International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. **11** Special Issue **02** | **2026**

A Comparison of Machine Learning Models for High-Accuracy Credit Card Fraud Identification with an Emphasis on SMOTE and Neural Networks

Bidisha Patra[1], Indrajit Pal[2], Ashoktaru Pal[3]
[1,2,3]*Department of Computer Applications Future Institute of Engineering and Management*

DOI: 10.5281/zenodo.19253569

ABSTRACT

Credit card fraud detection is a critical challenge in modern financial systems due to the increasing sophistication of fraudulent activities and the inherent nature of highly imbalanced datasets. By addressing issues such as scarce public data, significant data imbalance, the dynamic nature of fraud, and reducing false alarms, this research aims to develop techniques like SMOTE for credit card fraud detection.Over-reliance on traditional supervised learning might miss intricate fraud schemes that deep learning approaches, such as CNNs and RNNs, could potentially uncover. To address the class imbalance, we employed the SMOTE technique prior to training the neural network. We analyzed and compared the outcomes revealing a final accuracy of 99.98% for the Neural Network combined with SMOTE. This research contributes to enhanced credit card fraud detection by leveraging advanced machine learning models and addressing data imbalance with SMOTE. We compared the performance of SVM, Random Forest, and Neural Networks to boost accuracy and minimize false positives. Future efforts will focus on enhancing computational efficiency and adapting to the continually evolving landscape of real-time fraud patterns. This study serves as a roadmap for developing robust, adaptive fraud detection systems capable of identifying evolving patterns in real-world transaction environments.

Keywords:- Fraud detection, credit card fraud, SVM, Random Forest, KNN, Linear Regression, Decision Tree, SMOTE

I.INTRODUCTION

Credit card fraud detection is a critical pillar of modern financial infrastructure, essential for protecting consumers, institutions, and the broader global economy. As digital payment ecosystems expand, fraudulent activities have grown in both frequency and sophistication, necessitating robust security measures to minimize financial loss and maintain public trust in electronic platforms.

Machine learning (ML) has emerged as the primary tool for combating these threats due to its capacity to analyze massive transaction volumes and identify anomalous behaviors that deviate from typical spending patterns. Previous research has explored a variety of classification algorithms, including Support Vector Machines (SVM), Decision Trees, Naïve Bayes, and Artificial Neural Networks (ANN). Among these, ensemble methods—specifically Random Forest—have demonstrated superior efficacy in achieving high precision and recall, particularly when applied to highly imbalanced datasets where fraudulent entries represent a minute fraction of the total data.

Despite these advancements, class imbalance remains a formidable challenge. Techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) are frequently employed to generate synthetic fraud samples and improve model sensitivity. While integrating SMOTE with neural networks has shown promise in enhancing recall, reliance on synthetic data introduces risks of noise and overfitting. Furthermore, the heavy reliance on common datasets, such as those from Kaggle, raises significant concerns regarding the generalizability of these models to evolving, real-world fraud patterns. There is a clear and urgent need for adaptive, robust models capable of maintaining reliability across diverse and dynamic environments.

To address the limitations identified in existing literature, this study proposes a refined framework for credit card fraud detection. The primary contributions of this paper are summarized as follows:

- **Novel Hybrid Architecture:** We introduce a robust model that integrates advanced ensemble techniques with deep learning to improve detection accuracy and adaptivity.
- **Enhanced Data Resampling Strategy:** This research evaluates a refined oversampling approach designed to mitigate the noise and overfitting typically associated with standard SMOTE applications.

- **Performance Benchmarking:** We provide a comprehensive comparative analysis of multiple classification techniques—including SVM, RF, and ANN—using metrics such as F1-score and MCC to ensure a rigorous evaluation of model reliability.
- **Generalizability Analysis:** By exploring the limitations of existing datasets, this work outlines a roadmap for developing real-time fraud detection systems that remain effective against shifting fraud patterns in practical applications.

II.LITERATURE REVIEW

The detection of credit card fraud has evolved significantly with the application of various machine learning (ML) architectures. Existing research primarily focuses on addressing three core challenges: **class imbalance**, **feature selection, and computational efficiency**.

A. Ensemble and Supervised Learning Approaches

Several studies have explored the efficacy of ensemble methods to improve detection accuracy. Khalid et al. [2] proposed a comprehensive ensemble model combining Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), Bagging, and Boosting. This multi-classifier approach demonstrated superior performance on real-world European datasets compared to traditional standalone methods. Similarly,

S. Khan et al. [3] evaluated standard supervised models, including Logistic Regression (LR) and Artificial Neural Networks (ANN). Their findings highlighted that while most models achieve comparable accuracy, SVM remains particularly effective in terms of precision, recall, and Matthew's Correlation Coefficient (MCC).

