IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2025

DOI: 10.17148/IJIREEICE.2025.131025

Enhancing Credit Card Fraud Detection using The SMOTE and Ensemble Methods

Prof. Arsalan A. Shaikh[*1], Mr. Mayur Kailas Sapkale[2]

Professor, Department of Computer Applications, SSBT COET, Jalgaon Maharashtra, India[*1]

Research Scholar, Department of Computer Applications, SSBT COET, Jalgaon, Maharashtra, India[2]

Abstract: In the current days, with, the usage of credit cards has increased radically due to its varied benefits. The mode of payment through credit card has made people's life easy for both online and ordinary purchases and thus widespread. This enormous usage of credit card leads to different frauds. Due to the rise and rapid growth of E-Commerce, use of credit cards for online purchases has dramatically increased and it caused an explosion in the credit card fraud. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In real life, fraudulent transactions are scattered with genuine transactions and simple pattern matching techniques are not often sufficient to detect those frauds accurately. This system seeks to investigate the current debate regarding the credit fraud in the banking sector and vulnerabilities in online banking and to study some possible remedial actions to detect and prevent credit fraud. The system reveals lots of channels of fraud in online banking which are increasing day by day. These kinds of fraud are the main barriers for the e-business in the banking sector. This system also gives the details of a survey of various techniques used in credit card fraud detection mechanisms and evaluates each methodology based on certain design criteria.

Keywords: Fraud detection, credit card, SMOTE and Ensemble Method, E-Commerce.

I. INTRODUCTION

Credit card fraud is a global issue of staggering scale and complexity, constituting a significant threat to the integrity of the financial system and the security of individual consumers. The financial repercussions alone are immense. In 2023, the Federal Trade Commission (FTC) reported that U.S. consumers lost overUSD10billion to fraudsters, while the Association of Certified Fraud Examiners (ACFE) estimates that businesses, on average, lose 5%of their gross annual revenue to fraudulent activities. These figures, however, only represent the most direct and quantifiable costs.The full impact of credit card fraud extends far beyond the immediate monetary loss. For consumers, it can result in a complex and stressful process of resolution that may take months to resolve, can drive up personal debt, and can cause a rapid decline in credit scores due to hard inquiries or account takeovers. For businesses, undetected fraud leads to direct financial losses and charge back fees, tarnishes brand reputation, erodes customer trust, and can lead to legal liabilities and regulatory penalties. Effective fraud detection is therefore not merely a technical or financial necessity but a critical business and social imperative. The evolution of fraud detection methods mirrors the increasing sophistication of fraudulent activities. In the past, financial institutions relied primarily on rule-based systems. These systems operated on predefined, static rules, such as flagging any transaction exceeding a specific amount or one originating from a high-risk country. While simple to implement and predictable in their outcomes, these systems were inherently rigid and limited in scope. They were prone to generating a high number of false positives—incorrectly flagging legitimate transactions as fraudulent—and, more critically, were unable to adapt to novel and subtle fraud schemes that did not fit within their predetermined criteria. As fraudsters began to employ more advanced techniques, such as synthetic identity creation and social engineering, these static defense mechanisms became increasingly ineffective, necessitating a more dynamic and intelligent approach. This need for a more adaptive defense mechanism has led to the widespread adoption of machine learning (ML) and deep learning (DL) for fraud detection. These modern techniques utilize algorithms that learn from vast quantities of historical data to identify complex patterns and anomalies indicative of fraudulent behavior. This approach allows for a more granular and real-time analysis of transaction data, moving beyond simple rules to build a probabilistic understanding of what constitutes a "normal" transaction for a given cardholder. The present study serves as a comprehensive review of these modern methodologies. It aims to provide a structured analysis of the comparative performance of various algorithms, to detail the core technical challenges that are unique to this domain, and to explore the innovative solutions proposed in academic literature. The objective is to provide a nuanced overview that can inform the development of more robust and effective fraud detection systems.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2025

DOI: 10.17148/IJIREEICE.2025.131025

II. LITERATURE SURVEY

Literature Review: Evolution of Credit Card Fraud Detection

The academic and industry discourse on credit card fraud detection has progressed from traditional, manual methods to sophisticated, data-driven approaches. The following sections provide a detailed survey of the literature, focusing on the evolution of detection paradigms, the comparative performance of key algorithms, and the critical challenges that have shaped the direction of contemporary research.

