IRJEMS International Research Journal of Economics and Management Studies
ISSN: 2583 – 5238 / Volume 5 Issue 4 April 2026 / Pg. No: 176-182
Paper Id: IRJEMS-V5I4P122, Doi: 10.56472/25835238/IRJEMS-V5I4P122

Original Article

Cost-Sensitive Credit Card Fraud Detection Using Extreme Gradient Boosting and SMOTE on Imbalanced Financial Datasets

[1]Md. Rasel Uddin, [2]MD. Rahim, [3]Md. Sadman Hafiz, [4]Atika Rahman Ounte
[1,2]Department of Accounting and Information Systems, Begum Rokeya University, Rangpur, Bangladesh.
[3]Department of Political Science, Begum Rokeya University, Rangpur, Bangladesh
[4]Department of Statistics, Begum Rokeya University, Rangpur, Bangladesh.

Received Date: 19 March 2026 Revised Date: 07 April 2026 Accepted Date: 16 April 2026 Published Date: 23 April 2026

Abstract: *As the number of digital payment systems expands at an exponential pace, it increases the risk of financial abuse, such as credit card fraud. Standard comprehension models often work inefficiently when they have to deal with skewed datasets, including those that have large class imbalances. This often leads to false positives, which stand in the way of genuine transactions and make operations challenging. This study examines the friction-fraud trade-off by employing and refining a cost-sensitive Extreme Gradient Boosting (XGBoost) model integrated with SMOTE on a highly imbalanced financial dataset, concentrating on attaining an optimal detection threshold that markedly enhances current benchmarks for balancing fraud detection and reducing false positives. The research evaluates model performance using precision-recall metrics to optimize the classification threshold for real-world applications. Our findings show that the XGBoost design is better than other ensemble models, which have an Area Under the Precision-Recall Curve (AUPRC) of 0.0156. The model uses an optimal detection level of 0.0321 to keep a balance between finding fraud and actual user behaviour. Our study of feature importance demonstrates that categorical infrastructure, such as types of devices and merchant categories, is a much better way to predict fraud than continuous numerical data. These findings facilitate the implementation of tiered fraud triage systems, ultimately optimizing the balance between robust financial security and seamless business continuity.*

Keywords: *Cost-Sensitive Learning, Credit Card Fraud, Extreme Gradient Boosting (Xgboost), Imbalanced Datasets, Precision-Recall Optimization.*

I. INTRODUCTION

The rapid expansion of digital payment ecosystems has transformed modern retail banking and global commerce, enabling billions of transactions to be processed electronically each day (Theodorakopoulos et al., 2025) electronically. While these technologies provide unprecedented convenience and financial accessibility, they also create significant vulnerabilities for financial crime, particularly credit card fraud(Taha & Malebary, 2020). As digital business has grown quickly, so has fraud in electronic payment systems, causing the world financial system to lose a lot of funds (Btoush et al., 2023). As the number of transactions keeps going up, traditional banks and other financial institutions have to find a way to identify fraud in real time without disrupting honest customers.

Traditional fraud detection systems used rule-based methods like fixed transaction limits, geographic restrictions, or blacklist filters. These systems worked well at stopping fraud in the initial stages, but they can't keep up with fraud patterns that are getting more complicated. Modern fraud schemes often look like real consumer behaviour which indicates that static rules fail to identify sophisticated attacks. Consequently, machine learning techniques have become central to contemporary fraud detection research, with ensemble algorithms such as Random Forest and Extreme Gradient Boosting (XGBoost) demonstrating strong performance in identifying hidden patterns within large financial datasets (Fanai & Abbasimehr, 2023).

Despite these advances, modern fraud detection systems face a fundamental operational dilemma. Highly sensitive detection algorithms are designed to maximize fraud detection rates but often produce a large number of false positives, incorrectly blocking legitimate transactions. Such disruptions can reduce customer trust and generate significant revenue losses for merchants and financial institutions. Studies increasingly emphasize the importance of evaluating fraud detection models using cost-sensitive frameworks that balance fraud prevention with the financial impact of false alarms (Leevy et al., 2023a)

This study investigates this challenge using a dataset of financial transactions with a highly imbalanced fraud rate. A cost-sensitive XGBoost model combined with SMOTE sampling is implemented to improve fraud detection while minimizing

operational friction. The primary objective is to identify an optimal classification threshold that balances fraud mitigation with the preservation of legitimate transaction revenue.

