10.48047/jocaaa.2024.33.04.50

# A Comprehensive Comparative Review of Machine Learning Techniques for Credit Card Fraud Detection: Algorithms, Class Imbalance Handling, and Research Challenges

**Andrews Felicita S[1] Dr. Grasha Jacob[2]**

1. Research Scholar(PART TIME), Register No:22221172332009, Department of Computer Science, Rani Anna Government College for Women, Gandhi Nagar, Tirunelveli,627008, Tamil Nadu, India. Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-12,Tamil Nadu, India.
2. Research Supervisor, Principal& Head, Department of Computer Science, Government Arts and Science College, Ottapidaram, Thoothukudi. Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-12, Tamil Nadu, India.

**Received: 21.10.2024 Accepted: 09.11.2024 Published: 30.12.2024**

## Abstract

The rapid expansion of digital payment systems, electronic commerce, and online banking has significantly increased the use of credit cards for financial transactions worldwide. While these technological advancements have improved transaction speed and customer convenience, they have also created new opportunities for fraudulent activities. Credit card fraud has become a major concern for financial institutions, merchants, and consumers due to its substantial economic impact and the increasing sophistication of cybercriminals. Conventional rule-based fraud detection systems are no longer sufficient because they rely on predefined rules and are unable to adapt effectively to continuously evolving fraud patterns. Consequently, machine learning (ML) and artificial intelligence (AI) techniques have emerged as promising approaches for developing intelligent, adaptive, and data-driven fraud detection systems.

This paper presents a comprehensive comparative review of research on credit card fraud detection published between 2016 and June 2024. The review systematically examines traditional machine learning algorithms, ensemble learning techniques, deep learning models, graph-based learning approaches, explainable artificial intelligence (XAI), and class imbalance handling methods employed for fraud detection. Furthermore, the study compares the strengths, limitations, computational complexity, interpretability, and practical applicability of these techniques in realworld financial environments. Special emphasis is placed on the challenges associated with highly imbalanced datasets, evolving fraud strategies, feature engineering, real-time detection, and model transparency.

The reviewed literature indicates that traditional machine learning algorithms such as Logistic Regression, Decision Trees, Support Vector Machines, Random Forest, Naive Bayes, and K-Nearest Neighbours provide reliable baseline performance for fraud detection. Ensemble learning methods, including Random Forest, XGBoost, LightGBM, and CatBoost, generally demonstrate improved predictive performance by combining multiple learning models. Deep learning techniques such as Multi-Layer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Autoencoders have shown considerable potential in learning complex transaction patterns and identifying previously unseen fraud. Additionally, Explainable Artificial Intelligence methods, particularly SHAP and LIME, have enhanced the transparency and interpretability of machine learning models, thereby supporting regulatory compliance and increasing user trust.

The review also examines commonly adopted data balancing techniques, including Random Oversampling, Random Undersampling, Synthetic Minority Over-sampling Technique (SMOTE), Adaptive Synthetic Sampling (ADASYN), and hybrid sampling methods such as SMOTEENN, which address the severe class imbalance inherent in credit card transaction datasets. Based on a comparative analysis of the existing literature, this paper identifies current research trends, highlights unresolved challenges, and discusses future research opportunities for developing accurate, scalable, interpretable, and real-time fraud detection systems. The findings presented in this review provide valuable insights for

10.48047/jocaaa.2024.33.04.50

researchers, practitioners, and financial institutions seeking to understand the current state of credit card fraud detection and to guide the development50 of next-generation intelligent fraud prevention systems.

Keywords: Credit Card Fraud Detection; Machine Learning; Ensemble Learning; Deep Learning; Graph Neural Networks;Explainable Artificial Intelligence; Class Imbalance; SMOTE; Financial Fraud Detection; Cybersecurity.

# 1. Introduction

## 1.1 Background

The rapid digital transformation of the global financial sector has fundamentally changed the way individuals and organizations perform monetary transactions. The widespread adoption of electronic commerce, mobile banking, digital wallets, contactless payments, and online financial services has led to an unprecedented increase in the use of credit cards. Credit cards have become one of the most popular payment instruments because they offer convenience, flexibility, global accessibility, and secure electronic transactions. As digital payment infrastructures continue to expand, millions of financial transactions are processed every day across banking networks worldwide.

Despite these technological advancements, the increasing dependence on electronic payment systems has also intensified the risk of financial fraud. Credit card fraud has emerged as one of the most significant challenges faced by financial institutions, payment service providers, merchants, and consumers. Fraudulent transactions result in substantial financial losses, operational costs, reputational damage, and reduced customer confidence. Fraudsters continuously develop sophisticated attack strategies, making fraud detection an increasingly complex task. Consequently, the development of intelligent and adaptive fraud detection systems has become a critical research area within machine learning, artificial intelligence, and financial cybersecurity.

Traditionally, financial institutions employed rule-based fraud detection systems that relied on predefined rules, expert knowledge, and manually designed thresholds to identify suspicious transactions. Although these systems are effective in detecting known fraud patterns, they often generate a high number of false alarms and struggle to identify newly emerging fraud strategies. Maintaining large collections of handcrafted rules is also time-consuming and requires continuous updates as fraud patterns evolve. These limitations have motivated researchers to investigate datadriven approaches capable of learning complex fraud patterns directly from historical transaction data.

