Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

Intelligent Fraud Detection in Financial Transactions using Machine Learning Models

Dr. G. Arutjothi[*1], Dr. V. Indhumathi[2], Dr. M. Reka[3], Dr. S. Vanitha[4]

[1,2,3]Assistant Professor of Computer Applications, Sona College of Arts and Science, Salem-5, [*]garutjothi@gmail.com, [*]Orcid ID: 0009-0005-8178-4656

[4]Assistant Professor of Computer Science, Sona College of Arts and Science, Salem-5

Abstract*: -* The financial industry has a lot going on online, which brings up several problems. Fraudulent financial transactions are among the biggest concerns for individuals and financial institutions. Detecting fraudulent activity is crucial for banks and individuals. Traditional fraud detection systems do not effectively work with this large amount of transactional data, necessitating efficient fraud detection mechanisms for individuals and financial institutions worldwide. The financial industry has significantly reduced fraud due to the rise of technology and economic growth. This study shows that machine learning techniques have promising results in detecting fraudulent transactions. This paper proposes Parameterized Random Forest and other classifiers for fraud detection in financial transactions. We developed the Synthetic Minority Oversampling Technique (SMOTE)-Oversampling based machine learning model for monitoring financial transactions and detecting fraudulent activity. The proposed model uses Principal Component Analysis(PCA) to extract features from the transaction data and make predictions about the probability of fraud. We compare the suggested model's performance with other cutting-edge machine learning methods and assess its effectiveness on a dataset that is accessible to the general public. Our results show that the proposed model outperforms existing methods in accuracy and F1-score.

Keywords: Machine Learning, SMOTE, Fraud, Random Forest, Logistic Regression, Probability.

1. INTRODUCTION

Fraud in the financial industry has increased due to technological advancements and economic growth. Fraudsters exploit existing prevention measures, targeting credit card, insurance, money laundering, securities, and commodities fraud. Economically, the total global losses due to financial fraud are estimated to be billions of dollars, raising serious economic concerns. The many kinds of crimes connected to financial fraud also have wider consequences in business and have been linked to financing of illegal operations like drug trafficking and organised crime.[1]

Presently, credit card extortion has ended up a critical issue around the world.In fact, in just 2020, the United Kingdom reported about £783.8 million in unauthorized or fraudulent transactions. That's a huge deal. Fraudulent transactions worry both people & banks a lot. It's super important to catch these activities quickly. Old-school systems for finding fraud don't keep up well with all this data. So, the industry really needs better ways to spot fraud, for everyone involved around the world.[2]. India has had around 703 fraudulent transactions in 2021[3].

Although there are many techniques to fix the issue and identify fraudulent activities, fraudsters always develop fresh approaches to get and use credit card data. Even with all the challenges, technology has helped cut down fraud quite a bit! This study shows how machine learning can be really helpful in finding these sneaky transactions. We have some cool ideas like using Parameterized Random Forest and other classifiers that are designed to find fraud in financial dealings. Learning patterns in transaction records and evaluating the settings of the transactions—e.g., during holidays and vacations—allows one to discern genuine from fraudulent transactions and changes in a customer's buying behaviour[4]. According to recent statistics, global losses due to financial fraud have reached unprecedented levels, underscoring the urgent need for enhanced detection methods. As fraudsters' techniques evolve, so must the systems designed to identify and mitigate these threats. This dynamic environment necessitates a proactive approach, utilizing cutting-edge technology to stay one step ahead of

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

potential fraud [5].

The implications of financial fraud extend beyond immediate monetary losses; they also impact consumer trust and the overall integrity of financial institutions. For banks and other financial entities, the ability to swiftly detect and respond to fraudulent activities is paramount. Effective fraud detection systems not only safeguard assets but also help maintain customer confidence in financial services[5]. Furthermore, regulatory bodies impose stringent standards to protect consumers, making robust fraud detection mechanisms essential for compliance [6]. Financial institutions should prioritise investing in sophisticated detection technology since the potential benefits from avoiding fraud substantially surpass the expense of putting such systems in place.