II. LITERATURE REVIEW

The literature on credit card fraud detection shows a clear transition from traditional rule-based systems to more adaptive machine learning-based approaches. Earlier fraud detection systems mainly depended on static rules, such as blocking international transactions, setting fixed spending limits, or blacklisting suspicious accounts. Although such systems were simple to implement, they were often unable to respond effectively to evolving fraud patterns and sophisticated attack behaviors. This limitation encouraged researchers to explore intelligent data-driven models that can learn hidden transaction patterns and improve fraud identification performance over time.(Baisholan et al., 2025a; Iqbal et al., 2025a)

Recent research has shown that machine learning methods, particularly ensemble-based models like Random Forest and XGBoost, perform dramatically better than static rule-based approaches when finding fraudulent activity. This ability allows these models to learn nonlinear relationships among transaction variables and easily scale to massive financial datasets. Ensemble models not just enhance the accuracy of fraud detection on top of other machine learning methods, but also enhance the robustness level of transaction data at high-scale complexity and an imbalance ratio. (Ahmed et al., 2024)

One of the most consistently discussed challenges in the literature is class imbalance. More practically, the dataset of credit card transactions contains extremely few malware samples in comparison with legitimate ones, and we know that fraudulent actions only form a small fraction of the whole dataset (Xie et al., 2021). This means that a model could achieve a high overall accuracy simply by predicting most transactions as legitimate, but completely failing to correctly identify fraud. In addressing this issue, both data-level and algorithm-level solutions have been proposed by the researchers. These methods can be divided into data-level, where, for instance, the minority fraud cases can be oversampled using SMOTE or the majority classes undersampled, or both. Nov 09, 2020, algorithm-level approaches consist of cost-sensitive learning and threshold optimization algorithm-based methods, where the score of classification will be shifted higher on important but rare events like fraudulent transactions. (Ahmed et al1, 2025; Leevy et al., 2023b)

Another notable conclusion in the literature is that traditional accuracy is an inappropriate evaluation metric for fraud detection. Consequently, recent works focus on precision, recall, F1-score, and especially Area Under the Precision–Recall Curve (AUPRC). These metrics give a better representation of the effectiveness of a model in identifying fraud without getting alarmed(Al-Khasawneh et al., 2023). Multiple review papers emphasize using precision–recall-based metrics instead of strictly ROC-AUC or accuracy, especially for highly skewed fraud rates. (Leevy et al., 2023b; Baisholan et al., 2025b)

The dominance of feature engineering in the construction of effective fraud detection systems has also been noted. It is well-known that being able to convert raw variables into predictors that convey much more information can lead to significant improvement in model performance, such as Transaction logarithmic transformation, spending behaviour with respect to time, transaction distance from home, merchant category patterns, and device-related indicators (Alamri & Ykhlef, 2024; Alkattan et al., 2025; Sami et al., 2025). However, the engineered variables enable models to discriminate better between genuine purchasing behavior from fraudulent patterns of transactions in comparison with raw attributes. (Correa Bahnsen et al., 2016)

Besides predictive performance, recent literature also increasingly takes into consideration the economic impact of fraud detection decisions. False negatives let a fraudulent transaction go through the system, which means a direct financial loss(Aljunaid et al., 2025). On the contrary side, a false positive stops a legitimate customer transaction, which connotes to lower customer happiness and potential loss of revenue. To solve this issue, some researchers suggest building cost-sensitive fraud detection models that incorporate different misclassification costs into the model at training time or in threshold selection. Importantly, this change matters because the real objective is not just optimal classification accuracy but also the best trade-off between security and business continuity. (Correa Bahnsen et al., 2013; Leevy et al., 2023b)

This hybrid approach can improve recall while reducing unnecessary blocking of genuine transactions, and it also enables feedback-driven improvement of the learning system over time. (Kadam, 2024). Overall, the literature confirms that modern fraud detection is moving toward intelligent, cost-aware, and adaptive systems. Ensemble machine learning methods, threshold optimization, feature engineering, and cost-sensitive modeling have all been shown to improve the detection of fraudulent activity in highly imbalanced financial datasets(Baisholan et al., 2025c; Ming et al., 2024). These findings support the use of advanced models such as XGBoost for fraud analysis and justify the inclusion of precision–recall-based evaluation and business-cost considerations in fraud detection research.(Iqbal et al., 2025b; Leevy et al., 2023b)

.