Machine learning has emerged as one of the most effective approaches for automated fraud detection because it enables predictive models to learn from previously observed transaction behaviour. Unlike conventional rule-based systems, machine learning algorithms automatically discover hidden relationships among transaction attributes and adapt to changing fraud characteristics. Over the past decade, numerous supervised, unsupervised, and semi-supervised learning techniques have been applied to credit card fraud detection with the objective of improving classification accuracy while reducing false positive rates.

## 1.2 Problem Statement

Credit card fraud detection presents several technical challenges that limit the effectiveness of intelligent detection systems. The most significant challenge is the severe class imbalance present in real-world transaction datasets. In most publicly available datasets, fraudulent transactions constitute only a very small proportion of the total number of transactions. Consequently, machine learning algorithms tend to favour the majority class, producing high overall accuracy while failing to identify minority fraud cases effectively.

Another major challenge is the dynamic nature of fraudulent behaviour. Fraudsters continuously modify their attack strategies to bypass existing detection systems, causing previously trained models to become less effective over time. This phenomenon, commonly referred to as concept drift, requires fraud detection systems to adapt continuously to newly emerging fraud patterns.

Real-time transaction processing introduces additional constraints. Modern payment systems must analyse thousands of transactions every second while maintaining extremely low response times to avoid delaying legitimate customer payments. Therefore, fraud detection algorithms must achieve an appropriate balance between predictive accuracy and computational efficiency.

10.48047/jocaaa.2024.33.04.50

Furthermore, the increasing adoption of artificial intelligence in financial decision-making has created a growing demand for transparent and interpretable models. Financial institutions and regulatory authorities require explanations for automated decisions, particularly when legitimate customer transactions are declined. Consequently, model interpretability has become an essential consideration alongside predictive performance.

### 1.3 Motivation

The rapid evolution of machine learning has resulted in a wide variety of algorithms being applied to credit card fraud detection. Traditional machine learning algorithms, ensemble learning methods, deep learning architectures, graph-based learning approaches, and explainable artificial intelligence techniques have all demonstrated promising results under different experimental settings. However, individual studies often evaluate only a limited number of algorithms using specific datasets and evaluation criteria, making it difficult to determine the relative advantages and limitations of different approaches.

Although several review articles have examined particular categories of fraud detection techniques, there remains a need for a comprehensive comparative review that systematically analyses traditional machine learning, ensemble learning, deep learning, graph-based learning, explainable artificial intelligence, and class imbalance handling techniques within a unified framework. Such a review can provide researchers and practitioners with a broader understanding of existing methods, their practical applicability, and the challenges that remain unresolved.

### 1.4 Objectives of the Review

The primary objectives of this review are:

1. To provide a comprehensive overview of machine learning techniques used for credit card fraud detection.
2. To compare traditional machine learning, ensemble learning, deep learning, graph-based learning, and explainable artificial intelligence techniques.
3. To analyse commonly used class imbalance handling methods and their influence on fraud detection performance.
4. To identify the strengths and limitations of existing approaches based on published research.
5. To highlight current research challenges and recommend promising directions for future investigation.

### 1.5 Organization of the Paper

The remainder of this paper is organized as follows. Section 2 presents the background of credit card fraud detection and reviews the major categories of machine learning techniques reported in the literature. Section 3 provides a comparative analysis of the reviewed approaches, including datasets, evaluation metrics, advantages, and limitations. Section 4 discusses the major research gaps and outlines potential future research directions. Finally, Section 5 concludes the paper by summarizing the key findings and their implications for future research and practical fraud detection systems.

## 2. Literature Review

Credit card fraud detection has evolved from traditional rule-based systems to intelligent data-driven approaches capable of identifying complex and continuously changing fraud patterns. Advances in machine learning, ensemble learning, deep learning, and explainable artificial intelligence have significantly improved fraud detection performance. However, selecting an appropriate technique depends on several factors, including dataset characteristics, class imbalance, computational complexity, interpretability, and real-time processing requirements. This chapter reviews the major approaches employed for credit card fraud detection and critically discusses their strengths, limitations, and practical applicability.

10.48047/jocaaa.2024.33.04.50


Figure .1 Taxonomy of credit card Fraud Detection Techniques

2.1 Traditional Machine Learning Techniques

Traditional machine learning algorithms have formed the foundation of automated credit card fraud detection for more than a decade. These supervised learning methods classify transactions based on historical labelled data and remain attractive because they are computationally efficient, relatively easy to implement, and require limited computing resources. They are widely used as baseline models for evaluating newly proposed approaches.

Logistic Regression (LR) is one of the most frequently adopted classifiers for binary fraud detection. It estimates the probability that a transaction belongs to the fraudulent class using a logistic function. The model is simple, computationally efficient, and highly interpretable, making it suitable for realtime deployment. However, its ability to capture complex non-linear relationships is limited.

