International Journal of Engineering Science Invention Research & Development; Vol. 12, Issue 10, April 2026
www.ijesird.com, E-ISSN: 2349-6185

ENHANCING FRAUD DETECTION ACCURACY USING SUPERVISED MACHINE LEARNING ALGORITHMS

BRINDHA S[1] , GNANA SOWNDARI M[2] , RAJESWARI J[3]
[1,2]*UG Student , Department of Computer Science and Data Science, Nehru Arts and Science College, Coimbatore, Tamil Nadu, India.*
[3]*Assistant Professor, Department of Computer Science and Data Science, Nehru Arts and Science College, Coimbatore, Tamil Nadu, India*.

***ABSTRACT*- More people are using digital financial services and online transaction platforms. This has made it harder to spot fraud in real time. Traditional fraud detection systems rely on fixed rules, struggle to keep up with changing fraud patterns, and produce many false positives. This paper looks at how effective supervised machine learning algorithms are in improving fraud detection accuracy in financial transaction data. The study reviews four popular classification models: Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting Machine (GBM). It aims to see how suitable these models are for large-scale fraud detection. A detailed data preprocessing framework is used. This includes feature scaling, handling class imbalances with the Synthetic Minority Oversampling Technique (SMOTE), and creating behavioral features based on time and transaction patterns. The models are trained and assessed using metrics like Accuracy, Precision, Recall, F1-score, and Area Under the ROC Curve (AUC-ROC) to ensure a reliable evaluation under imbalanced conditions. Results show that models based on ensembles perform better than traditional linear classifiers. The results indicate that ensemble-based models outperform traditional linear classifiers in detecting fraudulent transactions under imbalanced conditions.These findings indicate that supervised learning techniques significantly improve fraud detection by reducing false positives and better identifying minority classes. The suggested framework offers a scalable and efficient solution for real-time monitoring of financial transactions. This research contributes to the growth of smart, data-driven fraud detection systems and provides ideas for future studies that combine hybrid and adaptive learning models.**

***KEYWORDS* : Fraud Detection, Supervised Machine Learning, Gradient Boosting, Random Forest, Support Vector Machine, Logistic Regression, Class Imbalance, SMOTE, Financial Transactions, AUC-ROC.**

1. INTRODUCTION

The rapid growth of digital financial ecosystems has changed how economic transactions happen around the world. Online banking, electronic payment gateways, mobile wallets, and e-commerce platforms are now essential parts of today's financial systems. While these technologies have increased transaction speed, access, and scalability, they have also introduced new risks that fraudsters are quick to exploit. Financial fraud has shifted from simple tricks to complex, tech-driven schemes that use automation, artificial intelligence, identity theft, and international digital networks. As transaction volumes surge, spotting fraudulent activity in real time has become a major challenge for financial institutions and regulators.

Traditional fraud detection methods mainly relied on rule-based systems that were built using expert knowledge and set logical conditions. These systems depended on fixed thresholds, such as unusually large transaction amounts, quick changes in user location, or frequent transactions in a short period. Although these early models offered some protection, they had serious drawbacks. They are static, need constant manual updates, and often fail to catch new fraud patterns that were not included in the rules. Additionally, rule-based systems tend to produce many false positives, leading to customer frustration, operational inefficiencies, and increased investigation costs. In a fast-changing fraud landscape, it is crucial to have adaptive and smart detection systems for maintaining reliability and effectiveness.

Supervised machine learning has become a strong alternative that can address many limitations of traditional methods. Unlike rule-based systems, supervised learning algorithms examine historical labeled transaction data to find hidden statistical relationships and behavioral patterns related to fraud. These models can pick up on complex interactions between various factors and can apply knowledge from past data to identify new fraudulent behavior. By continuously retraining on updated datasets, supervised models can adjust to new fraud tactics. This research explores how effective several supervised machine learning algorithms are in improving fraud detection accuracy, focusing on lowering false positives, enhancing recall for rare fraud cases, and ensuring they can scale for real- world use in financial systems.