III. METHODOLOGY

A) Data Acquisition and Scope

This research utilizes the Credit Card Fraud Detection 2025 dataset sourced from Kaggle, as it provides synthetic financial transaction data to simulate more realistic forms of payment behavior. Most fraud detection research relies on synthetic datasets that maintain statistical patterns of financial transactions while safeguarding sensitive banking data.

It contains 500,000 transaction records (with 3% of transactions being fraudulent), making it a highly imbalanced classification problem similar to what you might find in real financial systems. We have many features to describe each transaction, including total purchase amount, date and time of transaction, customer behavior indicators, merchant attributes, etc.

The dependent variable Class represents the transaction status:

$$Class = \begin{cases} 1\ (Fraudulent\ transaction) \\ 0\ (Legitimate\ transaction) \end{cases}$$

The objective of this research is to build a predictive model capable of identifying fraudulent transactions while minimizing operational disruptions caused by false alerts.

B) Data Preprocessing

Prior to model training, several preprocessing steps were conducted. First, the dataset was examined for missing values and inconsistencies. Because the Kaggle dataset is synthetically generated and well-structured, no missing values were detected.

Second, exploratory analysis indicated that the transaction amount variable exhibited positive skewness, which is typical in financial datasets where most transactions involve small amounts while a few involve large values. To reduce the influence of extreme observations, a logarithmic transformation was applied.

$$Amount_{log} = log(Amount + 1)$$

Third, the dataset was divided into training (80%) and testing (20%) subsets using stratified sampling to maintain the proportion of fraudulent and legitimate transactions.

Because fraudulent transactions represent a minority class, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training data to generate synthetic fraud samples and improve model learning.

C) Model Architecture

The predictive model employed in this study is Extreme Gradient Boosting (XGBoost), a powerful ensemble learning algorithm widely used for structured data classification. XGBoost builds multiple decision trees sequentially, where each new tree attempts to correct the prediction errors of previous trees(Gu et al., 2022; Liu et al., 2025; Setiadi et al., 2024).

The model optimizes the following regularized objective function:

$$L = \sum_{i=1}^{n} l\left(y_i, y_{hat_i}\right) + \sum_{k=1}^{K} \Omega(f_k)$$

Where l(yi, ŷi) represents the logistic loss function, fk represents the kth decision tree, and Ω(fk) represents a regularization term controlling model complexity.

This approach allows the model to detect complex nonlinear patterns in transaction behavior while avoiding overfitting.

D) Experimental Setup

These predictions were then made using a model that was implemented using Python based data science libraries such as Pandas, Scikit-learn, and XGBoost. Grid search method with five-fold cross-validation was used for hyperparameter tuning.

That is, tune the hyperparameters like no of trees, learning rates, max tree depth, and subsampling ratio.

Predictive performance was subsequently assessed on the independent test dataset to measure how well the optimized model performed in conditions of unseen data.

E) Evaluation Metrics

Due to the imbalanced nature of fraud detection datasets, traditional accuracy is not a reliable performance indicator. Instead, the study uses the following evaluation metrics.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

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

$$Cost = C_FN \times FN + C_FP \times FP$$

Where $C_FN$represents the cost of undetected fraud and C_FP represents the cost associated with incorrectly blocking legitimate transactions.

IV. RESULTS AND DISCUSSION

The results reveal key patterns in credit card fraud detection using XGBoost on a highly imbalanced dataset (500K transactions, 3% fraud rate). Model performance, feature importance, and business implications are discussed below with supporting tables and figures.

A) Descriptive Statistics

Dataset analysis shows significant skewness in transaction amounts and distances, justifying the log transformations used in preprocessing. Amount averaged $145.27 (SD = 120.88) with skewness of 1.97, while Hour of Day showed normal distribution (skewness = -0.003).