Decision Tree (DT) constructs hierarchical decision rules by recursively partitioning the feature space. The resulting model is easy to understand and provides transparent decision-making. Nevertheless, individual decision trees are prone to overfitting and may exhibit unstable performance when trained on noisy or highly imbalanced datasets.

Support Vector Machine (SVM) separates legitimate and fraudulent transactions by identifying an optimal hyperplane that maximizes the margin between classes. SVM performs well in highdimensional feature spaces but becomes computationally expensive for very large transaction datasets, limiting its applicability in large-scale payment systems.

K-Nearest Neighbours (KNN) classifies a transaction based on the majority class among its nearest neighbours. Although KNN is conceptually simple and does not require model training, prediction becomes computationally intensive as dataset size increases, making it less suitable for highthroughput financial applications.

Naive Bayes (NB) employs Bayes' theorem under the assumption of conditional independence among features. It offers fast training and prediction but may experience reduced accuracy when strong feature dependencies exist, which is common in financial transaction data.

Random Forest (RF) combines multiple decision trees through bootstrap aggregation to improve classification accuracy and reduce overfitting. Compared with individual decision trees, Random Forest generally provides better generalization and greater robustness to noisy data, making it one of the most widely adopted baseline models in fraud detection.
Overall, traditional machine learning algorithms remain valuable due to their simplicity, interpretability, and computational efficiency. However, their performance is often affected by severe class imbalance and their limited ability to model highly complex fraud patterns.

## 2.2 Ensemble Learning Techniques

Ensemble learning improves predictive performance by combining multiple base learners into a single robust classifier. Instead of relying on one model, ensemble methods aggregate the outputs of several models, thereby reducing variance, improving generalization, and enhancing resistance to overfitting.

Random Forest is the most widely adopted bagging technique for fraud detection because it combines numerous decision trees while maintaining good interpretability and stable performance. Boosting-

10.48047/jocaaa.2024.33.04.50

based algorithms, including AdaBoost, Gradient Boosting, XGBoost, LightGBM, and CatBoost, further improve prediction by sequentially correcting the errors of previous learners.

Among these approaches, XGBoost has received considerable attention due to its ability to model complex nonlinear relationships while maintaining high computational efficiency. LightGBM accelerates training through histogram-based learning and leaf-wise tree growth, whereas CatBoost effectively handles categorical variables and minimizes prediction bias.

Several comparative studies reported that ensemble learning consistently outperforms individual classifiers in terms of precision, recall, and overall classification performance, particularly when combined with imbalance handling techniques such as SMOTE or SMOTEENN. Nevertheless, increasing model complexity may reduce interpretability and increase computational requirements, especially in real-time financial systems.

## 2.3 Deep Learning Techniques

Deep learning has emerged as an effective approach for modelling complex transaction behaviour through automatic feature learning. Unlike conventional machine learning algorithms that rely heavily on manual feature engineering, deep neural networks learn hierarchical representations directly from transaction data.

Multi-Layer Perceptrons (MLPs) have been widely applied to binary fraud classification and provide improved representation learning compared with shallow models. Convolutional Neural Networks (CNNs) capture local feature interactions and have demonstrated promising performance in structured financial datasets. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are particularly effective for analysing sequential transaction histories because they preserve temporal dependencies across multiple transactions.

Autoencoders are frequently used as unsupervised anomaly detection models. They reconstruct normal transaction patterns and identify fraudulent transactions based on reconstruction error, making them particularly useful when labelled fraud samples are limited.

Although deep learning methods often achieve superior predictive performance, they require substantial computational resources, large training datasets, and longer training times. Their limited interpretability also presents challenges for practical deployment in highly regulated financial environments.

## 2.4 Explainable Artificial Intelligence

As machine learning models become increasingly complex, understanding their decision-making process has become an important requirement for financial institutions. Explainable Artificial Intelligence (XAI) improves model transparency by providing explanations that help analysts understand why a transaction has been classified as fraudulent.

Among the most widely adopted XAI techniques are Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP). LIME explains individual predictions by approximating the behaviour of a complex model locally, whereas SHAP quantifies the contribution of each feature using concepts derived from cooperative game theory.

Explainability improves regulatory compliance, increases user confidence, and supports fraud investigators during decision analysis. However, generating explanations introduces additional computational overhead, making real-time deployment challenging in high-volume transaction systems.

## 2.5 Class Imbalance Handling Techniques

One of the most significant challenges in credit card fraud detection is the severe imbalance between legitimate and fraudulent transactions. Conventional classifiers tend to favour the majority class, leading to poor fraud detection despite high overall accuracy. Consequently, numerous sampling techniques have been developed to improve minority class learning.

Random Oversampling increases minority class instances by duplicating existing observations, whereas Random Undersampling balances the dataset by reducing majority class samples. Although simple to implement, these approaches may introduce overfitting or information loss.

10.48047/jocaaa.2024.33.04.50

The Synthetic Minority Over-sampling Technique (SMOTE) addresses this limitation by generating synthetic minority class samples rather than exact duplicates. Adaptive Synthetic Sampling (ADASYN) further improves minority learning by generating additional synthetic observations in regions where classification is more difficult. Hybrid methods such as SMOTEENN combine oversampling with data cleaning to improve class separation and reduce noise.