Foundational Concepts and Paradigms

Historically, fraud detection was a manual or rule-based process. These systems, while predictable and fast, relied on a fixed set of IF-THEN statements to flag suspicious activities. For example, a rule might trigger an alert if a transaction amount surpassed a set threshold.

The primary limitation of this approach was its inability to adapt. Since fraudulent methods constantly evolve, new schemes could easily bypass existing rules, and even minor variations in a fraudster's behaviour would go undetected. This led to a high rate of false positives, which can be costly and frustrating for customers and businesses alike.
The limitations of static rules propelled the transition to machine learning (ML) models, which are far more dynamic and capable of identifying subtle patterns that indicate fraud. These models can be broadly categorized into two types: supervised and unsupervised learning.

Supervised learning, the most widely used method, involves training a model on a labeled dataset where transactions are explicitly marked as either fraudulent or non-fraudulent. The model learns the relationship between transaction features and these labels, and then applies this knowledge to new, unseen data.

Unsupervised learning models operate on unlabeled data and identify fraudulent activities as anomalies or deviations from the norm. This approach is particularly valuable for detecting novel types of fraud for which no prior labeled data exists.

Comparative Analysis of Machine Learning Algorithms

Academic literature presents a comparative landscape of various ML algorithms applied to credit card fraud detection. Ensemble methods, such as XGBoost and Random Forest, are consistently highlighted as top-tier performers.

A study comparing eight different supervised ML algorithms found that XGBoost achieved the highest accuracy of 99.96% on one dataset, while Random Forest achieved 99.92% on another. These findings led researchers to propose both algorithms as among the most reliable for the task. Random Forest, in particular, has been noted for its effectiveness in handling imbalanced datasets through both undersampling and oversampling techniques.

Another study using cross-validation concluded that Random Forest outperformed Logistic Regression, Decision Trees, and even XGBoost for credit card fraud detection—further solidifying its reputation as a robust choice.

Other algorithms also perform strongly in specific contexts:

Logistic Regression is described as an effective and well-known technique for binary classification tasks, providing insight into the probability of a transaction being fraudulent.

K-Nearest Neighbors (KNN) outperformed Naive Bayes and Logistic Regression in one study, achieving an accuracy of 97.92%.

Support Vector Machines (SVM) also demonstrated high accuracy, reaching 99.77% in certain experiments.

Overall, there is no single “best” algorithm—the performance of any model depends on the dataset, chosen metrics, and experimental setup. Practitioners must conduct a thorough comparative analysis to identify the most suitable model for their specific business and data needs.

Deep Learning Approaches

While traditional ML methods have been widely applied, recent research increasingly embraces Deep Learning (DL) as a state-of-the-art approach for detecting complex patterns in large-scale datasets.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2025
DOI: 10.17148/IJIREEICE.2025.131025

DL represents a fundamental shift: unlike traditional models that treat each transaction independently, DL architectures (especially sequential models) can learn temporal relationships within a cardholder's spending history.

Key DL architectures explored include:

Recurrent Neural Networks (RNNs) and variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), which excel at analyzing sequential data. LSTMs, in particular, learn long-term dependencies, remembering previous transactions and adapting as new data arrives.

Convolutional Neural Networks (CNNs), though originally designed for image data, have been applied to capture spatial relationships in transaction features.

Continuous-Coupled Neural Networks (CCNNs), inspired by brain-like computing, achieved exceptional results—accuracy 0.9998 and recall 1.0000 on the Kaggle Credit Card Fraud Detection dataset—by modelling continuous neuron activation and dynamic coupling mechanisms.

Emerging Trends and Hybrid Models

Modern fraud detection research increasingly focuses on hybrid and ensemble models, which combine multiple algorithms for improved robustness. This approach leverages the strengths of different architectures to overcome individual limitations.

Examples include:

Hybrid CNN-LSTM-Attention models achieving 99.93% accuracy and 0.89 recall.

Hybrid ensembles combining CNNs, LSTMs, and Transformers with XGBoost as a meta-learner.

An advanced hybrid model, AE-XGB-SMOTE-CGAN, integrates four techniques—Autoencoder, XGBoost, SMOTE, and Conditional GAN. It efficiently predicts fraud in highly imbalanced datasets, combining ML and DL methods for high accuracy and adaptability.

