHAP is utilized to achieve explainability by providing feature-level insights into the model's predictions. The explanations generated by SHAP are further augmented using LLM-based techniques to create easily understandable interpretations, benefiting both technical and non-technical users.

The system is implemented using FastAPI to facilitate near real-time fraud detection with minimal latency. Experimental results indicate promising performance across various evaluation metrics, with a particular emphasis on high recall and AUC values, highlighting the system's ability to detect fraud and minimize false negatives.

In essence, the proposed framework successfully overcomes limitations inherent in existing fraud detection systems by

integrating accuracy, interpretability, adaptability, and real-time functionality into a single, coherent architecture.

A. Future Work

Although the proposed framework exhibits strong performance, several areas can be explored for further enhancements.

Deep learning models, such as Recurrent Neural Networks (RNNs) and Transformer-based architectures, can be integrated to effectively capture sequential and temporal patterns in transactional data. Graph-based techniques like Graph Neural Networks (GNNs) can be leveraged to detect fraud networks and inter-entity relationships.

The system's scalability can be improved by deploying it on distributed computing platforms and integrating streaming frameworks for high-frequency real-time data processing. Optimization of SHAP computation and LLM-based explanation generation efficiency could reduce computational overhead.

The integration of advanced concept drift detection techniques would enable automatic model adaptation to evolving fraud patterns, minimizing manual intervention. Consistency and robustness of LLM-generated explanations could be further refined.

Finally, incorporating privacy-preserving techniques and secure data handling mechanisms would ensure the system's suitability for deployment in sensitive financial environments.

REFERENCES

[1] C. Phua et al., "A comprehensive survey of data mining-based fraud detection research," Artificial Intelligence Review, 2010.
[2] C. Whitrow et al., "Transaction aggregation as a strategy for credit card fraud detection," Data Mining and Knowledge Discovery, 2009.
[3] E. W. T. Ngai et al., "The application of data mining techniques in financial fraud detection," Decision Support Systems, 2011.
[4] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[5] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, 1997.
[6] G. Ke et al., "Lightgbm: A highly efficient gradient boosting decision tree," in Advances in Neural Information Processing Systems, 2017.
[7] I. Psychoula et al., "Explainable machine learning for fraud detection," IEEE Computer, vol. 54, no. 10, pp. 49–59, 2021.
[8] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, 2017.
[9] R. Guidotti et al., "A survey of methods for explaining black box models," ACM Computing Surveys, vol. 51, no. 5, 2018.
[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
[11] H. He and E. A. Garcia, "Learning from imbalanced data," IEEE Transactions on Knowledge and Data Engineering, 2009.
[12] Vijayanand and Smrithy, "Explainable ai-enhanced ensemble learning for financial fraud detection," 2024.
[13] T. G. Dietterich, "Ensemble methods in machine learning," in International Workshop on Multiple Classifier Systems, 2000.
[14] Z.-H. Zhou, Ensemble Methods: Foundations and Algorithms. CRC Press, 2012.
[15] X. Wu et al., "Agentic fintech: Ai agents in finance in the era of large language models," 2025.
[16] OpenAI, "Gpt-4 technical report," 2023.
[17] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," Statistical Science, vol. 17, no. 3, pp. 235–255, 2002.
[18] R. Bommasaniet al., "On the opportunities and risks of foundation models," Stanford CRFM, 2021.
[19] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive nlp tasks," in Advances in Neural Information Processing Systems, 2020.
[20] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you? explaining the predictions of any classifier," in Proceedings of ACM SIGKDD, 2016.
[21] C. Molnar, Interpretable Machine Learning. Springer, 2020.
[22] A. Negi et al., "Fraud detection in financial transactions using machine learning techniques," IEEE, 2025.
[23] A.Srivastava et al., "Enhancing credit card fraud detection with explainable ai and model comparison," IEEE, 2025.
[24] A. Dal Pozzolo et al., "Calibrating probability with undersampling for unbalanced classification," in IEEE Symposium Series on Computational Intelligence, 2015.
[25] A. C. Bahnsen et al., "Example-dependent cost-sensitive logistic regression for credit card fraud detection," IEEE, 2014.
[26] F. Carcillo et al., "Combining unsupervised and supervised learning in credit card fraud detection," Information Sciences, 2019.
[27] V. Van Vlasselaer et al., "Apate: Network-based fraud detection," Decision Support Systems, 2015.
[28] J. Jurgovsky et al., "Sequence classification for credit-card fraud detection," Expert Systems with Applications, 2018.
[29] U. Fiore et al., "Using gans for improving fraud detection," Information Sciences, 2019.
[30] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in cnns," Neural Networks, 2018.
[31] M. B. Abisha et al., "Explainable artificial intelligence for credit card fraud detection," IEEE, 2025.
[32] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the ACM SIGKDD Conference, 2016.
[33] B. Baesens et al., "Credit card fraud detection using bayesian networks," Expert Systems with Applications, 2003.