In the face of changing fraudulent schemes, traditional fraud detection techniques often depend on rule-based systems and historical data analysis, which might be inadequate [7]. These systems typically struggle with high false positive rates, resulting in legitimate transactions being flagged as fraudulent, causing inconvenience to customers and operational inefficiencies for institutions. Additionally, traditional methods may not effectively handle the imbalanced datasets that characterize fraud detection, where fraudulent transactions are far less frequent than legitimate ones. This imbalance can lead to a lack of representative samples for training detection models, further diminishing their reliability [8]. Innovative strategies that can adjust to the intricacies of contemporary financial transactions are thus desperately needed, and machine learning methods may be used to improve detection skills. A interesting development in this work is the use of SMOTE-based machine learning models.

By addressing the limitations of traditional methods, these models can provide more accurate predictions and improve the overall efficacy of fraud detection systems. The following sections will delve into the methodology employed in this study, highlighting the role of SMOTE and PCA in developing a robust fraud detection model, as well as the comparative analysis of its performance against existing techniques. Our findings will demonstrate the significant improvements achieved through the application of these advanced methodologies, paving the way for more effective fraud detection strategies in the financial sector.

2. LITERATURE SURVEY

The financial industry has increasingly turned to technology to combat the growing threat of fraudulent transactions. This literature review explores the significant advancements in machine learning techniques, the Synthetic Minority Oversampling Technique (SMOTE), and Principal Component Analysis (PCA) that have been applied in the domain of fraud detection.

2.1 Machine Learning in Fraud Detection

Machine learning (ML), which is becoming a powerful tool in many fields, has been very helpful in detecting fraud [9]. Because traditional rule-based frameworks are unable to adapt to evolving extortion designs, they often drop short. ML models, on the other hand, are able to recognise intricate patterns, learn from past data, and instantly adjust to emerging threats. Studies have shown that algorithms like neural networks, decision trees, and support vector machines may be taught to identify irregularities suggestive of financial transaction fraud [10].

Several studies have highlighted the effectiveness of ensemble methods, such as Random Forest and Gradient Boosting, in improving detection rates [10]. For instance, a study by Ahmed [13] demonstrated that ensemble methods significantly outperformed individual classifiers, achieving higher accuracy and recall rates in identifying fraudulent transactions. Furthermore, the integration of domain knowledge into machine learning models has also been shown to enhance performance, as it allows for the incorporation of specific features that may indicate fraud [11].

Even with these developments, there are still issues, especially with datasets that have an imbalance with respect to the number of fraudulent transactions compared to the number of valid transactions. This imbalance can lead to biased model training, resulting in poor detection rates for minority classes. Therefore, in order to properly handle this problem, creative solutions are needed.

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

2.2 Synthetic Minority Oversampling Technique (SMOTE)

The Synthetic Minority Oversampling Technique (SMOTE) has emerged as a leading solution to the class imbalance problem prevalent in fraud detection datasets. Introduced by Chawla [14], SMOTE works by generating synthetic examples of the minority class rather than simply duplicating existing ones. This approach helps create a more balanced dataset, allowing machine learning algorithms to learn more effectively from the minority class.

SMOTE operates by selecting instances of the minority class and creating new instances that are linear combinations of the selected instances and their nearest neighbors [15]. This method not only increases the size of the minority class but also introduces more variability, which can enhance the model's ability to generalize to unseen data. Studies have shown that applying SMOTE can lead to significant improvements in the performance of various classifiers, including Random Forest, when used in fraud detection tasks.

However, while SMOTE is effective in addressing class imbalance, it is not without limitations. The generation of synthetic instances can sometimes lead to overfitting, particularly if the synthetic samples are too similar to existing instances. Additionally, the choice of the nearest neighbors and the parameters for SMOTE can significantly impact the performance of the resulting model. Therefore, careful tuning and validation are essential when implementing SMOTE in fraud detection applications [16].