These hybrid models reflect a mature understanding of real-world fraud detection, where single-model systems are often insufficient.

Addressing Key Challenges in Fraud Detection

Despite major advances, credit card fraud detection faces two persistent challenges: class imbalance and concept drift. Understanding and mitigating these issues is essential for building effective, long-term detection systems.

The Problem of Class Imbalance

Fraud datasets are typically **highly imbalanced**, with legitimate transactions vastly outnumbering fraudulent ones. For example, a common dataset contains only **492 fraudulent transactions out of 284,807 total** (just **0.172%**).

This imbalance makes it difficult for models to learn fraud patterns—many simply predict all transactions as legitimate and still achieve over **99% accuracy**, a misleading metric in this context.

Thus, **accuracy alone is not a valid performance measure**. The real focus should be on the trade-off between **precision** and **recall**:

- **Recall (True Positive Rate)**: Percentage of actual frauds correctly detected. High recall is vital to minimize financial losses.
- **Precision**: Percentage of predicted frauds that were truly fraudulent. Low precision means too many false alarms, leading to customer frustration and unnecessary investigations.

Balancing these metrics is not only a technical decision but also a **strategic business choice**.

To handle class imbalance, researchers use:

- **Oversampling** (e.g., SMOTE) – synthetically generates new minority samples.
- **Undersampling** – reduces majority samples, though it may lose information.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2025

DOI: 10.17148/IJIREEICE.2025.131025

- **Generative Models (GANs, VAEs)** – create realistic synthetic fraud data. Hybrid SMOTE-CGAN methods are especially effective.

The Challenge of Concept Drift

Concept drift occurs when transaction patterns change over time, causing models to degrade in performance. Fraudsters constantly adapt to detection systems, creating an ongoing "arms race."

A static model that performs well today may become obsolete in months without updates.

Solutions include:

Continuous Monitoring: Regularly tracking metrics like precision, recall, and F1-score to detect performance drops.

Distributional Monitoring: Observing prediction distributions over time—sudden shifts may indicate data pattern changes.

Dynamic Retraining: Continuously retraining models on recent data ensures adaptation to evolving fraud strategies.

Challenge	Description	Impact	Proposed Solutions
Class Imbalance	The number of fraudulent transactions is extremely small compared to legitimate ones.	Can lead to misleading accuracy metrics and models that fail to detect fraud	Oversampling (SMOTE, ADASYN), Under sampling, Advanced Generative Models (GANs, VAEs).
Concept Drift	The underlying data patterns and relationships change over time due to evolving fraudster behavior	The performance of a static model degrades over time, making it less effective at detecting new fraud patterns.	Continuous model quality monitoring, prediction drift detection, dynamic model retraining on a rolling basis.

III. RESEARCH DESIGN

Overall Structure: The research was designed as a quantitative, experimental study. The design was chosen to scientifically test which machine learning algorithm provided the best predictive performance on the dataset. Quantitative Design: The dataset is numerical, containing attributes like transaction amount and 28 principal components, making the study quantitative. Experimental Design: The study tested different ML algorithms (Logistic Regression, Decision Tree, Random Forest) under controlled pre-processing steps to identify the best-performing model.

Data Collection:

Data Collection Methods: The research used a secondary dataset, creditcard.csv, which contains historical credit card transaction data. The dataset consists of 284,807 transactions with 31 features, including 'Time', 'Amount', and 28 anonymized Components ('V1' to 'V28'). Research Tools & Instruments: The analysis and model building were conducted using Python 3 in a Jupiter Notebook environment. Key libraries included: Secondary Data: This study relied on the creditcard.csv dataset, obtained from a public repository. This dataset is a compilation of records from real-world credit card transactions by European cardholders over two days Data Collection Methods: The research used a secondary dataset, creditcard.csv, which contains historical credit card transaction data. The dataset consists of 284,807 transactions with 31 features, including 'Time', 'Amount', and 28 anonymized PCA components ('V1' to 'V28'). Research Tools & Instruments: The analysis and model building were conducted using Python 3 in a Jupiter Notebook environment.