The selection of an imbalance handling technique depends on dataset characteristics and the intended fraud detection objective. Most comparative studies indicate that hybrid sampling strategies generally achieve better recall while maintaining acceptable precision, particularly when integrated with ensemble learning algorithms

## 2.6 Summary of the Literature

The reviewed studies demonstrate a clear progression from traditional machine learning models toward more sophisticated ensemble and deep learning approaches. Traditional algorithms remain attractive because of their simplicity, computational efficiency, and interpretability, whereas ensemble learning generally provides improved predictive performance through model aggregation. Deep learning has expanded the capability of fraud detection systems to learn complex behavioural patterns but requires greater computational resources and larger datasets.50 Explainable Artificial Intelligence has addressed the growing demand for transparent decision-making, while class imbalance handling techniques have become essential for improving minority class detection. Despite these advances, challenges related to evolving fraud patterns, real-time processing, dataset diversity, and model interpretability continue to motivate further research in intelligent credit card fraud detection.

Table 1 Comparative study

| Re | Autho | Yea | Algori Techn | Data | Imbala Handl | Key Findings |
|---|---|---|---|---|---|---|
| 1 | Bahnsen et | 2016 | Feature Enginee ML | Real-wor CCFD | None | Demonstrated the importance of feature engineering for improving fraud detection performance. |
| 2 | Awoyemi ( | 2017 | LR, NB | Europear Card Dat | None | Random Forest outperformed conventional classifiers in fraud classification. |
| 3 | Srivastava | 2019 | Neural Networl XGBoo: | Kaggle | SMOTEEN | Hybrid learning improved fraud detection compared with individual models. |
| 4 | Wang et al | 2020 | LR, RF, | Kaggle | None | Random Forest achieved higher classification accuracy than LR and KNN. |
| 5 | Khatri et a | 2020 | SVM, R | Kaggle | SMOTE | Ensemble methods consistently produced better predictive performance. |
| 6 | Kumar et a | 2022 | ML Compar | Kaggle | ADASYN | Class balancing significantly improved fraud recall. |
| 7 | Singh et al | 2023 | Ensemb Learnin | Financial Dataset | SMOTE | Boosting algorithms demonstrated strong overall performance. |

# 3.Comparative Analysis of Credit Card Fraud Detection Techniques

The studies reviewed indicate that no single machine learning algorithm consistently outperforms all others under every experimental condition. Model performance depends on several factors, including dataset characteristics, class imbalance, feature engineering, computational complexity, interpretability, and deployment requirements. Consequently, selecting an appropriate fraud detection technique requires balancing predictive performance with operational constraints such as latency, scalability, and model transparency.

10.48047/jocaaa.2024.33.04.50

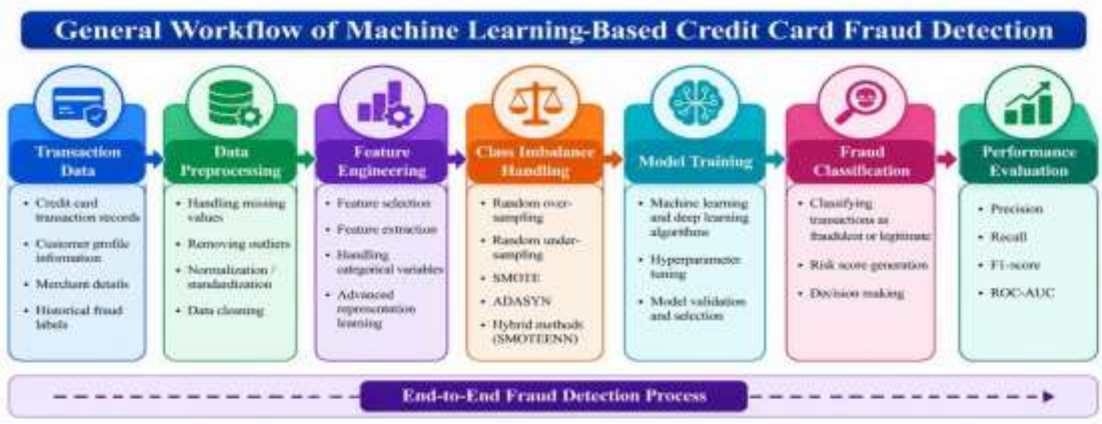


**Figure 2.** General Workflow of Machine Learning-Based Credit Card Fraud Detection

## 3.1 Comparative Analysis of Learning Paradigms

Traditional machine learning algorithms remain valuable because of their simplicity, low computational cost, and ease of implementation. Logistic Regression, Decision Trees, Naïve Bayes, Support Vector

Machines, and K-Nearest Neighbours provide reliable baseline performance50 and are suitable for applications where computational efficiency and model interpretability are primary concerns. However, their ability to capture complex nonlinear relationships is limited, particularly when fraud patterns continuously evolve.

Ensemble learning methods consistently demonstrate superior predictive performance by combining multiple weak or strong learners into a single robust classifier. Random Forest reduces overfitting through bootstrap aggregation, while boosting algorithms such as XGBoost, LightGBM, and CatBoost improve classification by sequentially correcting prediction errors. These methods generally achieve higher recall and Area Under the Receiver Operating Characteristic Curve (ROC-AUC) than individual machine learning algorithms, making them well suited for large-scale financial fraud detection.