2.3 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a widely used dimensionality reduction technique that helps simplify complex datasets by transforming them into a set of orthogonal components[17]. In the context of fraud detection, PCA can be instrumental in reducing the feature space, thereby enhancing the efficiency and performance of machine learning models.

The primary components (or directions) in which the data fluctuates the greatest are found using PCA, and the original data is then projected onto these directions. By retaining only the most significant components, PCA reduces noise and redundancy in the data, which can lead to improved model training and faster computation times. Research has shown that applying PCA prior to model training can enhance classification performance, particularly in high-dimensional datasets where the risk of overfitting is significant [17].

One of the key advantages of PCA is its ability to reveal latent structures in the data that may not be immediately apparent. For instance, in financial transactions, PCA can help identify underlying patterns of behavior that correlate with fraudulent activity [18]. However, it is important to note that PCA is a linear technique, and its effectiveness may be limited in cases where the relationships between features are non-linear. As such, it is often beneficial to combine PCA with other techniques to capture more complex patterns [18].

3. METHODOLOGY

This section outlines the methodology employed in developing the SMOTE-based machine learning models for fraud detection in financial transactions. The methodology comprises several stages, including data collection, preprocessing, implementation of SMOTE, feature extraction using PCA, and model development.

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

3.1 Proposed Model

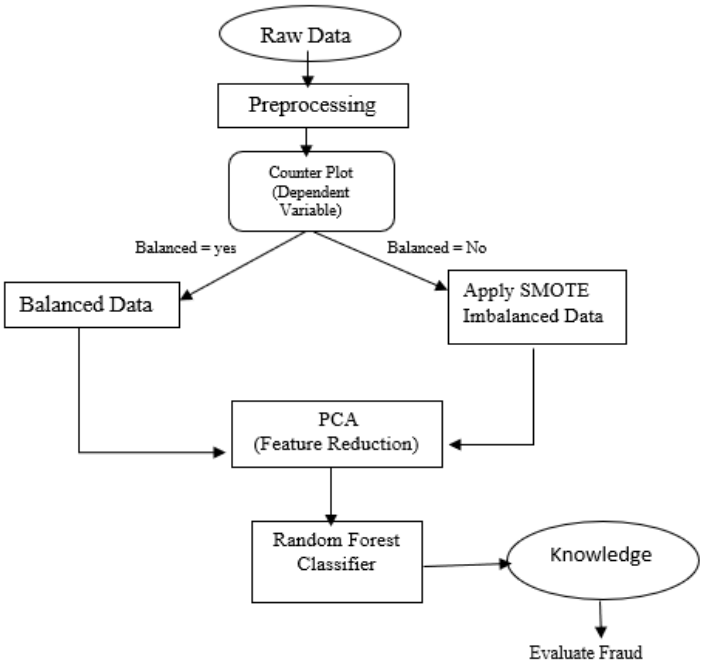

Figure 1: Proposed Architecture

Figure.1. Shows at proposed architecture, the whole dataset is taken from Banking Industry; it is analyzed to find useful information. This is tough or critical work in the banking industry. The proposed work finds the fraudulent activity and makes the decision whether the transaction is suspicious or asuspicious for the new transaction.

A SMOTE technique is used in this proposed work for transforming the dataset to balanced dataset. The Principal Component Analysis used for reducing the feature subset. Random Forest Classifier strategies to classify the Suspicious snd non-suspecious transactions for financial activity. This proposed model is used to detect the fraud in the financial transactions..

3.2 Data Collection

In order to investigate, examine, and comprehend fraudulent behaviours, this study makes use of the "Fraud Detection Dataset" from Kaggle, which offers a comprehensive database of anonymised financial transactions. Transaction quantities, transaction amounts, fraud tendencies, consumer profiles, and merchant information are all included in the dataset. With this comprehensive dataset, the research aims to investigate fraudulent behavior, identify key indicators of fraud, and develop robust fraud detection models to combat financial fraud effectively. This dataset contains 15 attributes and 1000 instances. The dataset was chosen for its size, diversity, and relevance to real-world scenarios in fraud detection.