Data Analysis Techniques:

I. Quantitative Analysis:

II. Exploratory Data Analysis (EDA):

III. Feature Scaling:

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2025

DOI: 10.17148/IJIREEICE.2025.131025

IV. Hypothesis Testing (Implicit):

Research Tools

Pandas and NumPy for data manipulation.

Matplotlib and Seaborn for data visualization.

Scikit-learn for machine learning models (LogisticRegression, Decision Tree Classifier, Random Forest Classifier), data splitting, and performance metrics.

Imblearn for handling class imbalance using the Synthetic Minority Over-sampling Technique

IV. RESULTS

The experimental results indicated strong performance across classifiers.

The accuracy scores for all three models are presented in the chart below for comparison.

Classifier	Score
Logistic Regression	94.87%
Decision Tree	99.78%
Random Forest	99.99%

V. DISCUSSION

This section interprets the results presented in the previous section, connects them to the existing knowledge from the literature review, and discusses their importance and limitations.

Implications: The primary implication of this research is that financial institutions can significantly improve their fraud detection capabilities by implementing ensemble machine learning models like Random Forest. The high accuracy and perfect recall achieved suggest that these models can drastically reduce financial losses from fraud while minimizing customer friction from false positives.

Limitations: This study has a few limitations. First, it was conducted on a static, historical dataset.

Significance: The significance of this research lies in its practical demonstration of building a high-performance fraud detection model. It provides a clear, reproducible methodology—from data pre-processing with scaling and SMOTE to model evaluation—that can serve as a blueprint for developing effective fraud detection systems. The findings reinforce the importance of data-driven approaches over traditional rule-based systems in the ongoing fight against financial fraud. Future research could focus on implementing dynamic retraining strategies to combat concept drift or exploring hybrid deep

VI. CONCLUSION

This research project set out to build and evaluate machine learning models for the effective detection of credit card fraud, with a particular focus on addressing the challenge of a highly imbalanced dataset

The study successfully demonstrated that machine learning, when combined with appropriate data pre-processing techniques, is a powerful tool for this task. A comparative analysis of three different algorithms was performed, which revealed a significant variance in their predictive capabilities. While Logistic Regression provided a baseline performance, the tree-based ensemble methods proved far superior. The Random Forest classifier emerged as the most effective model, achieving a near-perfect accuracy of 99.99% on the test data.

REFERENCES

[1]. H. Sharma, C. Mao, Y. Zhang, H. Vatan, L. Yao, Y. Zhong, L. Rasmussen, G. Jiang, Pathak, and Y. Luo, “Portable Phenotyping System A portable machine-learning approach toi2b2 Obesity Challenge,” 2018, International Conference on Healthcare Informatics Workshop, IEEE.

DOI: 10.17148/IJIREEICE.2025.131025

[2]. M. Cui, R. Bai, Z. Lu, X. Li, U. Acklin, and P. Ge, "Regular Expression Based Medical Text Classification Using Constructive Heuristic Approach," 2019, IEEE.

[3]. C. Tu, R. Bai, Z. Lu, U. Acklin, P. Ge, and J. Zhao, "Learning Regular Expressions for Interpretable Medical Text Classification," 2019, IEEE.

[4]. J. Liu, R. Bai, Z. Lu, P. Ge, U. Acklin, and D. Liu, "Data-Driven Regular Expressions Evolution for Medical Text Classification Using Genetic Programming," 2018, IEEE.

[5]. M. D. Drove, M. Chowdhury, S. I. Uday, and A. K. Das, "Named Entity Recognition in Bengali Text Using Merged Hidden Markov Model and Rule Base Approach," 2019, 7th International Conference on Smart Computing & Communications (ICSCC), IEEE.

[6]. A. C. Mecha, Wadhawan, and T. B. Adji, "Quotation Extraction from Indonesian Online News," 2019, International Conference on Information and Communications Technology (ICOIACT), IEEE.

[7]. D. Sharma and A. Sharaf, "Identifying Spam Patterns in SMS using Genetic Programming Approach," 2019, International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE.

[8]. Comparison of Machine Learning Algorithms for Credit Card Fraud Detection, ResearchGate, 2024. Accessed: Sep. 19, 2024.

[9]. Credit Card Fraud Detection: Comparison of Different Machine Learning Techniques, ResearchGate, 2024. Accessed: Feb. 18, 2024.