Deep learning approaches have further enhanced fraud detection by automatically learning hierarchical feature representations from transaction data. Neural network architectures such as Multi-Layer Perceptrons, Convolutional Neural Networks, Long Short-Term Memory networks, and Autoencoders effectively capture complex behavioural patterns and temporal dependencies. Despite their high predictive capability, these models require larger datasets, greater computational resources, and longer training times. Their limited interpretability also presents practical challenges in regulated financial environments.

Explainable Artificial Intelligence does not replace predictive models but complements them by improving transparency. SHAP and LIME provide explanations that help analysts understand model predictions and support regulatory compliance. Although these techniques improve user confidence, explanation generation increases computational overhead, particularly in real-time applications.

## 3.2 Comparative Analysis of Class Imbalance Handling Techniques

Class imbalance remains one of the most critical challenges in credit card fraud detection because fraudulent transactions represent only a small fraction of total transactions. Without appropriate balancing strategies, most classifiers become biased toward legitimate transactions, resulting in poor fraud detection despite high overall accuracy.

Random Oversampling and Random Undersampling are simple balancing techniques that improve class distribution but may introduce overfitting or information loss. Synthetic Minority Over-sampling Technique (SMOTE) generates artificial minority class samples and has become one of the most widely adopted preprocessing methods in fraud detection research. Adaptive Synthetic Sampling (ADASYN) further improves minority learning by generating additional samples in difficult-to-classify regions.

Hybrid approaches such as SMOTEENN combine synthetic oversampling with noise removal, producing cleaner decision boundaries and improving minority class classification. Comparative studies generally

10.48047/jocaaa.2024.33.04.50

report that hybrid balancing techniques provide better recall than single sampling methods when integrated with ensemble learning algorithms.

## 3.3 Comparative Analysis of Evaluation Metrics

The literature demonstrates that overall classification accuracy alone is not an appropriate performance indicator for highly imbalanced fraud detection datasets. A model predicting all transactions as legitimate may achieve high accuracy while failing to detect fraudulent transactions.

Consequently, researchers commonly evaluate fraud detection systems using Precision, Recall, F1-score, ROC-AUC, and confusion matrices. Recall is particularly important because it measures the proportion of fraudulent transactions correctly identified, thereby minimizing financial losses. Precision evaluates the proportion of predicted frauds that are genuinely fraudulent, reducing unnecessary investigations. The F1-score provides a balanced measure of precision and recall, whereas ROC-AUC evaluates the overall discriminative ability of the classifier across different decision thresholds.

The selection of evaluation metrics should therefore reflect the operational objectives of financial institutions rather than relying solely on classification accuracy.

## 3.4 Comparative Analysis of Publicly Available Datasets

Most published studies utilise the Kaggle Credit Card Fraud Detection dataset due to its public availability, standardized format, and realistic class imbalance. This dataset contains 284,807 transactions, of which only 492 represent fraudulent activity. Its widespread adoption has facilitated objective comparison among machine learning algorithms.

Although publicly available datasets promote reproducible research, excessive reliance on a single dataset limits the generalizability of reported findings. Real-world banking environments differ considerably in transaction characteristics, customer behaviour, fraud patterns, and geographic distribution. Consequently, models demonstrating strong performance on50benchmark datasets may not necessarily perform equally well in operational financial systems.

Future research should therefore evaluate fraud detection algorithms using multiple datasets representing different financial institutions, transaction environments, and fraud scenarios.

## 3.5 Strengths and Limitations of Existing Approaches

Traditional machine learning methods provide fast prediction, low computational complexity, and high interpretability, making them suitable for resource-constrained environments. However, they often struggle to identify complex fraud patterns and are sensitive to severe class imbalance.

Ensemble learning methods generally achieve the best balance between predictive performance and computational efficiency. Their primary limitations include reduced interpretability and increased training complexity compared with simpler machine learning models.

Deep learning approaches effectively learn complex behavioural relationships and sequential transaction patterns without extensive feature engineering. Nevertheless, they require substantial computational resources, larger training datasets, and specialized hardware, making deployment more challenging.

Explainable Artificial Intelligence enhances transparency and regulatory compliance but introduces additional computational overhead during explanation generation. Similarly, data balancing techniques substantially improve minority class detection but require careful selection because inappropriate sampling may introduce synthetic bias or remove valuable transaction information.

## 3.6 Comparative Summary

Based on the reviewed literature, traditional machine learning algorithms remain suitable for baseline modelling and interpretable fraud detection systems. Ensemble learning methods consistently provide stronger predictive performance and represent the most practical choice for many real-world applications. Deep learning offers significant advantages for modelling complex transaction behaviour but requires greater computational resources and extensive training data. Explainable Artificial Intelligence improves model transparency, while class imbalance handling techniques remain essential for reliable fraud detection.

10.48047/jocaaa.2024.33.04.50

Rather than relying on a single algorithm, current research increasingly combines robust classifiers with effective data balancing and appropriate evaluation metrics to achieve reliable and scalable fraud detection systems capable of supporting modern digital payment environments.

