JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Comparative Study of Machine Learning Model for Credit Card Fraud Detection.

1st Aniketh Chawdhari
MCA, School of IT
The ICFAI University, Sikkim
Gangtok, India
aniketzyan54325@gmail.com

2nd Pravesh Prasad
MCA, School of IT
The ICFAI University, Sikkim
Gangtok, India
kunaldx28@gmail.com

3rd Vivek Thapa
Assistant Professor , IT
The ICFAI University, Sikkim
Gangtok, India
vivekthapa@iusikkim.edu.in

***Abstract*— Credit card use has grown to become more widespread thus leading to higher incident rates of fraudulent transactions. The current standard detection systems encounter problems with both high numbers of wrong alarms and slow reactions that have negative effects on institution security protocols and user confidence levels. The research measures the performance of Logistic Regression, Decision Tree, Random Forest, and XGBoost in identifying fraud in real-time conditions. The evaluation of the models focuses on their accuracy rates and precision levels and recall measures and F1-score along with their computational speed and throughput. A hybrid ensemble model serves as a proposal to achieve performance- speed balance which makes it fit for deployment in real-world financial applications. The analysis studies the challenges between complex models and easy interpretation of financial systems and reveals their necessity for making transparent decisions. All results show different models lead at specific points since assessments vary with context which underlines the requirement for operation-dependent optimization. Research in development will employ adaptive learning to detect new types of evolving fraudulent patterns.**

Keywords— Credit card fraud, machine learning, real-time detection, classification models, fraud prevention, precision.

I. INTRODUCTION

Credit cards operate as financial instruments for authorized buying on borrowed money while delivering practicality along with cost-control options to consumers [1]. Banks together with financial institutions provide these cards to their users who can access borrowing funds that reach a specific spending limit and pay the funds back before interest accumulates within the defined payment period. Through credit cards users can obtain short-term loan benefits that include [2] cashback programs coupled with reward points and payment installment capabilities

Modern commerce extensively relies on credit cards which let users perform e-commerce transactions and obtain foreign products as well as obtain quick money during emergencies. The increase in credit card user numbers has resulted in major growth of credit card fraud along with unauthorized transactions and elevated cybersecurity dangers. In the digital economy of today users' financial information security along with their trust represent essential challenges that must be addressed[3],[4]. Credit card fraud increased in modern times continues as fraudsters endlessly develop new ways exploiting vulnerabilities in financial systems [5]. Various types of credit card fraud exist that use crooks steal physical credit cards for making unauthorized purchases. Hidden devices allow skimming scammers to harvest card information from ATMs and payment terminals[6]. By fake emails and websites users are tricked by phishing and social engineering into revealing their credit card information. Card- not-present fraud is a common online offense whereby thieves utilize stolen card information to make online purchases without physical card access[7]. Account takeover fraud is perpetrated by fraudsters who gain access to a victim's credit card account to change personal details and make unauthorized transactions. Application fraud is carried out by criminals who apply for new credit cards using stolen or fake identities. Chargeback fraud occurs when legitimate cardholders report genuine purchases as unauthorized in order to obtain a refund. Financial institutions and cybersecurity professionals collaborate to create sophisticated fraud detection and prevention techniques to safeguard consumers and businesses from financial loss as credit card fraud methods become increasingly sophisticated[8],[9].

Successful fraud detection protects financial operations from theft while protecting both money value and transaction safety. These systems use transaction pattern insights to separate valid transactions from improper ones. Organizations use two types of fraud detection approaches including predefined systems based on known fraud signatures together with algorithms that analyze previous transaction data to identify irregularities. Deep learning together with artificial intelligence technology enables dynamic detection systems that have scalable capabilities [11].

Organizations achieve successful fraud detection through which they block financial losses and guarantee security of consumer identities and earn customer trust for payment systems while fulfilling regulatory requirements. Such technological systems prevent wrong positive detections which allows the uninterrupted processing of genuine transactions [12].