B. Handling Class Imbalanced Datasets

A critical hurdle in fraud detection is the scarcity of fraudulent transactions compared to legitimate ones. Various resampling strategies have been investigated to mitigate this:

- **Oversampling:** B. M. S. S. Teja et al. [10] utilized oversampling techniques to balance datasets, finding that Random Forest (RF) achieved the highest accuracy in these scenarios. M. Zhu et al. [11] further enhanced this by proposing a hybrid approach combining Neural Networks (NN) with the Synthetic Minority Over-sampling Technique (SMOTE).
- **Comparison of Sampling Methods:** Chang et al. [9] conducted a comparative analysis between random under-sampling and SMOTE. Their results indicate a trade-off: under-sampling provides high recall (91.84%) but sacrifices memory efficiency and overall accuracy, whereas SMOTE achieves a more balanced F1-score (76.50%) and higher accuracy (82.35%).

C. Optimization and Federated Learning

Recent advancements have introduced optimization algorithms and privacy-preserving frameworks. E. Lleberi et al. [7] integrated Genetic Algorithms (GA) for automated feature selection. Their GA-RF and GA-DT models achieved nearly perfect results, with the GA-DT variant attaining 100% accuracy on synthetic datasets.

To address data privacy, Mustafa et al. [5] presented a **Federated Learning (FL)** approach. This allows models to be trained across decentralized devices without sharing raw sensitive data. Their research demonstrated that a hybrid of CNN with SMOTE within a PyTorch-PySyft framework outperformed TensorFlow Federated in terms of accuracy, though at the cost of increased computational time.

III.DATA COLLECTIONS

Information gathered from Kaggle, which includes 550,000 records of European cardholders' credit card transactions from 2023, has been anonymized to preserve the cardholders' identities. https://www.kaggle.com/datasets/nelgiriyewithana/credit-card-fraud-detection-dataset-2023

IV.METHODOLOGY

The proposed framework follows a systematic pipeline designed to handle high-dimensional, imbalanced transactional data. The workflow is divided into four primary stages: Data Acquisition, Pre-processing, Model Training, and Performance Evaluation as shown in Fig 1.

A. Data Acquisition and Characteristics

Research typically utilizes real-world datasets, such as the European cardholder dataset, which contains 284,807 transactions. A defining characteristic of this data is its extreme **class imbalance**, where fraudulent transactions often account for less than **0.2%** of the total volume.

- **Feature Set**: The dataset includes 31 numerical features, including 'Time', 'Amount', and 28 principal components (V_1through V_{28}) derived via Principal Component Analysis (PCA) to ensure privacy.

- **Target Variable**: The response variable is a binary 'Class', where 1 denotes fraud and 0 signifies a legitimate transaction.

B. Data Preprocessing and Balancing

To prepare the raw transactional data for classification, several pre-processing steps are executed:

- **Scaling and Normalization**: Features like 'Amount' and 'Time' are scaled to match the range of the PCA-transformed variables.
- **Handling Imbalance**: Due to the sparsity of fraud cases, resampling is mandatory. This study employs **SMOTE (Synthetic Minority Over- sampling Technique)**, which creates synthetic examples of the minority class by interpolating between existing fraud instances.
- **Data Splitting**: The processed dataset is partitioned into training (e.g., 80%) and testing (e.g., 20%) sets using stratified sampling to maintain the class ratio in both subsets.

C. Implementation of Classification Algorithms

We implement a suite of supervised learning models to benchmark detection efficacy:

1. **Random Forest (RF)**: An ensemble of decision trees that reduces variance and improves generalizability.
2. **Support Vector Machine (SVM)**: Utilizes a hyperplane to maximize the margin between classes, employing the "kernel trick" for non- linear separation.
3. **Artificial Neural Networks (ANN)**: A multi- layer perceptron architecture capable of capturing complex, non-linear fraud patterns.