Table 2 Comparative Analysis of Algorithms

| Technique | Advantages | Limitations | Computational Cost | Interpretability | Suitable Applications |
|---|---|---|---|---|---|
| Logistic Regression | Simple, fast, easy to interpret | Limited nonlinear modelling | Low | High | Baseline fraud detection |
| Decision Tree | Transparent decision rules | Overfitting | Low | High | Small datasets |
| SVM | Effective in highdimensional data | High computational complexity | High | Moderate | Medium-sized datasets |
| Random Forest | Robust and accurate | Reduced interpretability | Medium | Moderate | General fraud detection |
| XGBoost | High predictive performance | Parameter tuning required | Medium–High | Moderate | Large-scale fraud detection |
| LightGBM | Fast training | Sensitive to parameter settings | Medium | Moderate | Large financial datasets |
| CNN | Learns feature representations | Requires substantial training data | High | Low | Complex transaction data |
| LSTM | Captures temporal behaviour | Long training time | High | Low | Sequential transaction analysis |
| Autoencoder | Effective anomaly detection | Limited interpretability | High | Low | Unsupervised fraud detection |
| SHAP/LIME | Improves model transparency | Additional computational overhead | Medium | Very High | Decision explanation and auditing |

# 4. Research Gaps

Despite significant advancements in machine learning-based credit card fraud detection, several challenges remain unresolved. The literature published up to June 2024 demonstrates considerable progress in improving detection accuracy; however, many proposed solutions are evaluated under controlled experimental conditions and face limitations when deployed in real-world financial environments. The following research gaps were consistently identified across the reviewed studies.


Fig.3 Research Gaps and Future Research Directions

10.48047/jocaaa.2024.33.04.50

## 4.1 Limited Dataset Diversity

A large proportion of existing studies rely on the publicly available Kaggle Credit Card Fraud Detection dataset. While this dataset has become a standard benchmark for algorithm evaluation, its extensive use limits the generalizability of reported findings. Real-world banking systems differ in transaction volume, customer behaviour, fraud strategies, and regional payment patterns. Consequently, models trained and evaluated on a single benchmark dataset may not maintain the same performance across different financial institutions. Future research should incorporate multiple datasets and real-world transaction environments to improve model robustness and external validity.

## 4.2 Severe Class Imbalance

Although various sampling techniques such as SMOTE, ADASYN, and SMOTEENN have improved minority class detection, class imbalance remains a fundamental challenge in fraud detection. Many studies primarily optimize overall classification performance, while the impact on false positives and operational costs receives comparatively less attention. Developing methods that effectively balance fraud detection capability with acceptable false alarm rates remains an important research problem.

## 4.3 Limited Model Interpretability

The increasing adoption of ensemble and deep learning models has improved predictive performance but reduced model transparency. Financial institutions require explanations for automated decisions to support auditing, regulatory compliance, and customer trust. Existing explainability techniques such as SHAP and LIME improve interpretability; however, they often introduce additional computational overhead and may not be suitable for high-volume real-time transaction processing. More efficient explainable learning approaches are therefore needed.

## 4.4 Real-Time Deployment Challenges

Most studies evaluate fraud detection algorithms using offline datasets, whereas operational payment systems require continuous real-time transaction analysis. Practical deployment must satisfy strict latency requirements while maintaining high detection accuracy. Balancing computational efficiency with predictive performance remains a significant challenge, particularly for resource-intensive ensemble and deep learning models.

## 4.5 Adaptation to Evolving Fraud Patterns

Fraudulent behaviour continuously evolves as attackers develop new strategies to bypass existing detection systems. However, many published models are trained using static historical datasets and assume that fraud patterns remain unchanged over time. This limitation reduces long-term model effectiveness in dynamic financial environments. Adaptive learning strategies capable of responding to changing fraud behaviour remain an important area for future investigation.

10.48047/jocaaa.2024.33.04.50

### 4.6 Standardized Evaluation and Benchmarking

The reviewed studies employ different datasets, preprocessing methods, sampling strategies, and evaluation metrics, making direct comparison between algorithms difficult. Inconsistent experimental settings often lead to conflicting conclusions regarding algorithm performance. Establishing standardized evaluation protocols and benchmark datasets would improve the reproducibility and comparability of future research.

### 4.7 Summary

The literature indicates that future research should focus on improving dataset diversity, handling class imbalance more effectively, enhancing model interpretability, supporting real-time deployment, adapting to evolving fraud patterns, and establishing standardized evaluation frameworks. Addressing these challenges will contribute to the development of more accurate, reliable, and practical credit card fraud detection systems capable of meeting the demands of modern digital payment ecosystems. Table 3 Research Gaps and Future Research Direction

| Research | Current Limita | Future Research Direction |
|---|---|---|
| Dataset Divers | Heavy reliance on benchmark da | Evaluate models using diverse real-world financial datasets |
| Class Imbalanc | Existing sampling methods may | Develop adaptive and cost-aware balancing strategies |
| Model Interpre | Complex models operate as blacl | Design efficient explainable AI methods for financial applications |
| Real-Time Dep | High computational complexity l | Develop lightweight models with low inference latency |
| Evolving Frauc | Static models struggle with conc | Investigate adaptive and continuously updated learning systems |
| Standardized E | Different metrics and datasets hir | Establish common benchmarking protocols and evaluation frameworks |

## 5. Future Research Directions

The rapid evolution of digital payment systems and fraud techniques requires continuous improvement in credit card fraud detection models. Although existing machine learning approaches have achieved encouraging results, several opportunities remain for developing more accurate, efficient, and practical fraud detection systems. Based on the research gaps identified in the reviewed literature, the following future research directions are recommended.