2. BACKGROUND AND THEORETICAL FOUNDATIONS

Fraud detection is fundamentally a binary classification problem where each transaction must be identified as either legitimate or fraudulent. However, unlike typical classification tasks, fraud detection has several unique challenges. One major issue is extreme class imbalance, where fraudulent transactions make up only a small portion of the total dataset. In many real-world financial datasets, fraud cases account for less than one percent of all transactions. This imbalance complicates model training because machine learning algorithms often favor the majority class. This results in high overall accuracy but poor fraud detection performance. Thus, evaluation metrics beyond simple accuracy, such as precision, recall, and F1-score, become vital in judging model effectiveness.

Another important aspect of fraud detection is concept drift. This term describes how fraudulent behavior evolves over time. Fraudsters continuously change their tactics to evade detection systems, making static models ineffective unless they are frequently updated. This shifting environment demands learning algorithms that can adjust to new data distributions. Moreover, fraud detection systems need to work under strict real-time conditions. Financial institutions often need transaction decisions made within milliseconds to stop fraudulent payments from being approved. Thus, computational efficiency, scalability, and model optimization are key factors when choosing suitable algorithms.

From a mathematical standpoint, supervised learning aims to estimate a predictive function that connects input feature vectors with corresponding class labels. During training, model parameters are improved by minimizing a loss function that reflects the difference between predicted and actual results. In classification tasks like fraud detection, cross-entropy loss is often used due to its probabilistic nature and strong gradient characteristics. The optimization process usually involves iterative methods like gradient descent, which modify model parameters to decrease prediction errors. Understanding these foundational concepts is crucial for analyzing how different algorithms perform with imbalanced and high-dimensional financial datasets.

3. DATA PREPROCESSING AND FEATURE ENGINEERING

The performance of any machine learning model relies heavily on the quality of input data. Raw financial transaction logs often have inconsistencies, missing values, extra records, and noise that need to be fixed before training the model. Data preprocessing starts with careful cleaning. This includes removing duplicate transactions, correcting corrupted entries, and checking that timestamps are consistent. Keeping data accurate helps avoid bias and improves model reliability.

Feature scaling is another important preprocessing step, especially for algorithms that react strongly to changes in feature size. Techniques like standardization adjust variables to have a zero mean and unit variance. This prevents features with large numerical ranges from having too much influence on the model. This normalization speeds up training and helps maintain numerical stability.

Feature engineering is particularly important in fraud detection. Rather than just using raw transaction details like amount and time, engineered features look for behavioral patterns and unusual contexts. For instance, tracking how often transactions happen within a certain time frame can show suspicious rapid activity. Additionally, calculating the average transaction amount for each customer can reveal sudden changes in spending habits. Features that indicate late-night transactions or unusual weekend activity can also suggest potential fraud. By using knowledge from the field in these engineered features, models build a deeper understanding of user behavior, which helps classify complex fraudulent patterns more accurately.

4. HANDLING CLASS IMBALANCE

Class imbalance is a major challenge in fraud detection systems. When fraudulent transactions make up

only a small part of the dataset, machine learning models can reach high overall accuracy by predicting the majority, which is the legitimate class. However, these models fail because they ignore fraudulent cases. To tackle this imbalance, we need specific data-level and algorithm-level strategies.

One common method is undersampling, which lowers the number of majority class samples to balance the dataset. While this technique makes training easier, it might throw away important information from legitimate transactions, possibly harming the model's ability to generalize. Oversampling, in contrast, boosts minority class representation by duplicating existing fraud samples. Though oversampling helps balance the classes, it can cause overfitting because the model sees the same minority examples too often.

A more advanced approach is the Synthetic Minority Oversampling Technique (SMOTE). This method generates new fraud samples by blending existing minority instances. Instead of just copying data points, SMOTE creates new examples along the line segments that connect the nearest neighbors within the minority class. This technique increases diversity among fraud samples and lowers the risk of overfitting. By improving the balance of class distribution, SMOTE helps supervised learning algorithms better identify subtle fraud characteristics without losing information about the majority class.