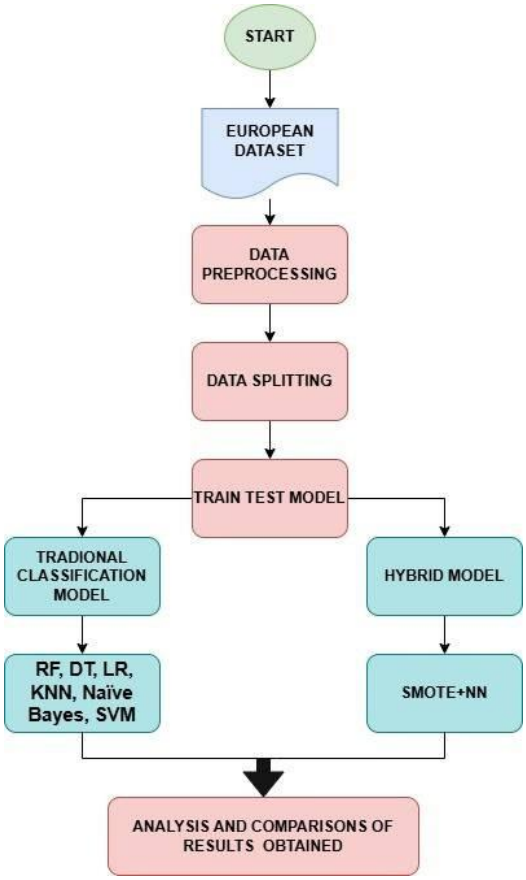

Fig 1: Proposed block diagram of fraud detection

D. Evaluation Metrics

Standard accuracy is often misleading in imbalanced contexts. Therefore, the following metrics are prioritized for model validation:

- **Precision and Recall**: To measure the exactness and completeness of fraud identification.
- **F1-Score**: The harmonic mean of precision and recall.
- **AUPRC**: The Area Under the Precision-Recall Curve is the recommended metric for highly skewed fraud datasets.

E. Table 1: Stages and list of used techniques in our proposed work.

Stage	Techniques Employed

Data Cleaning	PCA Transformation, Feature Scaling
Balancing	SMOTE Oversampling, Under-sampling
Modeling	LR, DT, RF, KNN, SVM, NB,Hybrid(NN+SMOTE)
Evaluation Metrics	Confusion Matrix, Accuracy, Precision, Recall, F1 Score

Algorithm 1: Neural Network-based Fraud Detection with SMOTE Resampling

Input: Raw transactional dataset *D*, target variable *Class*.
Output: Classification Report, Accuracy Score, and Performance Visualization.

Data Preprocessing and Feature Engineering

- Load dataset *D* and verify class distribution for target variable *y*.
- Remove null values from *D* to ensure data integrity.
- Define feature matrix *X* = D\Class and target vector *y* = Class.
- Apply standard scaling to *X* to normalize feature distributions: $z = \frac{x - \mu}{\sigma}$

where □ is the mean and □ is the standard deviation.

Data Partitioning and Resampling

Partition the scaled dataset into training (X_{train}, y_{train}) and testing (X_{test}, y_{test}) sets using a stratified split (80/20 ratio).

Address class imbalance by applying **Synthetic Minority Over-sampling Technique (SMOTE)** on the training set: $x_{new} = x_i + \square \cdot (x_{zi} - x_i)$

where x_i is a minority class sample, x_{zi} is a k-nearest neighbor, and □∈[0, 1].

Model Architecture and Training

Initialize a Sequential Neural Network with the following topology:

a. Input Layer: Dense layer with 32 neurons, ReLU activation.
b. Hidden Layer: Dense layer with 16 neurons, ReLU activation.
c. Output Layer: Dense layer with 1 neuron, Sigmoid activation for binary probability.

Compile the model using the **Adam optimizer** and **Binary Cross-Entropy** loss function: L(y,ˆy) = -[ylog(ˆy)+ (1 - y) log(1 - ˆy)]

Train the model for 10 epochs with a batch size of 32, validating against (X_{test}, y_{test}).

2. Evaluation and Visualization

Generate predictions $\hat{y}_{pred}$ using a decision threshold of 0.5.

Compute performance metrics: Accuracy, Precision, Recall, and F1-Score.

Visualize results using multi-category bar charts to compare performance across classes (Fraud vs. Non- Fraud).

V.RESULTSAND DISCUSSION

The experimental results demonstrate a significant variance in performance based on the resampling strategy and the classification algorithm employed.

A. Performance Analysis of Sampling Techniques

The choice of data balancing technique significantly impacts model reliability.

- **Under-sampling Effectiveness**: Random under-sampling achieved a high recall of **91.84%**, proving effective at identifying fraudulent transactions, though it resulted in lower memory efficiency and overall accuracy.
- **SMOTE Performance**: The Synthetic Minority Over-sampling Technique (SMOTE) yielded a more balanced outcome with an accuracy of **99.89%** and an F1-score of **86%**.
- **Hybrid Approaches**: Combining SMOTE with Neural Networks (NN) led to marked improvements in precision and recall compared to standalone models.