### 5.1 Development of Adaptive Learning Models

Most existing fraud detection models are trained using historical transaction data and remain unchanged after deployment. However, fraud patterns continuously evolve as cybercriminals adopt new attack strategies. Future research should focus on adaptive learning models capable of updating their knowledge from newly available transaction data while maintaining stable predictive performance. Such approaches can improve the long-term effectiveness of fraud detection systems in dynamic financial environments.

### 5.2 Improved Handling of Class Imbalance

The highly imbalanced nature of credit card transaction data continues to be a major challenge. Although techniques such as SMOTE, ADASYN, and SMOTEENN have demonstrated improved minority class detection, no single approach performs consistently across different datasets. Future studies should investigate more robust balancing strategies and evaluate their effectiveness under diverse transaction distributions while minimizing false positive rates.

### 5.3 Explainable and Trustworthy Artificial Intelligence

The adoption of complex machine learning and deep learning models has increased the need for transparent decision-making. Financial institutions require understandable explanations to support auditing, regulatory compliance, and customer confidence. Future research should focus on developing lightweight explainable

10.48047/jocaaa.2024.33.04.50

artificial intelligence techniques that provide meaningful explanations with minimal computational overhead, enabling practical deployment in real-time payment systems.

### 5.4 Real-Time Fraud Detection

Modern electronic payment platforms process thousands of transactions every second, requiring fraud detection systems to produce accurate predictions within strict time constraints. Future research should emphasize computationally efficient algorithms that maintain high detection performance while reducing processing latency. Model optimization and efficient deployment strategies will play an important role in supporting real-time financial applications.

### 5.5 Evaluation Using Diverse Datasets

The majority of existing studies evaluate algorithms using a limited number of publicly available datasets. Although these datasets facilitate benchmarking, they may not adequately represent the diversity of real-world financial transactions. Future research should validate fraud detection models across multiple datasets collected from different banking environments, transaction types, and customer populations to improve model robustness and generalizability.

### 5.6 Integration of Multiple Learning Approaches

The reviewed literature indicates that different learning paradigms offer complementary advantages. Traditional machine learning methods provide simplicity and interpretability, ensemble learning improves predictive performance, deep learning captures complex behavioural patterns, and explainable artificial intelligence enhances transparency. Future research should investigate integrated frameworks that effectively combine these complementary capabilities while maintaining computational efficiency and practical applicability.

### 5.7 Standardized Evaluation Frameworks

Meaningful comparison among fraud detection techniques remains difficult because existing studies employ different preprocessing methods, sampling strategies, evaluation metrics, and experimental settings. Developing standardized benchmarking procedures and common evaluation protocols would improve reproducibility, facilitate objective comparison, and support the identification of genuinely effective fraud detection methods.

### 5.8 Summary

Future credit card fraud detection research should prioritize adaptive learning, effective class imbalance handling, efficient explainable artificial intelligence, real-time processing, evaluation across diverse datasets, integration of complementary learning approaches, and standardized experimental methodologies. Addressing these challenges will contribute to the development of intelligent, reliable, and scalable fraud detection systems capable of supporting the growing demands of digital financial services.

## 6. Conclusion

Credit card fraud continues to pose a significant challenge to the global financial industry due to the rapid growth of digital payment systems, online banking, and electronic commerce. The increasing volume of financial transactions and the evolving nature of fraudulent activities have made traditional rule-based detection systems insufficient for protecting modern payment infrastructures. Consequently, machine learning and artificial intelligence have become the primary technologies for developing intelligent and adaptive fraud detection systems.

This paper presented a comprehensive comparative review of credit card fraud detection techniques reported in the literature from 2016 to June 2024. The review examined the evolution of fraud detection approaches, including traditional machine learning algorithms, ensemble learning methods, deep learning models, explainable artificial intelligence techniques, and class imbalance handling strategies. Their working principles, strengths, limitations, and practical applicability were critically analyzed to provide a comprehensive understanding of current research developments.

The comparative analysis indicates that traditional machine learning algorithms remain valuable because of their simplicity, computational efficiency, and interpretability. However, their performance is often limited when dealing with complex fraud patterns and highly imbalanced datasets. Ensemble learning techniques generally provide improved predictive performance by combining multiple learning models, making them suitable for many real-world fraud detection applications. Deep learning approaches have demonstrated the capability to

10.48047/jocaaa.2024.33.04.50

learn complex behavioural patterns and temporal relationships, although their computational requirements and limited interpretability remain important challenges. In addition, explainable artificial intelligence has emerged as50an essential component for improving transparency and supporting regulatory compliance, while class imbalance handling techniques continue to play a critical role in enhancing fraud detection performance.

