International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

AI-Driven Fraud detection in Online Transaction

Shravani Vaibhav Bhalerao
Department of Master of Computer Applications (MCA)
JSPM University, Pune, Maharashtra, India
Email: Shravanibhalerao26@gmail.com

----------**********----------

Abstract:

The rapid expansion of digital payment platforms, including online banking, credit cards, and UPI based transactions, has transformed the modern financial ecosystem. Although these technologies provide speed and convenience, they have also increased the risk of sophisticated online financial fraud. Fraudulent activities such as identity theft, phishing attacks, transaction manipulation, and unauthorized fund transfers are becoming more complex and difficult to detect using traditional rule-based detection systems. Conventional methods rely on predefined rules and manual monitoring, which often fail to identify emerging fraud patterns in real time.

This research proposes an AI-driven fraud detection system that enhances the security of online transactions using machine learning techniques. The system analyses multiple transaction attributes, including transaction amount, frequency, geographical location, device information, and user behavioural patterns. By training classification models on historical transaction datasets, the system learns to differentiate between legitimate and fraudulent transactions effectively. The proposed architecture follows a three-tier structure consisting of a presentation layer, an application processing layer, and a secure data storage layer. Data preprocessing techniques such as normalization, feature selection, and dataset balancing are applied to improve model accuracy and reliability.

Experimental results show that the proposed model achieves higher detection accuracy and reduces false positive rates compared to traditional rule-based approaches. The system can identify suspicious transactions in real time and generate alerts for preventive action. The findings indicate that artificial intelligence based fraud detection significantly improves transaction security, reduces financial losses, and strengthens customer trust in digital payment systems. Furthermore, the system can adapt to evolving fraud patterns through continuous learning and future integration with advanced deep learning techniques.

----------**********----------

I. INTRODUCTION

In recent years, digital payment systems such as online banking, credit cards, debit cards, and mobile payment applications have become an essential part of daily financial activities. Customers prefer digital transactions because they are fast, convenient, and easily accessible from anywhere. However, the rapid growth of online financial services has also increased the risk of cybercrime and fraudulent activities. Fraud in online transactions can result in significant financial losses, reduced customer confidence, and serious reputational damage for banks and financial institutions.

Financial fraud in digital platforms includes activities such as identity theft, phishing attacks, account takeovers, and unauthorized fund transfers. As technology advances, fraudsters are also developing more sophisticated techniques to bypass security systems. This makes fraud detection a critical requirement maintaining secure digital payment environments.

Traditional fraud detection mechanisms are primarily rule-based systems. These systems operate using predefined conditions, such as setting

transaction amount limits, detecting unusual geographic locations, or monitoring sudden changes in user behaviour. Although rule-based systems are simple to implement, they have several limitations. They cannot easily adapt to new or unknown fraud patterns and often generate a high number of false alarms. Additionally, manually updating rules requires continuous monitoring and expert intervention, which reduces efficiency. Artificial Intelligence (AI) and Machine Learning (ML) provide a more advanced and adaptive approach to fraud detection. These techniques analyze large volumes of transaction data to identify hidden patterns and unusual behaviours. Machine learning models can be trained using historical transaction records, allowing them to distinguish between legitimate and fraudulent transactions with higher accuracy. Over time, these models improve their performance as they learn from new data. This research aims to design and implement a machine learning-based fraud detection system that enhances the security of online transactions. The proposed system focuses on improving detection accuracy while minimizing false positives, ensuring that genuine customers are not unnecessarily affected. By integrating AI techniques into financial security systems, the study contributes toward building a more reliable and intelligent fraud detection framework for modern digital payment platforms.

II. LITERATURE REVIEW

Researchers have studied different approaches for fraud detection over the years. Fraud detection is generally treated as a classification problem where transactions are categorized as genuine or fraudulent. However, detecting fraud is challenging because fraudulent transactions are very few compared to normal transactions.

Rule-Based Systems

Initial systems used predefined rules to detect fraud. These systems were simple and easy to implement but could not adapt to new fraud patterns. They often generated many false alerts.

Statistical Methods

Logistic Regression is a supervised classification algorithm used to estimate the probability that a given transaction belongs to a specific class, such as fraudulent or legitimate. Unlike linear regression, Logistic Regression applies a sigmoid (logistic) function to transform predicted values into a probability range between 0 and 1.

Let the input feature vector for a transaction be represented as:

$$X = (x1, x2, x3, \dots , xn)$$

where

$x1, x2, \dots , xn$represent transaction features such as transaction amount, frequency, geographical location, and device information.

The linear combination of input features is expressed as:

$$z = \beta 0 + \beta 1x1 + \beta 2x2 + \cdots + \beta nxn$$

where

$\beta 0$is the intercept term,

$\beta 1, \beta 2, \dots , \beta n$are the model coefficients.