5. LOGISTIC REGRESSION

Logistic Regression is a commonly used supervised learning method for binary classification problems. It is a basic model in fraud detection research. Despite being simple, Logistic Regression offers a solid statistical foundation and clear outcomes, making it especially useful in finance where compliance and transparency are important. Unlike linear regression, which predicts continuous values, Logistic Regression estimates the likelihood that a transaction is fraudulent by using the logistic (sigmoid) function on a linear combination of input features.

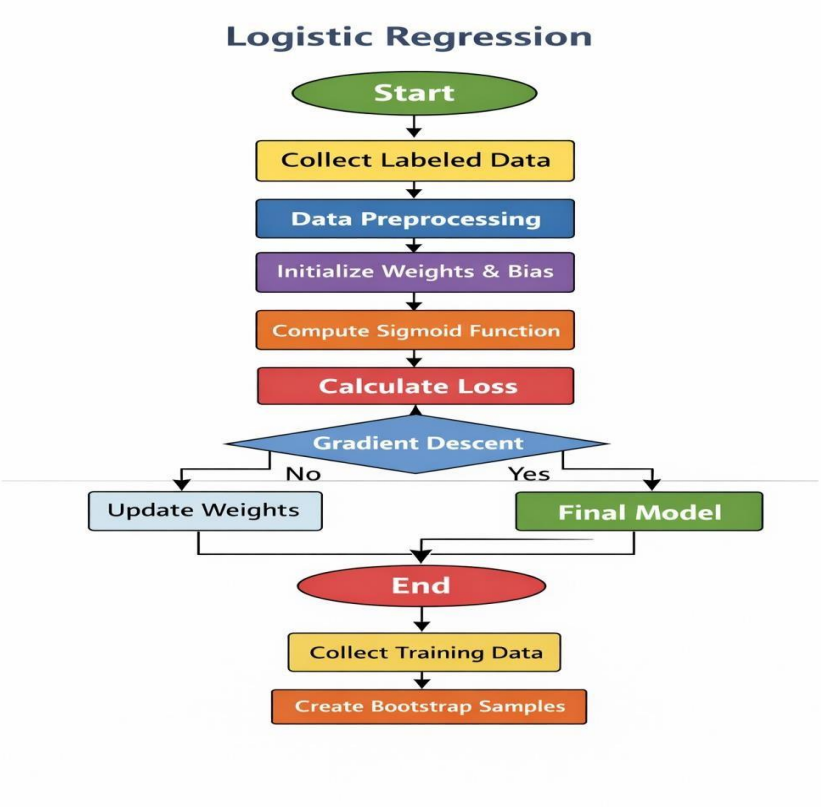

Fig 1 : Logistic Regression

Mathematically, the model calculates a weighted sum of input variables and then transforms the result

with the sigmoid function. This function converts real-valued inputs into a probability range from zero to one. This probabilistic approach allows for adjusting decision thresholds based on business needs. For instance, a lower threshold may boost sensitivity to fraud detection, while a higher threshold could lessen false alarms. Model parameters are improved by minimizing cross-entropy loss. This loss function penalizes incorrect probability estimations more when confidence is high. This process ensures the model not only classifies transactions correctly but also provides reliable probability scores. In fraud detection, Logistic Regression works well when the relationships between variables and the likelihood of fraud are roughly linear. It is efficient and suitable for large transaction systems that need real-time predictions. Moreover, the interpretability of regression coefficients helps analysts understand how specific features—like transaction amount, frequency, or geographic deviation—affect fraud risk. However, Logistic Regression struggles to capture complex nonlinear interactions between features. Fraudulent behavior often shows intricate patterns that a simple linear boundary cannot represent. This limitation can lessen the model's ability to detect highly sophisticated fraud situations. Therefore, while Logistic Regression is valuable as a benchmark and understandable model, more advanced ensemble methods usually provide better predictive accuracy.