A. GENERAL OVERFLOW OF CREDIT CARD FRAUD DETECTION

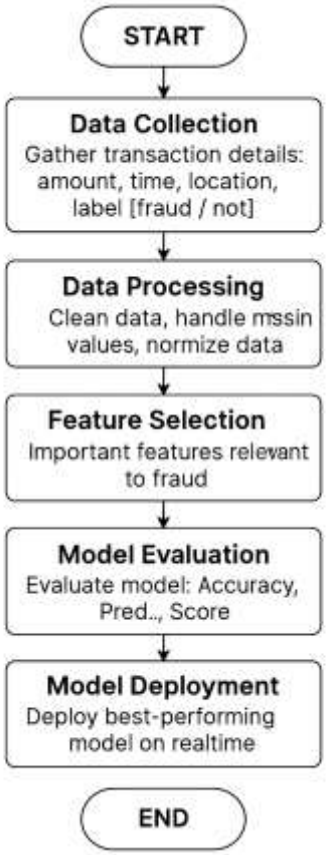

Fig 1: General Flow of Credit Card Fraud Detection

Steps for Credit Card Fraud Detection:

1.Data Collection

Gather transaction data (amount, time, location, status: fraud/not fraud).

2.Data Processing

Clean the data (fix errors, handle missing values).

Normalizethe data (scale numbers for model understanding).

3.Feature Selection

Select the most important features that influence fraud detection.

Remove irrelevant or redundant features.

4.. Model Selection

Choose machine learning models like Logistic Regression, Random Forest, Neural Networks, etc.

5. Model Training

Train the selected model on historical (past) transaction data.

6. Model Evaluation

Test and evaluate model performance using metrics like:

- Accuracy
- Precision
- F1-Score
- Recall

7. Model Deployment

Deploy the best-performing model into the real-world system for real-time transaction monitoring.

8.Continuous Monitoring and Improvement

Regularly monitor model performance.

Update and retrain the model with new data to keep it effective against evolving fraud techniques

B. OVERVIEW OF MACHINE LEARNING MODELS USED

Algorithm	Type	Strengths	Weaknesses
Random Forest (RF)	Ensemble (Tree-based)	Accurate, handles imbalance, avoids overfitting	Slow with large datasets
Logistic Regression	Linear Classifier	Fast, simple, interpretable	Poor with complex/non-linear data
XGBoost	Boosting (Tree-based)	High performance, avoids overfitting	Complex, requires tuning

Table 1: Overview of machine learning model used

C. TYPES OF CREDIT CARD FRAUD

- **Card Theft**

Physical stealing of a credit card to make unauthorized purchases.

- **Card Skimming**

Using hidden devices at ATMs or payment terminals to steal card information.

- **Phishing**

Fraudsters send fake emails, messages, or websites to trick users into giving up card details.

- **Card-Not-Present (CNP) Fraud**
  Fraudulent transactions made online or over the phone without physical access to the card.
- **Account Takeover**
  Hackers gain control of a credit card account by stealing login credentials and personal details.
- **Application Fraud**
  Criminals apply for a new credit card using stolen or fake identities.
- **Hargeback Fraud (Friendly Fraud)**
  A legitimate cardholder falsely claims a valid transaction was unauthorized to get a refund.
- **Counterfeit Cards**
  Fake cards created by cloning the information from a legitimate card's magnetic stripe.
- **Mail Theft**
  Stealing credit card statements or new cards from a victim's mailbox.
- **Data Breaches**
  Large-scale theft of credit card information from businesses or online platforms.

II. Related Work

The following are the investigations of the different scholarly articles for the credit card fraud detection.

Xuetong Niu et al. (2019): Proposed a fraud detection system using multiple machine learning algorithms. The system automatically identifies fraudulent transactions through training and testing on transaction data. Their approach emphasizes efficient transaction pattern recognition using historical transaction analysis. It demonstrates the effectiveness of combining multiple ML techniques. The study highlights the importance of data-driven insights in financial anomaly detection. The authors emphasize future work in incorporating real-time transaction monitoring.