The review also identified several research challenges that require further investigation. These include limited dataset diversity, severe class imbalance, model interpretability, real-time deployment constraints, adaptation to evolving fraud patterns, and the lack of standardized evaluation frameworks. Addressing these challenges is essential for developing fraud detection systems that are accurate, scalable, reliable, and suitable for practical deployment in financial institutions.

Overall, the findings of this review demonstrate that no single technique is universally optimal for all fraud detection scenarios. The effectiveness of a fraud detection system depends on selecting appropriate algorithms, preprocessing methods, evaluation metrics, and deployment strategies based on the characteristics of the application environment. Future research should continue to focus on developing adaptive, explainable, and computationally efficient fraud detection models while promoting standardized evaluation practices and validation across diverse real-world datasets. It is expected that continued advances in machine learning and artificial intelligence will further strengthen the capability of financial institutions to detect and prevent fraudulent transactions, thereby improving the security and reliability of modern digital payment systems.

# References

[1] A. C. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten, "Feature engineering strategies for credit card fraud detection," Expert Systems with Applications, vol. 51, pp. 134-142, 2016.

[2] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in Proc. Int. Conf. Comput. Netw. Inf. Secur., pp. 684689, 2017.

[3] A. Srivastava and R. Gupta, "Credit card fraud detection using neural network and XGBoost," in Proc. 4th Int. Conf. Mach. Learn. Data Sci., pp. 45-52, 2019.

[4] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, "Credit card fraud detection - Machine learning methods," in Proc. 18th Int. Symp. INFOTECH-JAHORINA, pp. 1-5, 2019.

[5] S. Wang, Y. Li, and X. Zhang, "Credit card fraud detection using logistic regression, random forest and KNN," Journal of Financial Data Science, vol. 2, no. 3, pp. 78-89, 2020.

[6] S. Khatri, A. Arora, and A. P. Agrawal, "Supervised machine learning algorithms for credit card fraud detection: A comparison," in Proc. 10th Int. Conf. Cloud Comput. Data Sci. Eng., pp. 147-150, 2020.

[7] T. F. Cosoiu, "Credit card fraud detection using logistic regression," Journal of Applied Computational Science and Mathematics, vol. 14, no. 2, pp. 45-50, 2020.

[8] H. Farvaresh and M. M. Sepehri, "A data mining framework for detecting credit card fraud by considering fraud types," Computers & Mathematics with Applications, vol. 62, no. 3, pp. 1340-1352, 2021.

[9] N. S. Kumar, G. S. Rao, and V. K. Singh, "An empirical study of imbalanced classification for credit card fraud detection," International Journal of Computer Applications, vol. 182, no. 4, pp. 12-18, 2022.

[10] L. Zheng, Z. Li, and Y. Wang, "Real-time credit card fraud detection using LightGBM and feature engineering," in Proc. IEEE Int. Conf. Big Data, pp. 3012-3017, 2022.

[11] R. Singh, P. Sharma, and A. Verma, "Fraud detection in credit card transactions using ensemble methods," Procedia Computer Science, vol. 218, pp. 1245-1252, 2023.

[12] F. Niu, X. Chen, and Y. Wang, "SMOTE-GBDT: An imbalanced credit card fraud detection method based on SMOTE and gradient boosting decision tree," IEEE Access, vol. 11, pp. 23451-23460, 2023.

[13] Y. Xiang, J. Li, and H. Chen, "GTAN: Gated temporal attention network for credit card fraud detection," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 8, pp. 8123-8135, 2023. [14] Z. Xia and S. C. Saha, "FinGraphFL: Financial graph-based federated learning for enhanced credit card fraud detection," Electronics, vol. 13, no. 16, pp. 1-18, 2024.

[15] P. Reddy, S. Kumar, and N. Rao, "Hybrid CNN-LSTM model for credit card fraud detection," Applied Sciences, vol. 14, no. 5, pp. 1-15, 2024.

[16] Y. Duan, L. Wu, and M. Zhang, "CaT-GNN: Enhancing credit card fraud detection via causal temporal graph neural networks," arXiv preprint arXiv:2402.14708, Feb 2024.

[17] M. Aljunaid, Y. Tang, and Z. Liu, "Explainable federated learning for financial fraud detection using SHAP and LIME," Information Sciences, vol. 678, pp. 1-14, 2024.

[18] Kaggle, "Credit Card Fraud Detection Dataset," 2018. [Online]. Available: https://www.kaggle.com/mlg-ulb/creditcardfraud

10.48047/jocaaa.2024.33.04.50

[19] I. Y. Hafez, M. A. El-dosuky, and M. Zaki, "A systematic review 50of AI-enhanced techniques in credit card fraud detection," Journal of Big Data, vol. 11, no. 1, pp. 1-28, 2024.
[20] A. H. Salem, S. M. Ezzat, and K. M. Elsayed, "BGVOA-LS: Binary griffon vulture optimization algorithm with local search for feature selection in credit card fraud detection," Journal of Big Data, vol. 11, no. 1, pp. 1-22, 2024.