6. RANDOM FOREST

Random Forest is an ensemble learning method that builds multiple decision trees during training and combines their predictions to enhance overall classification performance. It is part of the bagging methods, which aim to lower variance and prevent overfitting by training models on randomly sampled subsets of data. In fraud detection, where datasets are large and complex, Random Forest shows strong reliability.

The algorithm starts by generating several bootstrap samples from the original training dataset. A decision tree is built for each bootstrap sample using a randomly selected subset of features at every split. This random feature selection adds diversity among the trees, making sure that individual models do not depend too much on the same dominant attributes. After all trees are trained, predictions are combined using majority voting for classification tasks. This method helps minimize the chance that noise or outliers in the training data will significantly affect the final prediction.

Random Forest is especially useful in fraud detection because it can model nonlinear relationships and feature interactions without needing much preprocessing. Financial transaction data often includes complex dependencies between variables, such as correlations among spending patterns, merchant types, and geographic behavior. Random Forest captures these interactions automatically through hierarchical tree structures. Additionally, the algorithm shows feature importance measures, which help identify the most influential predictors in fraud detection. These insights can inform further feature engineering and risk assessment strategies.

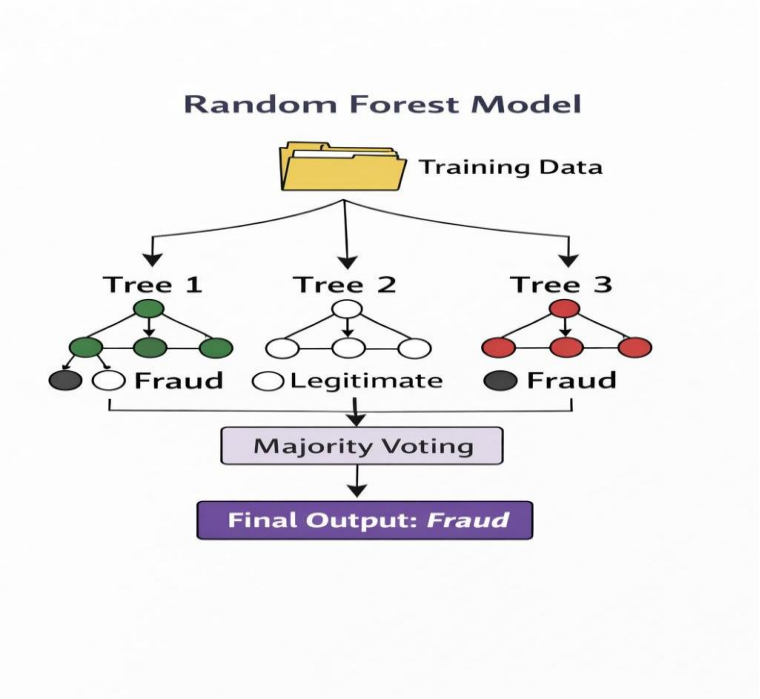

Fig 2 : Random Forest

However, Random Forest can become computationally heavy when the number of trees is large. While it offers better generalization than single decision trees, it may still face challenges with extreme class imbalance unless it is combined with resampling techniques or cost-sensitive learning. Still, in many real-world fraud detection systems, Random Forest provides a good balance between interpretability, performance, and computational efficiency.

7. SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a strong supervised learning method that finds the best separating hyperplane. This hyperplane maximizes the distance between two classes. Unlike probabilistic classifiers, SVM focuses on minimizing structural risk to improve generalization by managing model complexity. The core idea of SVM is to identify support vectors, which are key data points near the decision boundary. It then constructs a hyperplane that maximizes the distance between these points and the boundary. In fraud detection, SVM works well with high-dimensional feature spaces. Transaction datasets often include many behavioral, temporal, and engineered features, which can complicate the model. SVM handles these high-dimensional spaces effectively by using kernel functions. Kernels change input data into higher-dimensional forms, allowing the model to separate nonlinear patterns that would be hard to distinguish in the original feature space. The Radial Basis Function (RBF) kernel is often used in fraud detection because it captures local variations and subtle nonlinear connections. One major benefit of SVM is its strong theoretical base in convex optimization, which ensures a global optimum solution under specific conditions. This feature provides stability during training and lowers the risk of getting stuck in local minima. However, SVM can be computationally demanding when used with very large datasets, especially in real-time fraud detection systems that handle millions of transactions each day. Additionally, picking the right kernel parameters requires careful adjustment since poor choices can significantly hurt performance. While SVM often achieves high classification accuracy in moderately sized datasets, its scalability issues restrict its use in large financial environments. Still, when optimized correctly, SVM remains a competitive

method for fraud detection tasks that involve complex features.
Consider a training dataset $\{(x_i, y_i)\}_{i=1}^{n}$

where $x_i \in \mathbb{R}^d$represents feature vectors and $y_i \in \{-1, +1\}$denotes class labels (fraud or legitimate). The objective of SVM is to find a hyperplane defined by:

$$w \cdot x + b = 0$$

where

- wis the weight vector,
- bis the bias term.

For linearly separable data, SVM seeks to maximize the margin, which is defined as:

$$\text{Margin} = \frac{2}{\| w \|}$$

Maximizing the margin is equivalent to minimizing:

$$\frac{1}{2} \| w \|^2$$

subject to the constraint:

$$y_i(w \cdot x_i + b) \geq 1 \forall i$$

However, financial fraud data are rarely perfectly separable. Therefore, slack variables $\xi_i$are introduced to allow misclassification:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i$$

The optimization problem becomes:

$$\min_{w,b,\xi} \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{n} \xi_i$$

where Cis a regularization parameter controlling the trade-off between margin maximization and classification error.
For nonlinear fraud patterns, SVM employs kernel functions that map data into higher-dimensional space. A commonly used kernel in fraud detection is the Radial Basis Function (RBF):

$$K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2)$$

where γcontrols the influence radius of each training point.
In the fraud detection algorithms, SVM is highly effective in modeling the complex nonlinear transaction patterns. However, the computational complexity of SVM is dependent on the size of the dataset, which may pose scalability challenges in large-scale banking applications. However, SVM has excellent theoretical foundations and classification performance.

8. GRADIENT BOOSTING MACHINE (GBM)

Gradient Boosting Machine (GBM) is an advanced form of ensemble learning. GBM creates predictive models by trying to optimize a differentiable loss function. Unlike the Random Forest algorithm, which creates decision trees independently using bagging, GBM creates decision trees one by one, where the next decision tree tries to compensate for the errors made by the previous decision tree. This makes GBM a very efficient algorithm for fraud detection.
The model begins with an initial prediction:

$$F_0(x) = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma)$$

where L(y, F(x))is a differentiable loss function, commonly the logistic loss for binary classification:

$$L(y, F(x)) = \log(1 + e^{-yF(x)})$$

At each iteration m, GBM computes the negative gradient of the loss function with respect to the current model predictions. These residuals represent the direction of steepest descent:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]$$

A new weak learner $h_m(x)$, typically a shallow decision tree, is trained to approximate these residuals. The model is then updated as:

$$F_m(x) = F_{m-1}(x) + \eta h_m(x)$$

where

- ηis the learning rate controlling the contribution of each tree.

The final prediction is obtained after Miterations:

$$F(x) = \sum_{m=1}^{M} \eta h_m(x)$$

GBM is very effective in fraud detection because it has the ability to learn complex nonlinear relationships between the features of a transaction. The ability of GBM to iteratively minimize the residual error enables it to focus on the most difficult to classify fraud transactions, which are normally in the minority class. Moreover, regularization techniques such as shrinkage (small learning rate), tree depth, and subsampling can be used to prevent overfitting.
The empirical evidence has shown that GBM outperforms single classifiers in fraud detection due to its robust optimization environment and flexibility.