Emmanuel Ileberi et al. (2022): Developed a fraud detection engine using genetic algorithms (GA) for feature selection. Compared classifiers like Decision Tree, Random Forest, Logistic Regression, ANN, and Naive Bayes, achieving high accuracy with European credit cardholder data. Their GA approach optimizes relevant feature selection, reducing noise. This improves detection speed and accuracy across several ML models. The paper also investigates class imbalance challenges. Future work could enhance GA-driven ML pipelines with deep learning.

Omkar Dabade et al. (2022): Designed a detection system using Random Forest, AdaBoost, and XGBoost combined through majority voting. Real-world banking data was used to validate the model's accuracy. The hybrid ensemble showed strong resilience to fraudulent data irregularities. Their method improves classification performance by leveraging multiple algorithm strengths. It outperforms standalone models on benchmark metrics. They recommend testing on more diverse datasets.

Dr. K. Maithili et al. (2023): Focused on machine learning-based fraud detection. Highlighted the limitations of traditional rule-based systems and used data preprocessing to improve model accuracy. Their approach integrates data balancing techniques to optimize training results. The study emphasizes enhancing feature extraction and transformation. It highlights model robustness in handling evolving fraud tactics. Future improvements could explore ensemble models.

Sreelekshmi S. & Shilpa A. (2023): Proposed a multi-algorithm fraud detection system that identifies fraudulent activities automatically using transaction data, enhancing detection with effective model training and testing. The study emphasizes the combination of classification and anomaly detection techniques. Model evaluation includes key metrics like sensitivity and specificity. The research supports scalable real-time fraud detection deployment. Future scope includes deep learning model experimentation.

Syeda Farjana Farabi et al. (2024): Evaluated nine ML algorithms including Logistic Regression, Decision Trees, Random Forest, Naive Bayes, KNN, and ANN. Measured performance using accuracy, F1-score, sensitivity, and specificity. This comparative analysis helped in identifying the best-performing model. The research stressed the role of precision in fraud identification. Ensemble models emerged as top contenders. Further enhancement can come from feature optimization strategies.

Yao Zou & Dawei Cheng (2025): Introduced a HOGRL model using mixture-of-expert attention and high-order graph learning. It outperformed baselines in fraud camouflage detection, recommending adaptive GNNs for future work. Their system uses advanced graph learning for relationship modelling. The attention mechanism prioritizes key features in detection. It effectively uncovers hidden fraud patterns. The study calls for continued research into graph-based fraud solutions.

Mir Mohtasam Hossain Sisan et al. (2025): Studied ML-based real-time fraud detection using supervised and unsupervised methods. Suggested integrating AI identity systems with blockchain for secure financial systems. Their framework evaluates transaction legitimacy on-the-fly. This reduces decision latency in online payments. Blockchain integration offers added transparency and traceability. Their study promotes fusion of AI and cybersecurity techniques.

Angel Jones & Marwan Omar (2025): Employed the LOF algorithm on unbalanced data for anomaly detection. Recommended further work on threshold tuning and integrating LOF with other ML methods. Their preprocessing pipeline improves detection accuracy. LOF showed robustness against minority class suppression. Model tuning significantly affected false positive rates. Future work includes real-time LOF deployment.

Weddou Mohamedhen et al. (2025): Combined Federated Learning (FL), LSTM, and SMOTE for privacy-preserving, imbalanced data fraud detection. Suggested further tuning of FL parameters and enhancing privacy with differential privacy techniques. Their framework enables collaborative model training without data sharing. LSTM captured sequential transaction dependencies effectively. SMOTE balanced fraud class distribution. The approach promotes secure and accurate fraud systems

Btoush et al. (2025): Similar to Weddou's work, combined FL, LSTM, and SMOTE for effective fraud detection while preserving data privacy across financial institutions. Their system benefits from distributed intelligence. It ensures scalability and compliance with data protection laws. SMOTE further strengthened class balance. Their results recommend continuous model updates for evolving patterns.