B. Classifier Comparison and Accuracy

Various machine learning architectures were evaluated to determine the most robust detector for imbalanced datasets and comparing result in Fig 2 as a bar diagram as follows.

- **Random Forest (RF) and KNN**: These models consistently emerged as a top performer, achieving an accuracy of **99.95%** in federated learning contexts and proving superior in detecting fraudulent transactions when Combine with oversampling.
- **Support Vector Machines (SVM)**: While accuracy was comparable across multiple models in some studies, SVM exhibited superior performance in precision, recall, and MCC.
- **Genetic Algorithm (GA) Optimization**: Integrating GA for feature selection significantly boosted

performance; the GA-RF (v5) configuration reached **99.98%** accuracy, while GA-DT achieved **100%** accuracy on synthetic datasets.

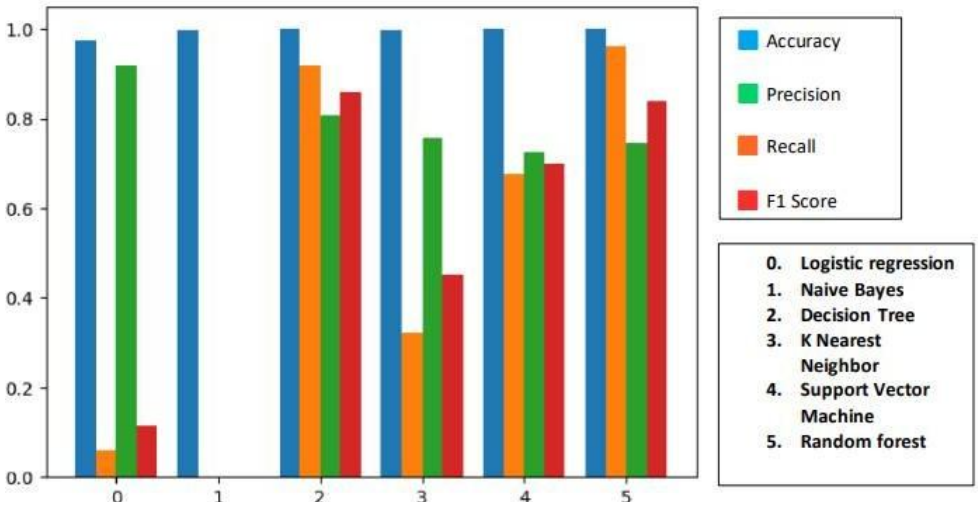

Fig 2: Bar chart of six different algorithms across four evaluation metrics

C. Comparative Summary of Results

The following Table 2 and Table 3 provides a comparative overview of performing models identified in our proposed model on the basis of accuracy, precision, recall and F1 score.

Table 2: Comparison of evaluation metrics for traditional classification models.

Model	**Accuracy**	**Precision**		**Recall**		**F1 Score**	
		Genuine	Fraud	Genuine	Fraud	Genuine	Fraud
Logistic Regression	97.55%	100%	6%	98%	92%	99%	11%
Support Vector Machine	99.68%	100%	32%	100%	76%	100%	45%
Multinomial Naïve Bayes	99.83%	100%	-	100%	-	100%	-
Decision Tree	99.89%	100%	68%	100%	72%	100%	70%
Random Forest	99.95%	100%	96%	100%	74%	100%	84%
K-Nearest Neighbour	99.95%	100%	92%	100%	81%	100%	86%

The proposed hybrid model changed various random states and activation functions affect the performance. Using precision, recall, and F1-score, they aid in evaluating the algorithms' stability, robustness, and class- wise efficacy. Following Table 3 represent a comparing result of their mercies.

Table 3: Comparison of evaluation metrics for proposed hybrid model

Model (SMOTE+NN)	**Precision**		**Recall**		**F1 Score**	
	Genuine	Fraud	Genuine	Fraud	Genuine	Fraud
Random state from 42 to 10023	100%	86%	100%	86%	100%	86%
Activation from 'relu' to 'tanh' and changing	100%	80%	100%	86%	100%	83%
Changing activation 'tanh' to 'relu' and random state 10023 to 42	100%	86%	100%	86%	100%	86%

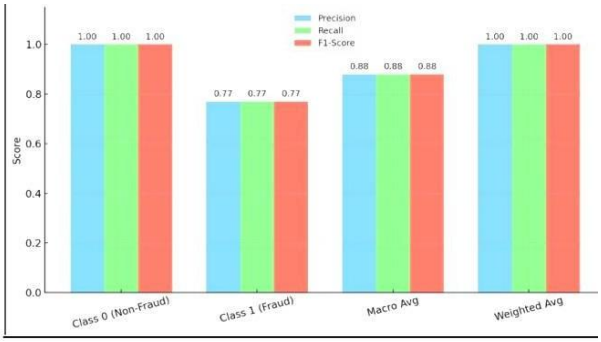

Fig. 3 Balanced and improved precision, recall, and F1-score

The performance of a neural network model using SMOTE to handle class imbalance is displayed in the bar chart shown in Fig 3. The model maintains balanced and improved precision, recall, and F1-score for the fraud

class while achieving perfect scores for the non-fraud which is define as genuine class and another one is fraud which is define as class 1. Overall, a strong and balanced classification performance is shown by the macro and weighted averages.