9. MODEL EVALUATION METRICS

Accuracy is not a sufficient measure for evaluating fraud detection models because of the class imbalance problem. Let:
True Positives (TP): Correctly identified fraud cases
True Negatives (TN): Correctly identified legitimate cases
False Positives (FP): Legitimate transactions incorrectly flagged
False Negatives (FN): Fraud cases missed Accuracy is
defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

However, in highly imbalanced datasets, accuracy may be misleading. Therefore, precision and recall are more informative.
Precision measures how many predicted fraud cases were actually fraud:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall (Sensitivity) measures how many actual fraud cases were detected:

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-score balances precision and recall:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Another critical metric is the Area Under the ROC Curve (AUC). The True Positive Rate (TPR) and False Positive Rate (FPR) are defined as:

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

The ROC curve is a graph of the TPR versus the FPR for different values of the threshold, and the AUC is a measure of the overall discriminative ability. An AUC of 1.0 is a very strong indication of having better fraud detection capability.

10. EXPERIMENTAL SETUP

The current study assesses the effectiveness of numerous supervised learning techniques in the detection of fraud cases via the implementation of a well-structured experimental design. For this purpose, a dataset containing labeled information regarding financial transactions has been utilized. This dataset contains information regarding legitimate and fraudulent transactions. Due to the imbalanced nature of fraud detection datasets, special care has been taken to ensure the reliability of model evaluation.

For model development and evaluation purposes, the dataset has been divided into training and testing sets via an 80:20 stratified sampling technique. Before training the models, all preprocessing techniques, including feature scaling and normalization, have been implemented. To ensure there was no data leakage during model development and evaluation, the Synthetic Minority Oversampling Technique (SMOTE) has been implemented on the training set only.

Hyperparameter tuning was carried out using cross-validation methods to improve predictive performance. Grid Search was used to find the best combination of parameters, such as the regularization factor in Logistic Regression, the parameters of the kernel function in Support Vector Machine, the number of estimators and maximum depth in Random Forest, and the learning rate and maximum depth of the tree in Gradient Boosting Machine. Cross-validation was used in the k-fold strategy to obtain robust and generalizable performance for different partitions of the data.

The experiments were carried out under identical conditions to compare the performance of the models. Performance is evaluated using several metrics.

11. RESULTS AND DISCUSSION

The experimental analysis shows that there are performance variations between the supervised learning algorithms considered. Logistic Regression is a good baseline performer due to its ease of use and interpretability. Nevertheless, the linear decision surface of Logistic Regression makes it difficult to capture complex patterns of fraudulent behavior that exist in financial transactions.

The performance of the Support Vector Machine is good, especially when nonlinear surfaces are used. This is due to the ability of the model to create hyperplanes that are optimal for classifying legitimate and fraudulent transactions. Nevertheless, computational difficulties arise when the size of the dataset is large. This is because Random Forest has a high overall performance, which is a result of the ensemble properties of the model and its ability to prevent overfitting. Additionally, the feature importance score of the model helps in understanding the critical fraud indicators.

Gradient Boosting Machine has shown better detection properties compared to other models. This is because the boosting properties of the model allow it to reduce errors, especially for the minority class of the dataset.

From the results, it has been confirmed that the use of ensemble-based models, including SMOTE, has a positive impact on the accuracy of fraud detection models.

12. MODEL COMPARISON AND PRACTICAL IMPLICATIONS

From the comparative assessment of the applied models, it is noted that trade-offs exist between the interpretability of the model, efficiency of the model in computations, and the effectiveness of the model in prediction. For Logistic Regression, it provides ease of implementation and interpretability, which is essential in a regulated environment. However, it may not guarantee the best detection results.