Kibet & Tonui (2025): Compared CNNs, LSTMs, and Autoencoders for fraud detection. Used SMOTE to handle class imbalance and found CNN+LSTM outperform traditional models. Their deep learning models captured spatial and temporal transaction features. Results highlight generalization and robustness. The study also focused on minimizing false positives. Future work includes hybrid architectures with blockchain.

Ghosh Dastidar (2025): Proposed a context-aware fraud detection method using Neural Aggregate Generator (NAG) and GANs to generate synthetic data. Suggested using attention-based transformers in future work. The contextual approach improved fraud signature recognition. GANs enriched model learning with diverse data. The research promotes adaptive learning in fraud detection. Future extensions involve real-time transformer-based models.

Lossan Bonde & Abdoul Karim Bichanga (2025): Developed a hybrid model combining CNN, GRU, and MLP with SMOTE-ENN. Achieved 100% accuracy and recommended developing real-time fraud detection systems. CNN extracted spatial features while GRU analyzed sequences. MLP acted as the final classifier. The SMOTE-ENN preprocessing balanced data and improved learning. Authors call for improved computation and deployment capabilities.

Mniai Ayoub et al. (2025): Introduced GrCF, combining CBR and FRS with BGWO for better parameter tuning and feature selection. Demonstrated high speed and accuracy in detecting new fraud patterns. Granular computing handled complex feature sets efficiently. FRS filtered redundant features while BGWO optimized performance. The system dynamically learns evolving fraud behavior. It sets a foundation for real-time adaptive systems.

Ahmed Samer et al. (2025): Reviewed the GrCF model by Ayoub et al. and analyzed its effectiveness compared to traditional ML methods, focusing on its feature selection and hyperparameter optimization. Their evaluation validated GrCF's practical efficiency. The study highlights the importance of optimized parameter tuning. Compared with conventional methods, it showed better speed and reliability. The paper recommends expanding GrCF across varied fraud scenarios.

Xuetong Niu et al. (2019) A credit card fraud detection system which employs several machine learning algorithms constitutes the main proposal of this research. The system seeks automatic fraudulent transaction detection through transaction data analysis. Testing and training procedures help the system identify regular transactions from fraudulent ones effectively.

Table 1: Major Contributions of Research in the Field of credit card fraud detection

Author (s)	Year	Method/Focus	Outcome	Limitation
Xuetong Niu et al.	2019	Multiple ML Algorithms	Successfully detects fraud using combined models. Improves accuracy and automation.	May lack adaptability to evolving fraud techniques.
Omkar Dabade et al.	2022	RF, AdaBoost, XGBoost (Voting)	Ensemble methods improve fraud detection accuracy in real-world data.	Performance may drop on imbalanced datasets.
Emmanuel Ileberi et al.	2022	GA + DT, RF, LR, ANN, NB	Feature selection via GA boosts model performance using European dataset.	Limited validation on diverse geographies.
Dr. K. Maithili et al.	2023	ML with Data Preprocessing	Balancing and feature enhancement strengthens fraud detection.	Focuses mostly on preprocessing, less on model innovation.
Sreelekshmi S. & Shilpa A.	2023	Multiple ML Algorithms	Effective classification of fraud via supervised training/testing.	Lacks real-time application and hybrid techniques.
Syeda Farjana Farabi et al.	2024	LR, DT, RF, NB, KNN, ANN	Compared 9 ML models; RF and ensemble performed best.	Further model tuning and deeper feature engineering needed.
Yao Zou & Dawei Cheng	2025	HOGRL (Graph Learning)	Outperforms other models in camouflage fraud detection using graph learning.	Complexity in implementing adaptive GNN frameworks
Ahmed Samer et al.	2025	Review of GrCF Framework	Validates GrCF's superiority over traditional ML systems.	Lacks practical implementation data in diverse settings.
Mniai Ayoub et al.	2025	GrCF: CBR + FRS + BGWO	Uses granular computing and optimization for faster fraud detection.	May require more tuning on diverse datasets.