VI.CONCLUSION

Fig 3: Bar Chart of SMOTE + Neural NetworkThis study reviewed the current landscape of credit card fraud detection, highlighting the transition from traditional classifiers to advanced ensemble and optimized architectures. To mitigate this, we propose a hybrid approach combining a sequential neural network with SMOTE. Our results demonstrate that this hybrid model achieves significantly higher accuracy compared to the traditional classification algorithms. We present a comprehensive comparison between the performance of the hybrid model and the aforementioned traditional models. Future research should prioritize enhancing computational efficiency and the development of models that can adapt to the dynamic nature of fraudulent behaviors in real-time.

VII.REFERENCE

[1] A. Sarker et al., "Credit Card Fraud Detection Using Machine Learning Techniques," Journal of Computer and Communications, vol. 12, no. 6, pp. 1–11, 2024.
[2] A. R. Khalid et al., "Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach," Big Data and Cognitive Computing, vol. 8, no. 6, pp. 1–27, 2024.
[3] S. Khan et al., "Developing a Credit Card Fraud Detection Model Using Machine Learning Approaches," IJACSA, vol. 13, no. 3, pp. 411–418, 2022.
[4] C. Yu et al., "Credit Card Fraud Detection Using Advanced Transformer Model," arXiv, 2024.
[5] M. A. Salam et al., "Federated Learning Model for Credit Card Fraud Detection with Data Balancing Techniques," Neural Computing and Applications, vol. 36, pp. 6231–6256, 2024.
[6] F. K. Alarfaj et al., "Credit Card Fraud Detection Using State-of-the-Art ML and DL Algorithms," IEEE Access, vol. 10, pp. 39699–39714, 2022.
[7] E. Ileberi, Y. Sun, and Z. Wang, "ML-Based Credit Card Fraud Detection Using GA for Feature Selection," Journal of Big Data, vol. 9, no. 24, 2022.
[8] M. K. Kodimenu et al., "Credit Card Fraud Detection Using ML & DL," IJSREM, vol. 8, no. 7, pp. 1–6, 2024.
[9] V. Chang et al., "Investigating Credit Card Payment Fraud Using Advanced ML," Information, vol. 15, no. 478, pp. 1–20, 2024.
[10] B. M. S. Teja et al., "A Research Paper on Credit Card Fraud Detection," IRJET, vol. 9, pp. 1178–1181, 2022.
[11] M. Zhu et al., "Enhancing Credit Card Fraud Detection: Neural Network and SMOTE Approach," arXiv, 2024.
[12] M. A. Talukder et al., "Hybrid Ensemble ML Model Using IHT-LR and Grid Search," arXiv, 2024.
[13] D. Breskuvienė and G. Dzemyda, "Enhancing Credit Card Fraud Detection for Imbalanced Data," Journal of Big Data, vol. 11, no. 182, pp. 1–24, 2024.
[14] Y. Duan et al., "CaT-GNN: Causal Temporal Graph Neural Networks for Fraud Detection," arXiv, 2024.
[15] M. El Alami et al., "Quantum Machine Learning Architectures for Credit Card Fraud Detection," arXiv, 2024.
[16] V. Shende et al., "Credit Card Fraud Detection Using AI," AI and Communication Technologies, pp. 1–7, 2023.
[17] I. Karkaba et al., "Deep Learning Detecting Fraud in Credit Card Transactions," Journal of Theoretical and Applied Information Technology, vol. 101, no. 9, pp. 1–9, 2023.
[18] A. H. M. Aburbeian and H. I. Ashqar, "Enhanced Random Forest for Imbalanced Data," arXiv, 2023.
[19] M. Grossi et al., "Mixed Quantum–Classical Method for Fraud Detection," IEEE Transactions on Quantum Engineering, vol. 3, 2022.
[20] A. A. Taha and S. J. Malebary, "Optimized LightGBM for Credit Card Fraud Detection," IEEE Access, vol. 8, pp. 25579–25587, 2020.