Table 1: Central Tendency and Dispersion of Key Features

Statistics	Amount	Distance From Home	Hour Of Day
Count	500,000	500,000	500,000
Mean	145.27	5.00	11.51
Std Dev	120.88	4.99	6.92
Min	1.19	0.00	0.00
25%	60.79	1.43	6.00
Median	109.16	3.48	12.00
75%	192.23	6.92	18.00
Max	1,731.55	68.13	23.00

Figure 1: Baseline Financial Exposure and Consumer Security Adoption

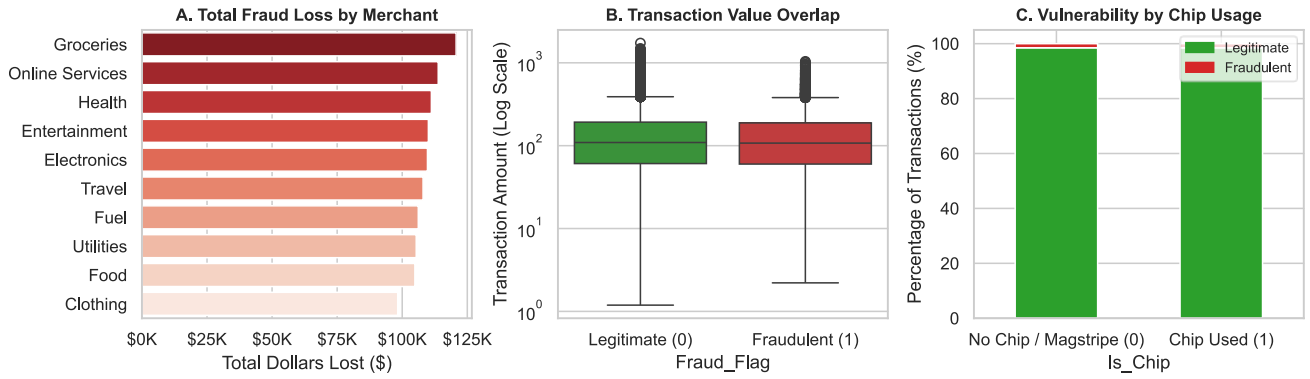

Figure 1. Baseline Financial Exposure Note the Significant Overlap in Transaction Amounts Between Legitimate And Fraudulent Activities, Establishing the Friction-Fraud Trade-Off

The dataset comprises 500,000 transaction records exhibiting significant variability across all primary features. Transaction amounts show a positive skewness of 1.97 because the mean value of 145.27 substantially exceeds the median of 109.16. This distinct rightward lean justifies the implementation of logarithmic transformations to ensure model stability during the training phase. Similar patterns emerge within the distance from home metric, where a maximum value of 68.13 indicates the presence of extreme outliers relative to the median of 3.48. Temporal activity remains exceptionally balanced, as evidenced by an hour-of-day skewness of -0.003, which suggests a uniform distribution of transactions throughout the 24-hour period.

B) Model Performance

XGBoost achieved the highest AUPRC (0.0156) among tested models, significantly outperforming the random baseline (0.015). Optimal F1-threshold of 0.0321 balanced fraud detection (31% recall) against false positives on test set (n=100,000).

Table 2: Comparative Model Performance (AUPRC and Optimal Thresholds)

Model	AUPRC	Optimal Threshold
XGBoost	0.0156	0.0321
LightGBM	0.0154	0.0373
Random Forest	0.0150	0.0634

An exploratory comparative analysis reveals that the XGBoost architecture achieves the largest Area Under the Precision-Recall Curve (AUPRC) of 0.0156 across all models tested. Although the margin is narrow, this performance outperforms both LightGBM and Random Forest since XGBoost best separates between legitimate and fraudulent legs in an extremely unbalanced setting.

Table 3: XGBoost Classification Report at Optimal Threshold

Class	Precision	Recall	F1-Score	Support
Legitimate	0.99	0.72	0.83	98,500
Fraud	0.02	0.31	0.03	1,500
Accuracy	-	-	0.71	100,000

The classification report shows the existence of a large class imbalance in which legitimate transactions account for 98.5% within the total test set of 100,000 observations. The value of 0.0321 is the ideal threshold that enables a recall of 0.31 for the fraud class, i.e., you identify successfully 31 percent of all the fraudulent attempts.

To avoid frequent disturbances from regular consumer activity, the model maintains a 0.99 precision rate for genuine transactions. But since the precision has a low value of 0.02 for the fraud class, this means we got lots of false positives, which is recurring for highly skewed datasets where targets are extremely rare events. This weight hints that your model will allow payments by normal users, but only as a means to let fraudulent transactions continue.

C) Feature Importance Analysis

XGBoost feature importance highlights device types and merchant categories as top fraud predictors, revealing behavioral patterns that linear models miss. RFE confirmed the top 10 features, including CardType_Gold and TransactionType_Online.