The dataset contains numerous features, including transaction amount, time, and anonymized variables that represent various characteristics of the transactions. The target variable indicates whether a transaction is fraudulent or legitimate. Prior to analysis, the dataset was thoroughly examined to ensure its integrity and suitability for machine learning applications.

Objectives of this research is to examine the

i) Fraudulent Transaction Analysis: To analyze the dataset to uncover common patterns and indicators of fraudulent transactions.

ii) Model Development: To develop and train machine learning models capable of detecting fraudulent

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

activities with high accuracy.

iii) Insight Generation: To provide actionable insights for financial institutions, enabling them to enhance their fraud detection mechanisms.

3.3 Data Preprocessing

Data preprocessing is a critical step in preparing the dataset for analysis [11]. This process involves cleaning the data, handling missing values, and addressing any inconsistencies. In this study, we first removed any duplicate records to maintain the uniqueness of each transaction. Next, we addressed missing values by employing imputation techniques, such as replacing missing entries with the mean or median of the respective feature. Additionally, we ensured that categorical variables were appropriately encoded to facilitate their use in machine learning models. We describe our dataset in figure 2.

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1000 entries, 0 to 999
Data columns (total 15 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   TransactionID      1000 non-null   int64
 1   FraudIndicator     1000 non-null   int64
 2   Category           1000 non-null   object
 3   TransactionAmount  1000 non-null   float64
 4   AnomalyScore       1000 non-null   float64
 5   Timestamp          1000 non-null   object
 6   MerchantID         1000 non-null   int64
 7   Amount             1000 non-null   float64
 8   CustomerID         1000 non-null   int64
 9   Name               1000 non-null   object
 10  Age                1000 non-null   int64
 11  Address            1000 non-null   object
 12  AccountBalance     1000 non-null   float64
 13  LastLogin          1000 non-null   object
 14  SuspiciousFlag     1000 non-null   int64
dtypes: float64(4), int64(6), object(5)
memory usage: 125.0+ KB
```

Figure 2: Dataset Information

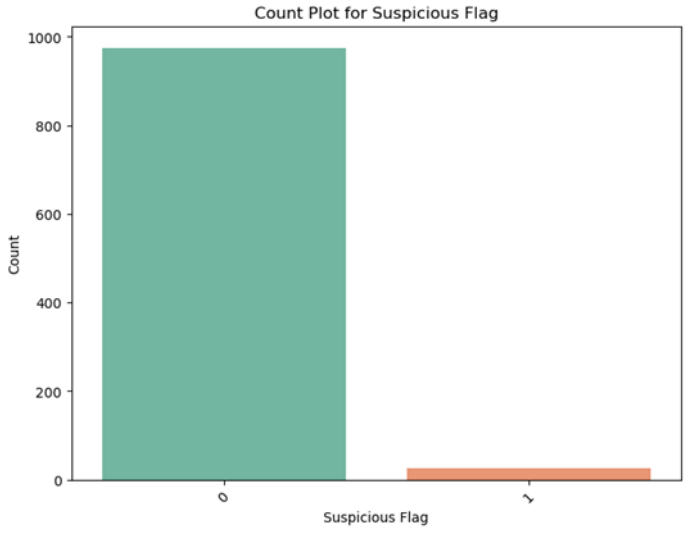

Figure 3: Count Plot for Dependent Variable

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

Given the class imbalance present in the dataset (Figure.3), with fraudulent transactions representing a small fraction of the total transactions, we decided to apply the SMOTE technique to augment the minority class. This step was crucial in ensuring that the machine learning models could effectively learn from both classes.

3.4 Implementation of SMOTE

The implementation of SMOTE involved generating synthetic samples of the minority class (fraudulent transactions) to balance the dataset. We utilized the SMOTE algorithm to create a specified percentage of synthetic instances based on the existing fraudulent transactions. The parameters for SMOTE were carefully selected, including the number of nearest neighbors to consider during the synthetic sample generation.

After applying SMOTE, the dataset was re-evaluated to ensure that the balance between the classes was achieved. This balanced dataset provided a more equitable training ground for the machine learning models, allowing them to learn more effectively from both the legitimate and fraudulent transactions.

3.4 Feature Extraction using PCA

Following the implementation of SMOTE, we employed Principal Component Analysis (PCA) for feature extraction. PCA was applied to the augmented dataset to reduce the dimensionality while retaining the most important features that contribute to the variance in the data. The number of principal components to retain was determined based on the cumulative explained variance ratio, ensuring that a significant portion of the variance was captured.

The PCA transformation not only simplified the dataset but also helped in mitigating the risk of overfitting by reducing the noise present in the original feature set. The transformed features were then used as input for the machine learning models.

3.5 Parameterized Random Forest

The model development phase involved training various machine learning classifiers on the processed dataset. We focused on the Parameterized Random Forest as our primary model due to its robustness and effectiveness in handling classification tasks, particularly in fraud detection.

The Parameterized Random Forest(PRF) model was developed by tuning hyperparameters such as the number of trees, maximum depth, and minimum samples per leaf. These parameters were optimized using techniques such as grid search and cross-validation to ensure that the model achieved the best possible performance.