III.PROPOSED WORK(METHODOLOGY)

THE FOLLOWING ARE THE STEPS FOR SYSTEM DEVELOPMENT:

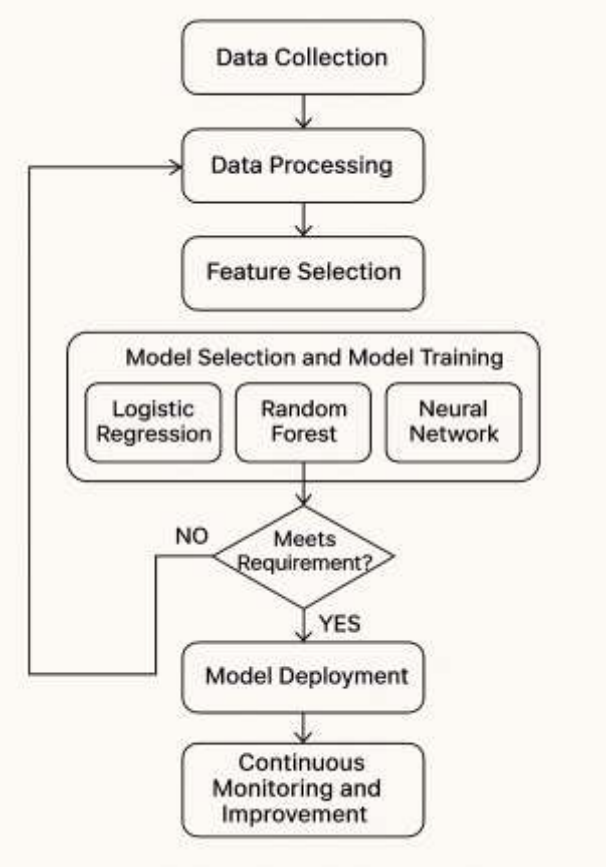

Figure 1: Driven Architecture for Real-Time Credit Card Fraud Detection

This diagram represents the process of detecting credit card fraud using machine learning. It starts with data collection, where transaction details such as amount, location, and time are gathered. Next, the data goes through processing, where

errors are fixed, missing values are handled, and numbers are scaled for better analysis.

After that, feature selection helps pick the most important factors that can indicate fraud while removing unnecessary ones. The system then moves to model selection, where different machine learning models, such as Logistic Regression, Random.

- **Data Collection:** Data was gathered from historical credit card transactions. Each transaction included features such as the transaction amount, time, location, and a label indicating whether it was fraudulent or not.
- **Data Processing:** Clean the data by removing any errors or missing information, and convert the data so that the model can understand it.
- **Feature Selection:** Select the most important features that help in predicting fraud and remove the ones that are not useful.
- **Model Selection and Training:** Choose a machine learning model like Logistic Regression, Random Forest, or Neural Network, and train it using the past transaction data.
- **Model Evaluation:** Once the model is trained, test how well it works using evaluation metrics like accuracy, precision, to choose the best one.
- **Model Deployment:** After selecting the best model, deploy it so it can start checking credit card transactions for fraud.

III. CHALLENGING IN CREDIT CARD FRAUD DETECTION

The Credit Card Fraud Detection (CCFD) systems have become much better with machine learning; there remind numerous issues to be address. Some of the significant ones are:

- **Imbalance Datasets**

  Fraudulent transactions are extremely rare compared to legitimate ones, making it difficult for fashions to learn meaningful fraud patterns without bias closer to majority classes.

- **Evolving Fraud processes**

Fraudsters continuously exchange their strategies, requiring detection fashions to be often up to date to stay powerful in opposition to new and sophisticated fraudschemes.

- **High fake Positives**

Many structures incorrectly flag legitimate transactions as fraudulent, main to consumer dissatisfaction and useless operational fees for financial establishments.