Support Vector Machine provides strong theoretical results and effective non-linear separation. However, it may pose some problems in a large-scale environment. Random Forest provides good trade- off between interpretability and effectiveness.
Gradient Boosting Machine is identified as the best model in terms of prediction strength and the ability to detect the minority class. Adaptive correction of errors in the model improves fraud identification without increasing false positives.
From a practical perspective, the implementation of ensemble-based approaches in combination with real-time preprocessing pipelines could be used to improve transaction monitoring systems in financial institutions. The integration of advanced supervised learning models in fraud detection systems promotes adaptable and data-driven financial security.

13. CONCLUSION

This research aimed to explore the role of supervised machine learning algorithms in improving the accuracy of fraud detection in financial transaction systems. The paper highlighted the significance of dealing with class imbalance problems using SMOTE and using various performance measures for evaluating the model.
From the comparative results, it was found that the ensemble learning techniques, like Gradient Boosting Machine and Random Forest, have better performance in detecting fraudulent transactions compared to other classifiers.
By effectively integrating different preprocessing techniques with powerful ensemble techniques, the risks of fraud can be minimized, and the efficiency of real-time transaction monitoring can be improved for financial institutions. In the future, research can be conducted based on different aspects, such as the use of a mixture of deep learning techniques, adaptive retraining of models, and streaming analytics for enhancing the fraud detection system against different types of frauds.

REFERENCES

1. Alonge, Enoch Oluwabusayo, Nsisong Louis Eyo-Udo, Bright Chibunna Ubanadu, Andrew Ifesinachi Daraojimba, Emmanuel Damilare Balogun, and Kolade Olusola Ogunsola. "Enhancing data security with machine learning: A study on fraud detection algorithms." *Journal of Data Security and Fraud Prevention* 7, no. 2 (2021): 105-118.
2. Afriyie, Jonathan Kwaku, Kassim Tawiah, Wilhemina Adoma Pels, Sandra Addai-Henne, Harriet Achiaa Dwamena, Emmanuel Odame Owiredu, Samuel Amening Ayeh, and John Eshun. "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions." *Decision Analytics Journal* 6 (2023): 100163.
3. Bello, O. A., Folorunso, A., Ejiofor, O. E., Budale, F. Z., Adebayo, K., & Babatunde, O. A. (2023). Machine learning approaches for enhancing fraud prevention in financial transactions. *International Journal of Management Technology*, *10*(1), 85-108.
4. Bello, Oluwabusayo Adijat, et al. "Analysing the impact of advanced analytics on fraud detection: a machine learning perspective." *European Journal of Computer Science and Information Technology* 11.6 (2023): 103-126.
5. Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised machine learning algorithms for credit card fraud detection: a comparison. In *2020 10th international conference on cloud computing, data science & engineering (confluence)* (pp. 680-683). IEEE.
6. Ismail, Mustafa Mohamed, and Mohd Anul Haq. "Enhancing enterprise financial fraud detection using machine learning." *Engineering, Technology & Applied Science Research* 14, no. 4 (2024): 14854-14861.
7. Eswar Prasad, G., G. Hemanth Kumar, B. Venkata Nagesh, S. Manikanth, and P. Kiran. "Enhancing Performance of Financial Fraud Detection Through Machine Learning Model." *J Contemp Edu Theo Artific Intel: JCETAI-101* (2023).
8. Angela, Omogbeme, Iyabode Atoyebi, Adesola Soyele, and Emmanuel Ogunwobi. "Enhancing fraud detection and prevention in fintech: Big data and machine learning approaches." *World Journal of Advanced Research and Reviews* 24, no. 2 (2024): 2301-2319.
9. Detthamrong, U., Chansanam, W., Boongoen, T., & Iam-On, N. (2024). Enhancing fraud detection in banking using advanced machine learning techniques. *International Journal of Economics and Financial Issues*, *14*(5), 177-184.