The logistic (sigmoid) function is defined as:

P where

$P(Y = 1 \mid X)$ represents the probability that the transaction is fraudulent, $e$is the exponential constant.

Substituting the value of z, the final probability function becomes:

P

If the predicted probability exceeds a predefined threshold (commonly 0.5), the transaction is classified as fraudulent; otherwise, it is classified as legitimate

Log-Likelihood Function

To estimate the optimal parameters β, Logistic Regression maximizes the log-likelihood function:

$$L(\beta) = \sum_{i=1}^{m} [yi \log(pi) + (1 - yi) \log(1 - pi)]$$

where

$m$is the number of training samples,

$yi$is the actual class label, $pi$is the predicted probability.

The parameters are optimized using gradient descent or other numerical optimization techniques to minimize classification error and improve prediction accuracy.

Machine Learning Techniques

Machine learning techniques have significantly enhanced fraud detection capabilities by enabling automated pattern recognition from large datasets. Algorithms such as Decision Trees, Random Forest, Support Vector

Machines (SVM), and k-Nearest Neighbours (kNN) are widely used in financial fraud detection systems.

Decision Trees classify transactions using hierarchical decision rules, making them interpretable and efficient. However, single decision trees may suffer from overfitting. Random Forest, an ensemble learning technique, overcomes this limitation by combining multiple decision trees to improve classification accuracy and generalization. Due to its robustness and stability, Random Forest is frequently reported as one of the most effective models for fraud detection.

Support Vector Machines construct optimal hyperplanes to separate classes in high dimensional feature space, while k-NN classifies transactions based on similarity measures. Although these algorithms provide higher accuracy compared to statistical models, computational complexity may increase with large datasets.

Deep Learning Approaches

Recent advancements in Artificial Intelligence have led to the adoption of deep learning models for fraud detection. Artificial Neural Networks (ANNs) and Recurrent Neural Networks (RNNs) are capable of modeling complex and nonlinear relationships within transaction data. These models automatically learn hierarchical feature representations and can detect subtle anomalies in user behaviour.

RNNs are particularly useful for analyzing sequential transaction patterns over time, allowing the system to identify evolving fraud strategies. Despite their high detection performance, deep learning models require extensive labeled datasets and significant computational resources. This requirement may limit their practical deployment in real-time systems with limited infrastructure.

Research Gap

Although various fraud detection approaches have been proposed, challenges remain in achieving an optimal balance between accuracy, adaptability, and computational efficiency. Rule-based systems lack flexibility, statistical models struggle with nonlinear patterns, and deep learning methods demand substantial computational power. Therefore, there is a need for an optimized machine learning framework that provides high detection accuracy while maintaining scalability and real-time processing capability.

This research addresses the identified gap by proposing an AI-driven fraud detection system that leverages efficient machine learning techniques to achieve accurate, fast, and reliable transaction classification in modern digital payment environments.

PROPOSED SYSTEM

The proposed system is an AI-driven fraud detection framework designed to monitor and analyze online financial transactions in real time. The primary objective of the system is to identify fraudulent transactions with high accuracy while maintaining low false positive rates. The architecture follows a

structured three-layer model to ensure scalability, security, and efficient processing.

Layer	Function	Key Components
Presentation Layer	Provides user interaction and displays transaction status	Web interface, alert system
Application Layer	Processes transaction data and performs classification	Preprocessing module, feature extraction, ML classifier
Data Layer	Stores and manages transaction and user data securely	Transaction database, fraud records, model parameters

Working Mechanism

When a user initiates a transaction, relevant features such as transaction amount, frequency, location, and device information are extracted in the Presentation Layer. These features are transmitted to the Application Layer, where data preprocessing techniques such as normalization and feature selection are applied.

The processed data is then passed to the trained machine learning model, specifically the Random Forest classifier. The model evaluates the transaction and computes the probability of fraud. If the predicted probability exceeds the defined threshold, the transaction is classified as fraudulent.

Upon detection of suspicious activity, the system generates an alert notification and can automatically block the transaction to prevent financial loss. All transaction details and classification results are stored securely in the Data Layer for future analysis and model retraining.

Algorithm

1. Collect transaction data from user input or dataset.
2. Perform data preprocessing including handling missing values, encoding categorical variables, and normalization.
3. Apply feature selection to identify important transaction attributes.
4. Split dataset into training and testing sets.
5. Train the classification model using Random Forest algorithm.
6. Evaluate model performance using accuracy, precision, recall, and F1-score.
7. Deploy the trained model for real-time fraud prediction.
8. Classify new transactions as genuine or fraudulent.
9. Generate alerts for suspicious transactions.