- **Real-Time Detection necessities**

reaching excessive accuracy while processing hundreds of thousands of transactions in real-time stays a technical and computational venture for fraud detection structures.

- **Statistics privateness and security worries**

gaining access to and sharing sensitive transaction statistics for model training and checking out is regularly restricted because of strict privateness rules, restricting model overall performance and pass-institutional collaboration

IV. FUTURE PROSPECTS IN CREDIT CARD FRAUD DETECTION(CCFD)

The credit card fraud detection has its challenges, the future is promising. As technology and research advanced, CCFD system will become smarter, faster, and more accurate. Here are some promising future possibilities:

a. Enhanced Feature Engineering

Specialists at Future Offline Fraud Detection Systems

Will Find Clusters of Abnormal Data in Static Datasets More Effectively Through Transaction Analysis.

b. Incorporation of Explainable AI (XAI) Banks need to know how their fraud detection models operate to approve systems that provide clear explanations about automatic actions.

c. Federated Learning for Offline Datasets Online detection systems build more secure and dependable fraud prediction models by letting multiplinstitutions pool their analytical knowledge.

d. Synthetic Data Generation for Model Training the mod When real fraudulent data is scarce GANs creates dependable fake records to help with offline model training

V CONCLUSION

The digital payment growth requires better security measures because criminals now exploit advanced means to access cardholder information online. This research analyzed various machine learning and deep learning standards along with multiple deep learning functional concepts. We analyzed Random Forest, Long Short-Term Memory (LSTM), Local Outlier Factor (LOF) and Convolutional Neural Network- GRU-Multi-Layer Perceptron combinations. The detection systems built with these systems spot fraudulent activities effectively and update themselves as new data arrives.
Our study confirms that multiple people working together produce superior outcomes while finding fraudulent transactions in mismatched data sets. Enhancing data defense plus system performance depend on how Federation Learning works with SMOTE for observation repetition and feature processing.

REFERENCES

[1] M. Ayoub, T. Abdelhamid, and J. Khalid, "Granular computing framework for credit card fraud detection," Alexandria Engineering Journal, vol. 121, no. February, pp. 387–401, 2025.

[2] D. Lunghi, Y. Molinghen, A. Simitsis, T. Lenaerts, and G. Bontempi, "FRAUD-RLA: A new reinforcement learning adversarial attack against credit card fraud detection," 2025, arXiv preprint arXiv:2502.02290.

[3] Y. Zou and D. Cheng, "Effective High-order Graph Representation Learning for Credit Card Fraud Detection," in Proc. IJCAI, pp. 7581–7589, 2024.

[4] A. S. I. Al-Dulaimi, I. R. Abdelmaksoud, S. Abdelrazek, and H. M. El-Bakry, "An intelligent credit card fraud detection model using data mining and ensemble learning," Edelweiss Applied Science and Technology, vol. 9, no. 2, pp. 1391–1405, 2025.

[5] M. M. H. Sizan et al., "Advanced Machine Learning Approaches for Credit Card Fraud Detection in the USA: A Comprehensive Analysis," Journal of Ecohumanism, vol. 4, no. 2, pp. 883–905, 2025.

[6] A. Jones and M. Omar, "Unveiling the Potential of Local Outlier Factor in Credit Card Fraud Detection," International Journal of Informatics, Information System and Computer Engineering, pp. 1– 13, 2026.

[7] W. Mohamedhen and M. Charfeddine, "Enhanced Credit Card Fraud Detection Using Federated Learning, LSTM Models, and the SMOTE Technique," in Proc. ICAART, vol. 3, pp. 368–375, 2025.

[8] E. Btoush, X. Zhou, R. Gururajan, K. C. Chan, and O. Alsodi, "Achieving Excellence in Cyber Fraud Detection: A Hybrid ML+DL Ensemble Approach for Credit Cards," Applied Sciences, vol. 15, no. 3, 2025.

[9] E. Oztemel and M. Isik, "A Systematic Review of Intelligent Systems and Analytic Applications in Credit Card Fraud Detection," Applied Sciences, vol. 15, no. 3, pp. 1–22, 2025.