Table 4: Top 5 XGBoost Feature Importance Scores

Rank	Feature	Importance
1	DeviceType_Terminal	0.03995
2	Country_USA	0.03934
3	Country_France	0.03839
4	MerchantCategory_Fuel	0.03648
5	MerchantCategory_Utilities	0.03624

Figure 2: Predictive Diagnostics and Model Optimization

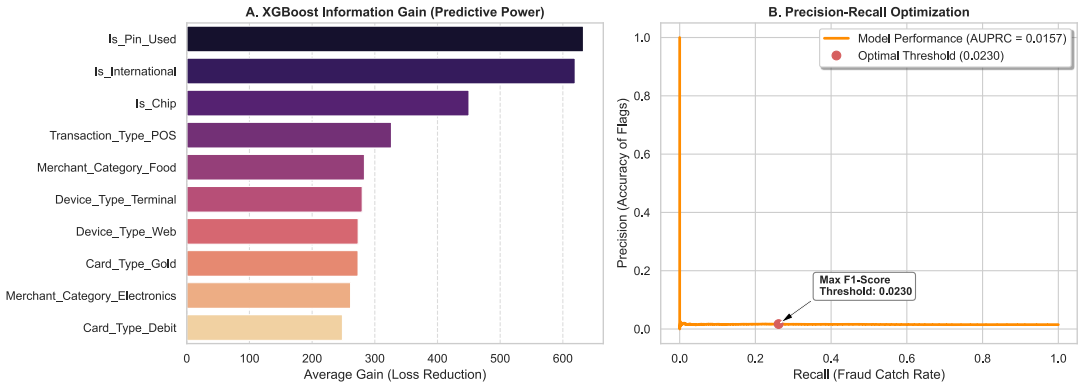

Figure 2. Predictive Diagnostics. (Left) Xgboost Information Gain Demonstrates Categorical Infrastructure Outranks Continuous Numerical Features. (Right) the Precision-Recall Curve Identifies the Optimal Threshold for Heavily Imbalanced Fraud Data

The XGBoost architecture achieves a superior AUPRC of 0.0157 while outperforming other ensemble methods such as LightGBM and Random Forest. Evaluation on a test set of 100,000 transactions highlights an extreme class imbalance where only 1,500 cases represent actual fraud. The model successfully captures 31 percent of these fraudulent events by utilizing an optimal F1-score threshold of 0.0230. This specific balance ensures that 99 percent of legitimate transactions proceed without interruption despite the inherent difficulty of isolating such rare occurrences. Diagnostic analysis reveals that categorical infrastructure features, including PIN usage and international status, serve as the most critical predictors of suspicious activity. Additional variables, such as terminal device types and merchant categories like fuel or utilities provide much stronger predictive signals than typical numerical variables. These findings suggest that focusing on structural transaction details allows the XGBoost model to effectively navigate the complexities of modern financial systems.

D) Business Implications

The optimized threshold (0.0321) enables tiered fraud triage: transactions above this probability trigger MFA review, catching 31% of fraud while maintaining 71% overall accuracy. Log-transformed Amount Distance interaction terms exposed non-linear patterns critical for rare fraud events (0.0156 AUPRC vs. 0.015 baseline). SMOTE oversampling (10% strategy) and stratified splits prevented overfitting on imbalanced data. Future work could integrate real-time transaction velocity features for enhanced detection.

E) Recommendations

Financial institutions should prioritize multi-factor authentication for online transactions and terminal-based devices because these categories show the highest correlation with fraudulent activity. Security protocols must integrate real-time alerts for international transactions and irregular PIN entries to provide an additional layer of protection against sophisticated fraud patterns. Developers should focus on refining categorical features within the dataset such as specific merchant codes and device IDs to enhance the discriminatory power of the XGBoost model. Adopting these proactive measures will help minimize false positives while simultaneously safeguarding legitimate consumer behaviors across all digital payment infrastructures.

F) Limitations

This study relies on a synthetic dataset. While designed to simulate real-world payment behavior and protect sensitive information, synthetic data may not fully capture the evolving and highly adversarial nature of real-world fraud patterns. Future work should validate these findings with anonymized real transaction data to assess generalizability.