Mathematical Model

Let:

$X = \{x_1, x_2, x_3, \ldots, x_n\}$ represent the transaction feature vector $Y \{0,1\}$ represent the output class

Where:

0 = Genuine Transaction

1 = Fraudulent Transaction

The classification function can be defined as:

$Y = f(X)$

Where f represents the trained machine learning classifier.

The objective is to minimize classification error:

$Error = | Y\_actual - Y\_predicted |$

The model performance is evaluated using:

$Accuracy = (TP + TN) / (TP + TN + FP + FN)$

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

Where:

TP = True Positive

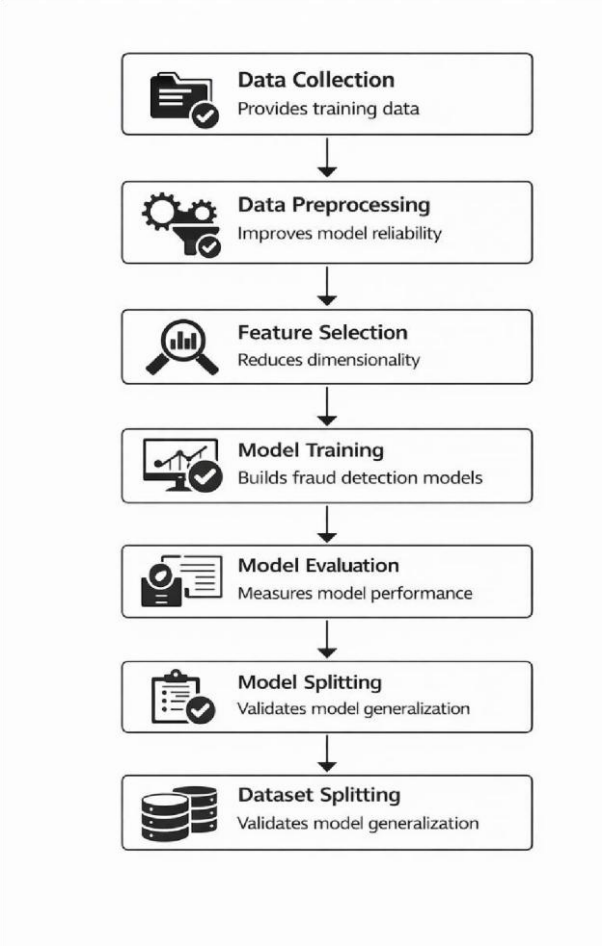
TN = True Negative

FP = False Positive

FN = False Negative

METHODOLOGY

The dataset used in this research consists of labeled transaction records containing features such as transaction amount, transaction time, location, and transaction type. The overall methodology follows a structured machine learning pipeline to ensure accurate fraud detection.



The above diagram illustrates the sequential workflow of the proposed fraud detection methodology. The process begins with data collection of labeled transaction records, followed by data preprocessing to handle missing values, normalization, and dataset balancing. Feature selection is performed to identify significant attributes influencing fraud detection. Subsequently, multiple machine learning models such as Logistic Regression, Decision Tree, and Random Forest are trained. The trained models are evaluated using performance metrics including Accuracy, Precision, Recall, and F1-score. Finally, dataset splitting into training and testing sets ensures proper validation and generalization of the model.

RESULTS AND DISCUSSION

The performance of the implemented machine learning models was systematically evaluated using widely accepted classification metrics, including Accuracy, Precision, Recall, and F1score. The dataset was divided into training (80%) and testing (20%) subsets to ensure unbiased model validation and proper generalization assessment.

Given the highly imbalanced nature of the dataset, particular emphasis was placed on Precision and Recall, as accuracy alone may provide misleading results in fraud detection problems.

Logistic Regression Performance

Logistic Regression served as the baseline model for classification. As a linear classifier, it estimates the probability of a transaction being fraudulent using the sigmoid function applied to a linear combination of input features.

The model achieved moderate accuracy and demonstrated stable convergence during training. However, due to its linear decision boundary, it struggled to capture complex, nonlinear relationships present in fraudulent transaction behaviour.

Key observations:

Reasonable overall accuracy

Moderate precision

Lower recall compared to ensemble methods Higher misclassification of minority (fraud) class

This indicates that while Logistic Regression performs well for linearly separable data, it is less effective for intricate fraud detection patterns.

Decision Tree Performance

The Decision Tree model improved classification performance by modeling nonlinear relationships between features. It constructs hierarchical decision rules based on feature splits that maximize information gain.

Compared to Logistic Regression, the Decision Tree:

- Improved recall for fraudulent transactions
- Reduced misclassification of minority class
- Provided interpretable decision rules

However, standalone Decision Trees are prone to overfitting, especially when trained on imbalanced datasets. Although performance improved, slight instability in predictions was observed due to high variance.