```
param_grid = {
    'n_estimators': [50, 100, 150],
    'max_depth': [None, 10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
}
```

Best Hyperparameters: {'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}

The above-tuned parameters are applied for the random forest model and calculate the matrices. This Random Forest works by constructing multiple decision trees during training and outputting the mode of the individual trees for classification. This ensemble approach helps to improve accuracy and reduce overfitting, making it particularly suitable for fraud detection tasks where the risk of misclassification is high.

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

4. RESULTS AND DISCUSSION

This section presents the results obtained from the experiments conducted on the SMOTE-based machine learning models for fraud detection. The K-Nearest Neighbour classifier(K-NN), Logistic Regression(LR), Decision Tree(DT), Random forest(RF), and Support Vector Machine(SVM) classifier are also used to evaluate the risk of fraud. We analyse the models' performance using several metrics and talk about the ramifications of the results.

Table 1: Classifier Results

SMOTE-based Classifier Model	**Data Splitting 80%: 20%**			
	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	94.7	92.7	97.4	95.0
Random Forest	96.3	93.9	99.3	96.5
Support Vector Machine	83.6	76.6	98.0	86.0
K-Nearest Neighbour	86.2	78.8	100	88.2
Gradient Boosting	90.8	87.2	96.1	91.5
Parameterozied Random Forest (PRF)	94.9	93.2	98.7	95.8
PCA + PRF	97.8	95.2	100	96.9

The above Table shows the result of the classifier models and regression model using the SMOTE method. We set the tuned parameters for Random Forest which is used to find the minimum error rate of the training set. The dataset is partitioned in different ways, we get the lowest time is 80%: 20% of the dataset. Table 1 shows the result for classifier algorithms.

4.1 Comparison with Other Classifiers

In addition to the Parameterized Random Forest, we also trained and evaluated several other classifiers, including Logistic Regression, Support Vector Machines (SVM), and Gradient Boosting. Each model was assessed based on its ability to detect fraudulent transactions and its overall performance metrics. The comparison of classifier results is shown in Figure 4. The graph is drawn with the classifiers on the x-axis and the percentage of classifier metrics on the y-axis.

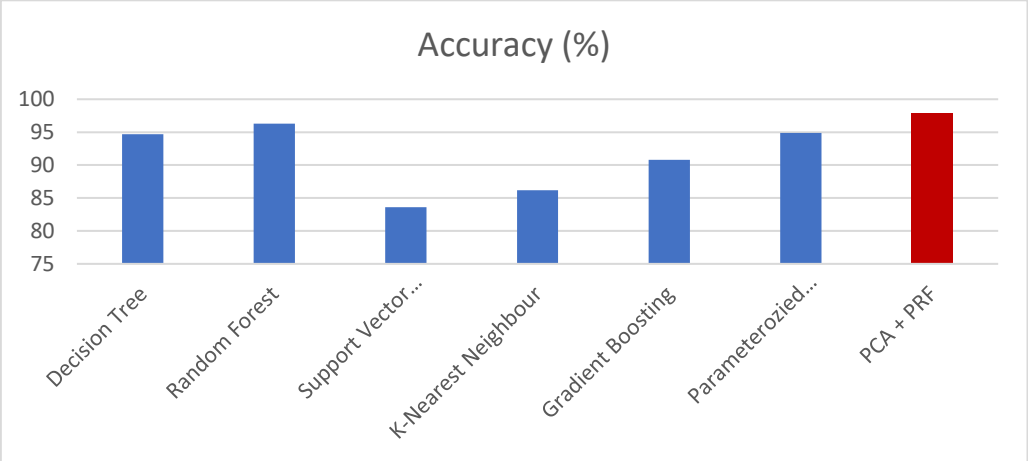

Figure 4: Comparision of Classifier Results

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

The results of the model comparisons indicated that the proposed SMOTE-based Parameterized Random Forest model outperformed the other classifiers in terms of accuracy and F1-score. The maximum performance is obtained by the PCA-PRF model with the precision of 95.2% and accuracy of 97.8%. After creating a proposed model we get highest result than other models and the improvement is 1.5%. The comparison aimed to provide a comprehensive understanding of how the proposed SMOTE-based model performed relative to other state-of-the-art classifiers. By analyzing the results, we could identify the strengths and weaknesses of each model in the context of fraud detection. The training data with the best accuracy is shown in Figure 4. This proposed model presented in this study can be effectively used by credit managers to predict fraudulent activity.