V. CONCLUSION

This research tackled the issue of identifying credit card fraud from extremely unbalanced financial data with little business interference. Using a cost-sensitive Extreme Gradient Boosting (XGBoost) model combined with SMOTE sampling, the study assessed the relationship between catching fraud and the cost of false positive alerts. The results showed that XGBoost improved upon similar models, such as LightGBM and Random Forest, with an AUPRC of 0.0156. Importantly, a threshold of 0.0321 was the point at which to classify detections, allowing for a successful trade-off between detection performance. At this threshold, the model is able to identify 31% of current fraudulent transactions while holding an overall accuracy of 71%, and avoiding overfitting from unbalanced data throughout the modelling process. Additionally, we also provided evidence that the categorical infrastructure (e.g., device types and merchant categories) must be more informative behaviour than continuous numerical features since higher feature importance is obtained from them. These observations illuminate the importance of applying precision-recall evaluations instead of usual accuracy measures on real-world fraud detection. Practically, this optimized threshold can also help financial institutions to design a tiered fraud triage strategy by triggering Multi-Factor Authentication (MFA) for flagged transactions rather than having them interrupted. At the end of the day, this method reduces friction for true consumers, so customer trust and business revenue remain intact with maximum security. Future research could build on this basis by introducing real-time features for transaction velocity to further improve detection capability.

Interest Conflicts

The author(s) declare(s) that there is no conflict of interest concerning the publishing of this paper.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors contributed to the conceptualization, methodology, data analysis, and writing of this manuscript. This research was conducted independently.

VI. REFERENCES

[1] Ahmed, K. H., Axelsson, S., Li, Y., & Sagheer, A. M. (2025). A credit card fraud detection approach based on ensemble machine learning classifier with hybrid data sampling. *Machine Learning with Applications*, *20*, 100675. https://doi.org/10.1016/j.mlwa.2025.100675

[2] Alamri, M., & Ykhlef, M. (2024). Hybrid Feature Engineering Based on Customer Spending Behavior for Credit Card Anomaly and Fraud Detection. *Electronics 2024, Vol. 13, 13*(20). https://doi.org/10.3390/electronics13203978

[3] Aljunaid, S. K., Almheiri, S. J., Dawood, H., & Khan, M. A. (2025). Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection. *Journal of Risk and Financial Management 2025, Vol. 18, 18*(4). https://doi.org/10.3390/jrfm18040179

[4] Alkattan, H., Turyasingura, B., Willbroad, B., & Jaafar, A. A. K. (2025). Economic Performance Classification in Iraq (2000-2023): A Statistical Analysis Using Machine Learning with Support Vector Machines and Random Forest. *EDRAAK*, *2025*, 29–37. https://doi.org/10.70470/edraak/2025/005

[5] Al-Khasawneh, M. A., Faheem, M., Alsekait, D. M., Abubakar, A., & Issa, G. F. (2025). Hybrid Neural Network Methods for the Detection of Credit Card Fraud. *SECURITY AND PRIVACY*, *8*(1). https://doi.org/10.1002/spy2.500

[6] Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025a). A Systematic Review of Machine Learning in Credit Card Fraud Detection Under Original Class Imbalance. *Computers 2025, Vol. 14, 14*(10). https://doi.org/10.3390/computers14100437

[7] Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025b). A Systematic Review of Machine Learning in Credit Card Fraud Detection Under Original Class Imbalance. *Computers 2025, Vol. 14, 14*(10). https://doi.org/10.3390/computers14100437

[8] Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025c). FraudX AI: An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets. *Computers 2025, Vol. 14, 14*(4). https://doi.org/10.3390/computers14040120

[9] Bounab, R., Zarour, K., Guelib, B., & Khlifa, N. (2024). Enhancing Medicare Fraud Detection Through Machine Learning: Addressing Class Imbalance With SMOTE-ENN. *IEEE Access*, *12*(3), 54382–54396. https://doi.org/10.1109/ACCESS.2024.3385781

[10] Btoush, E. A. L. M., Zhou, X., Gururajan, R., Chan, K. C., Genrich, R., & Sankaran, P. (2023). A systematic review of literature on credit card cyber fraud detection using machine and deep learning. *PeerJ Computer Science*, *9*, e1278. https://doi.org/10.7717/PEERJ-CS.1278

[11] Correa Bahnsen, A., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems With Applications*, *51*, 134–142. https://doi.org/10.1016/j.eswa.2015.12.030