Random Forest Performance

Random Forest achieved the highest overall accuracy of approximately 96%, significantly outperforming the other models. This ensemble technique builds multiple decision trees using bootstrapped samples and aggregates their predictions through majority voting. The superior performance can be attributed to:

1. Reduced Overfitting – Averaging multiple trees minimizes variance.

2. Robust Feature Handling – Random feature selection improves generalization.

3. Better Minority Class Detection – Higher recall for fraudulent transactions.

Performance characteristics:

- Highest accuracy (~96%)
- Lower false positive rate
- Higher recall (improved fraud detection rate)
- Balanced precision and recall
- Strong F1-score

The ensemble learning mechanism enhances predictive stability and reduces model bias, making Random Forest particularly effective for fraud detection tasks.

Comparative Analysis

When comparing all three models:

- Logistic Regression provides a strong baseline but lacks complexity handling.
- Decision Tree improves detection but may suffer from overfitting.
- Random Forest achieves optimal balance between bias and variance.

Importantly, Random Forest demonstrated:

- Better fraud identification capability
- Reduced false alarms (fewer legitimate transactions flagged as fraud)
- Higher robustness against data imbalance

This confirms that ensemble-based machine learning approaches are more suitable for realworld fraud detection systems.

Implications of Results

The experimental results clearly indicate that AI-based machine learning models significantly outperform traditional rulebased systems. Rule-based systems rely on static thresholds and predefined patterns, making them less adaptable to evolving fraud strategies.

In contrast, machine learning models:

- Learn dynamic fraud patterns from historical data

- Adapt to new fraud behaviours
- Provide scalable and automated detection

Therefore, the adoption of AI-driven fraud detection systems enhances:

- Financial security
- Operational efficiency
- Customer trust
- Risk mitigation capabilities

Conclusion

The rapid expansion of digital payment systems has significantly increased the risk of online financial fraud. Traditional rule-based fraud detection mechanisms are limited in their ability to detect new and complex fraud patterns. This project proposed and implemented an AI-Driven Fraud Detection System that uses machine learning techniques to identify suspicious transactions in real time.

The system was designed using a structured three-layer architecture consisting of presentation, processing, and data storage layers. Transaction data was preprocessed and analyzed using supervised machine learning algorithms, with Random Forest selected as the primary classifier due to its high accuracy and robustness. The experimental results demonstrated strong performance in terms of accuracy, precision, recall, and reduced false positive rate.

The proposed system successfully classifies transactions as genuine or fraudulent and generates alerts for suspicious activities. It improves detection speed, reduces financial losses, and enhances user trust in digital payment platforms. The project confirms that artificial intelligence based approaches provide a reliable and scalable solution for modern fraud detection challenges.

Future Enhancement

Although the proposed system performs effectively, several enhancements can further improve its performance and applicability:

- Integration with real-time banking APIs for live deployment.
- Implementation of advanced deep learning models such as Neural Networks or LSTM for improved pattern recognition.
- Use of real-time behavioural biometrics such as typing speed or device fingerprinting.
- Incorporation of anomaly detection techniques for identifying unknown fraud patterns.
- Deployment on cloud platforms for better scalability and distributed processing.
- Implementation of blockchain-based verification mechanisms for enhanced transaction security.
- Continuous model retraining using updated transaction datasets to adapt to evolving fraud techniques.

REFERENCES

[1] L. Breiman, “Random forests,” Machine

Learning, vol. 45, no. 1, pp. 5–32, 2001.

[2] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM

Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.

[3] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective,” Expert Systems with Applications, vol.

41, no. 10, pp. 4915–4928, 2014.

[4] T. Fawcett and F. Provost, “Adaptive fraud detection,” Data Mining and Knowledge

Discovery, vol.

1, no. 3, pp. 291–316, 1997.

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA:

MIT Press, 2016.

[6] D. J. Hand and W. E. Henley, "Statistical classification methods in consumer credit scoring: A review," Journal of the Royal

Statistical Society, vol. 160, no. 3, pp. 523– 541, 1997.

[7] S. Jha, M. Guillen, and J. C. Westland, "Employing transaction aggregation strategy to detect credit card fraud," Expert Systems with Applications, vol. 39, no. 16, pp.

12650–12657, 2012.

[8] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," Statistical Science, vol. 17, no.

3, pp. 235–255, 2002.

[9] P. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data miningbased fraud detection research," Artificial Intelligence Review, vol. 34, no. 1, pp. 1– 14, 2010.

[10] C. C. Aggarwal, Data Mining: The Textbook. New York, NY, USA: Springer, 2015.