10. Khalil, Ahmed Abdelreheem. "Enhancing Insurance Fraud Detection Accuracy with Integrated Machine Learning and Statistical Methods." *Computational Economics* (2025): 1-32.
11. Lakshmi, S. V. S. S., and Selvani Deepthi Kavilla. "Machine learning for credit card fraud detection system." *International Journal of Applied Engineering Research* 13, no. 24 (2018): 16819-16824.
12. Salem, Walaa Salah, Ibrahim El-Hasnony, Ahmed Abu Elfetouh, and Amira Rezk. "Enhancing fraud detection in imbalanced datasets: A comparative study of machine learning and deep learning algorithms with SMOTE preprocessing." *Mansoura Journal for Computer and Information Sciences* 20, no. 1 (2025): 1-21.
13. Alsuwailem, Alhanouf Abdulrahman Saleh, Emad Salem, and Abdul Khader Jilani Saudagar. "Performance of different machine learning algorithms in detecting financial fraud." *Computational Economics* 62, no. 4 (2023): 1631-1667.
14. Lee, C.W., Fu, M.W., Wang, C.C. and Azis, M.I., 2025. Evaluating machine learning algorithms for financial fraud detection: Insights from Indonesia. *Mathematics*, *13*(4), p.600.
15. Husnaningtyas, Nadia, and Totok Dewayanto. "FINANCIAL FRAUD DETECTION AND MACHINE LEARNING ALGORITHM (UNSUPERVISED LEARNING): SYSTEMATIC LITERATURE REVIEW." *Jurnal Riset Akuntansi dan Bisnis Airlangga (JRABA)* 8, no. 2 (2023).
16. Njoku, D.O., Iwuchukwu, V.C., Jibiri, J.E., Ikwuazom, C.T., Ofoegbu, C.I. and Nwokoma, F.O., 2024. Machine learning approach for fraud detection system in financial institution: A web base application. *Machine Learning*, *20*(4), pp.01-12.
17. Yee, Ong Shu, Saravanan Sagadevan, and Nurul Hashimah Ahamed Hassain Malim. "Credit card fraud detection using machine learning as data mining technique." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 10, no. 1-4 (2018): 23-27.
18. Ming, R., Abdelrahman, O., Innab, N., & Ibrahim, M. H. K. (2024). Enhancing fraud detection in auto insurance and credit card transactions: A novel approach integrating CNNs and machine learning algorithms. *PeerJ Computer Science*, *10*, e2088.
19. Patel, Amit, Manishkumar M. Patel, and Pankaj S. Patel. "Enhancing Credit Card Security Using Supervised Machine Learning Approach for Intelligent Fraud Detection." In *Advancing Cyber Security Through Quantum Cryptography*, pp. 397-412. IGI Global Scientific Publishing, 2025.
20. Al Rafi, M., 2024. AI-Driven Fraud Detection Using Self-Supervised Deep Learning for Enhanced Customer Identity Modeling. *International Journal of Humanities and Information Technology*, *6*(01).
21. Taha, A. A., & Malebary, S. J. (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE access*, *8*, 25579-25587.
22. Bin Sulaiman, Rejwan, Vitaly Schetinin, and Paul Sant. "Review of machine learning approach on credit card fraud detection." *Human-Centric Intelligent Systems* 2, no. 1 (2022): 55-68.
23. Wang, H., Wang, W., Liu, Y. and Alidaee, B., 2022. Integrating machine learning algorithms with quantum annealing solvers for online fraud detection. *Ieee Access*, *10*, pp.75908-75917.
24. Malik, Pankaj, Ankita Chourasia, Rakesh Pandit, Sheetal Bawane, and Jayesh Surana. "Credit risk assessment and fraud detection in financial transactions using machine learning." *Journal of Electrical Systems* 20, no. 3s (2024): 2061-2069.
25. Ileberi, Emmanuel, Yanxia Sun, and Zenghui Wang. "Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost." *Ieee Access* 9 (2021): 165286-165294.