[10] European Commission, "No Title," Volume 4, no. 1, pp. 1–23, 2016.

[11] S. Siddhish and C. Sekaran, "Identifying the ideal machine learning model for credit card fraud detection," Volume 12, no. 12, pp. 1–10, 2024.

[12] M. Sahu and R. Prasad, "Credit Card Fraud Detection: Survey and Discussion," Volume 3404, no. 1, pp. 1–6, 2025.

[13] L. Bonde, "Improving Credit Card Fraud Detection with Ensemble Deep Learning-Based Models: A Hybrid Approach Using SMOTE," 2025.

[14] H. Zheng, "Federated Learning-Based Credit Card Fraud Detection: A Comparative Analysis of Advanced Machine Learning Models," Paper No. 01022, pp. 1–6, 2025.

[15] M. Tayebi and S. El Kafhali, "Generative Modeling for Imbalanced Credit Card Fraud Transaction Detection," Journal of Cybersecurity and Privacy, vol. 5, no. 1, pp. 1–36, 2025.

[16] Y. R. K. Chakrabarti, "An intelligent framework for credit card fraud detection through data analytics," Volume 28, no. 1, pp. 139–149, 2025.

[17] [Y. Wu, L. Wang, H. Li, and J. Liu, "A Deep Learning Method of Credit Card Fraud Detection Based on Continuous-Coupled Neural Networks," Mathematics, vol. 13, no. 5, pp. 1–18, 2025.

[18] J. W. Alexander, "University of California, Los Angeles," Professional Geographer, vol. 9, no. 3, pp. 28–32, 1957.

[19] D. Salahudin-Mukeem and O. Ekundayo, "Hybrid Data Mining Technique for Credit Card Fraud Detection," Preprints, pp. 1–13, 2025.

[20] I. Y. Hafez, A. Y. Hafez, A. Saleh, A. A. Abd El-Mageed, and A. A. Abohany, "A systematic review of AI-enhanced techniques in credit card fraud detection," Journal of Big Data, vol. 12, no. 1, 2025.

[21] X. Fan and T. J. Boonen, "Machine Learning Algorithms for Credit Card Fraud Detection: Cost-Sensitive and Ensemble Learning Enhancements," unpublished manuscript, 2025.

[22] A. Hassan, A. Khader, J. Saudagar, S. Bhanja, and A. Das, "Data-Driven Methods for Credit Card Fraud Detection Using Machine Learning," Issue No. 3, 2021.

[23] I. P. Ojo and A. Tomy, "Explainable AI for credit card fraud detection: Bridging the gap between accuracy and interpretability," Volume 25, no. 2, pp. 1246–1256, 2025.

[24] A. Srivastava, A. Kundu, S. Sural, and A. Majumdar, "Credit CarFraud Detection Using Hidden Markov Model," IEEE Transactions on Dependable and Secure Computing, vol. 5, no. 1, pp. 37–48, 2008.

[25] V. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating Probability with Undersampling for Unbalanced Classification," in Proc. IEEE Symposium Series on Computational Intelligence, pp. 159–166, 2015.

[26] C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, "Transaction Aggregation as a Strategy for Credit Card Fraud Detection," Data Mining and Knowledge Discovery, vol. 18, no. 1, pp. 30–55, 2009.

[27] A. Bahnsen, D. Aouada, A. Stojanovic, and B. Ottersten, "Feature Engineering Strategies for Credit Card Fraud Detection," Expert Systems with Applications, vol. 51, pp. 134–142, 2016.

[28] [J. West and M. Bhattacharya, "Intelligent Financial Fraud Detection: A Comprehensive Review," Computers & Security, vol. 57, pp. 47–66, 2016.

[29] P. Phua, V. Lee, K. Smith, and R. Gayler, "A Comprehensive Survey of Data Mining-Based Fraud Detection Research," arXiv preprint arXiv:1009.6119, 2010.