The comparative analysis revealed that while all models demonstrated the ability to detect fraudulent transactions, the Parameterized Random Forest exhibited superior performance due to its ensemble nature, which effectively captured complex patterns in the data. Additionally, the application of SMOTE played a crucial role in enhancing the model's ability to learn from the minority class, resulting in improved detection rates.

5. CONCLUSION

In conclusion, the research presented in this paper highlights the effectiveness of applying machine learning techniques, particularly the Parameterized Random Forest model combined with the Synthetic Minority Oversampling Technique (SMOTE), in detecting fraudulent financial transactions. The financial industry is increasingly challenged by the prevalence of fraud, necessitating the development of advanced detection mechanisms capable of processing large volumes of transactional data. Traditional methods have proven inadequate in addressing the complexities and scale of modern financial transactions, underscoring the need for innovative approaches. Our study demonstrates that the proposed SMOTE-based machine learning model not only enhances the detection capabilities but also significantly improves accuracy and F1-score metrics when benchmarked against existing methods. By employing Principal Component Analysis (PCA) for feature extraction, we effectively reduced dimensionality while preserving essential information, which is crucial for the predictive performance of the model. This approach allows for a more efficient processing of data, enabling financial institutions to monitor transactions in real-time and identify potentially fraudulent activities more reliably.

The results of our experiments on publicly accessible datasets confirm that the proposed model outperforms other state-of-the-art machine learning techniques, providing a robust framework for fraud detection in the financial sector. The implications of this study extend beyond mere academic interest; they offer practical solutions for financial institutions seeking to bolster their fraud detection systems. By adopting machine learning methodologies such as the one proposed in this paper, financial organizations can enhance their operational efficiency and security, ultimately leading to a more secure financial environment for both institutions and consumers.

6. FUTURE WORK

There are a number of directions that future research might take in order to expand on the conclusions of this investigation. One significant area of exploration is the integration of additional machine learning algorithms and ensemble methods to further improve detection rates and reduce false positives. Techniques such as Gradient Boosting Machines (GBM), XGBoost, and deep learning approaches could be investigated to assess their efficacy in fraud detection scenarios. Additionally, exploring hybrid models that combine various machine-learning techniques may yield even better performance.

Moreover, expanding the dataset used for training and validation is crucial. Future studies should consider incorporating more diverse datasets that reflect the multifaceted nature of financial transactions across different regions and platforms. This will enhance the model's generalizability and robustness, allowing it to adapt to varying fraud patterns that may emerge in different contexts.