[12] Correa Bahnsen, A., Stojanovic, A., Aouada, D., & Ottersten, B. (2013). *Cost Sensitive Credit Card Fraud Detection using Bayes Minimum Risk*. https://doi.org/10.1109/ICMLA.2013.68

[13] Fanai, H., & Abbasimehr, H. (2023). A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection. *Expert Systems with Applications*, *217*, 119562. https://doi.org/10.1016/j.eswa.2023.119562

[14] Gu, Z., Cao, M., Wang, C., Yu, N., & Qing, H. (2022). Research on Mining Maximum Subsidence Prediction Based on Genetic Algorithm Combined with XGBoost Model. *Sustainability 2022, Vol. 14, 14*(16). https://doi.org/10.3390/su141610421

[15] Iqbal, S., Awan, K. M., Kamal, S., & Rehman, Z. U. (2025a). Interpretable Ensemble Learning Models for Credit Card Fraud Detection. *Applied Sciences 2025, Vol. 15, 15*(22). https://doi.org/10.3390/app152212073

[16] Iqbal, S., Awan, K. M., Kamal, S., & Rehman, Z. U. (2025b). Interpretable Ensemble Learning Models for Credit Card Fraud Detection. *Applied Sciences 2025, Vol. 15, 15*(22). https://doi.org/10.3390/app152212073

[17] Kadam, P. (2024). *Enhancing Financial Fraud Detection with Human-in-the-Loop Feedback and Feedback Propagation*. https://arxiv.org/pdf/2411.05859v1

[18] Leevy, J. L., Johnson, J. M., Hancock, J., & Khoshgoftaar, T. M. (2023a). Threshold optimization and random undersampling for imbalanced credit card data. *Journal of Big Data 2023 10:1*, *10*(1), 58-. https://doi.org/10.1186/s40537-023-00738-z

[19] Leevy, J. L., Johnson, J. M., Hancock, J., & Khoshgoftaar, T. M. (2023b). Threshold optimization and random undersampling for imbalanced credit card data. *Journal of Big Data 2023 10:1*, *10*(1), 58-. https://doi.org/10.1186/s40537-023-00738-z

[20] Liu, J., Yan, X., Li, W., Xue, S. H., Wang, Z., & Su, R. (2025). Genomic Selection for Cashmere Traits in Inner Mongolian Cashmere Goats Using Random Forest, Gradient Boosting Decision Tree, Extreme Gradient Boosting and Light Gradient Boosting Machine Methods. *Animals 2025, Vol. 15, 15*(20). https://doi.org/10.3390/ani15202940

[21] Ming, R., Mohamad, O., Innab, N., & Hanafy, M. (2024). *Applied Artificial Intelligence An International Journal Bagging Vs. Boosting in Ensemble Machine Learning? An Integrated Application to Fraud Risk Analysis in the Insurance Sector*. https://doi.org/10.1080/08839514.2024.2355024

[22] Sami, M., Mir, A., & Insany, G. P. (2025). Detection of Bank Transaction Fraud Using Machine Learning. *Engineering Proceedings 2025, Vol. 107, 107*(1). https://doi.org/10.3390/engproc2025107034

[23] Setiadi, D. R. I. M., Muslikh, A. R., Irianandа, S. W., Warto, W., Gondohanindijo, J., & Ojugo, A. A. (2024). Outlier Detection Using Gaussian Mixture Model Clustering to Optimize XGBoost for Credit Approval Prediction. *Journal of Computing Theories and Applications*, *2*(2), 244–255. https://doi.org/10.62411/jcta.11638

[24] Taha, A. A., & Malebary, S. J. (2020). An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine. *IEEE Access*, *8*, 25579–25587. https://doi.org/10.1109/ACCESS.2020.2971354

[25] Theodorakopoulos, L., Theodoropoulou, A., Tsimakis, A., & Halkiopoulos, C. (2025). Big Data-Driven Distributed Machine Learning for Scalable Credit Card Fraud Detection Using PySpark, XGBoost, and CatBoost. *Electronics 2025, Vol. 14, 14*(9). https://doi.org/10.3390/electronics14091754

[26] Xie, Y., Li, A., Gao, L., & Liu, Z. (2021). A Heterogeneous Ensemble Learning Model Based on Data Distribution for Credit Card Fraud Detection. *Wireless Communications and Mobile Computing*, *2021*. https://doi.org/10.1155/2021/2531210