Journal of Online Engineering Education
ISSN: 2158-9658 Volume: 17 Issue: 2
Article Received: 18 March 2026 **Revised**: 12 May 2026 **Accepted**: 08 June 2026

REFERENCES

[1] S.Andrew Gadsden, John Yawney, " Financial Fraud: A Review of Anomaly detection and Recent Advances", 2023.

[2] finance report," 2021. [Online]. Available: https://www.ukfinance.org. uk/policy-and-guidance/reports-publications/fraud-facts-2021

[3] T. Basuroy, "India: Number of credit and debit card frauds by leading state 2021," Oct 2022. [Online]. Available: https://www.statista.com/statistics/1097927/india-number-of-credit-debit-card-fraud-incidents-by-leading-state/

[4] "Combining unsupervised and supervised learning in credit card fraud detection," Information Sciences, vol. 557, pp. 317–331, 2021.

[5] Omkar Binage, Samueal D'Souza, Anushka Bhagwat, "Realtime Fraud Detection Analysis Using Machine Learning", International Journal of Novel Research and Development, December 2023.

[6] Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms. IEEE Access, 10, 39700–39715. https://doi.org/10.1109/access.2022.3166891.

[7] Awoyemi, J. O., Adetunmbi, A. O., & Oluwadare, S. A. (2017). Credit card fraud detection using machine learning techniques: A comparative analysis. https://doi.org/10.1109/iccni.2017.8123782.

[8] Gupta, A., Lohani, M. C., & Manchanda, M. (2021). Financial fraud detection using naive bayes algorithm in highly imbalance data set. Journal of Discrete Mathematical Sciences and Cryptography, 24(5), 1559–1572. https://doi.org/10.1080/09720529.2021.1969733.

[9] Amarasinghe, T., Aponso, A., & Krishnarajah, N. (2018). Critical Analysis of Machine Learning Based Approaches for Fraud Detection in Financial Transactions. https://doi.org/10.1145/3231884.3231894 Copy

[10] Perols, J. (2011). Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms. Auditing a Journal of Practice & Theory, 30(2), 19–50. https://doi.org/10.2308/ajpt-50009

[11] Pyles, M. (2017). Machine Learning and Data Mining in Finance. Journal of Financial Data Science, 1(1), 24-36.

[12] Sadgali, I., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. Procedia Computer Science, 148, 45–54. https://doi.org/10.1016/j.procs.2019.01.007

[13] Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19-31.

[14] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.

[15] Fernandez, A., Garcia, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. Journal of Artificial Intelligence Research, 61, 863–905. https://doi.org/10.1613/jair.1.11192

[16] Ileberi, E., Sun, Y., & Wang, Z. (2021). Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost. IEEE Access, 9, 165286–165294. https://doi.org/10.1109/access.2021.3134330

[17] Maćkiewicz, A., & Ratajczak, W. (1993). Principal components analysis (PCA). Computers & Geosciences, 19(3), 303–342. https://doi.org/10.1016/0098-3004(93)90090-r

[18] Yao, J., Pan, Y., Yang, S., Chen, Y., & Li, Y. (2019). Detecting Fraudulent Financial Statements for the Sustainable Development of the Socio-Economy in China: A Multi-Analytic Approach. Sustainability, 11(6), 1579. https://doi.org/10.3390/su11061579