Adaptive Real-Time Payment Fraud Detection Using Isolation Forest and Drift-Aware Learning

VENDRA PRASANNA

PG Scholar. Department of MCA, DNR College, Bhimavaram, Andhra Pradesh

A. Naga Raju

(Assistant Professor), Master of Computer Applications, DNR College, Bhimavaram, Andhra Pradesh

ABSTRACT

The rapid growth of digital transactions has significantly increased the risk of financial fraud, making real-time fraud detection systems essential for secure payment processing. Traditional rule-based systems are increasingly ineffective due to their inability to adapt to evolving fraud patterns. This research proposes an adaptive real-time payment fraud detection system leveraging machine learning techniques, specifically the Isolation Forest algorithm, combined with a drift detection mechanism using Adaptive Windowing (ADWIN). The proposed system is designed to identify anomalous transactions by analyzing behavioral and contextual features such as transaction amount, location score, device trust score, transaction frequency, and temporal patterns. The Isolation Forest model is utilized due to its efficiency in detecting anomalies in high-dimensional datasets without requiring labeled fraud data. This unsupervised learning approach makes the system scalable and suitable for real-world deployment where labeled datasets are often limited or imbalanced. To enhance adaptability, the system integrates a simplified ADWIN algorithm that continuously monitors incoming transaction streams to detect concept drift. Concept drift refers to changes in data distribution over time, often caused by evolving fraud strategies. By detecting drift in transaction patterns, the system can dynamically respond to changes and maintain high detection accuracy.The architecture includes an integrated preprocessing pipeline that performs feature engineering, such as extracting time-based attributes and scaling numerical features using StandardScaler. The processed data is then fed into the trained model for anomaly detection. The system outputs a fraud classification along with a confidence score, enabling decision-makers to assess the severity of each flagged transaction. Additionally, the system is implemented using a Django-based web interface, allowing users to input transaction details and receive real-time fraud analysis. This enhances usability and demonstrates the practical applicability of the solution. Experimental results show that the system effectively identifies suspicious transactions while maintaining low false-positive rates. The combination of Isolation Forest and ADWIN ensures both accuracy and adaptability, making the proposed model suitable for dynamic financial environments. In conclusion, this research contributes a robust, scalable, and adaptive fraud detection framework capable of handling real-time data streams and evolving fraud patterns. Future work may involve integrating deep learning models and expanding feature sets for improved performance.

KEYWORDS: Payment Fraud Detection, Isolation Forest, ADWIN, Anomaly Detection, Machine Learning, Concept Drift, Real-Time Analytics, Financial Security

I. INTRODUCTION

With the increasing adoption of online payment systems, financial transactions have become faster and more convenient. However, this digital transformation has also led to a rise in fraudulent activities, posing significant challenges to financial institutions and consumers. Fraudulent transactions not only result in financial losses but also damage customer trust and organizational reputation. Therefore, developing efficient and adaptive fraud detection systems has become a critical requirement in the financial sector. Traditional fraud detection systems rely heavily on rule-based approaches, where predefined rules are used to identify suspicious activities. While these systems are simple to implement, they lack flexibility and fail to detect new or evolving fraud patterns. Moreover, maintaining and updating rule sets is time-consuming and often inefficient. As fraudsters continuously adapt their techniques, there is a need for intelligent systems capable of learning and evolving in real time. Machine learning has emerged as a powerful tool for fraud detection, offering the ability to analyze large volumes of transactional data and identify hidden patterns. Supervised learning methods require labeled datasets, which are often difficult to obtain due to privacy concerns and the rarity of fraud cases. This limitation makes unsupervised learning techniques, such as anomaly detection, more suitable for real-world applications.

In this project, we propose a machine learning-based fraud detection system using the Isolation Forest algorithm. Isolation Forest is particularly effective for anomaly detection as it isolates anomalies instead of profiling normal data points. It is computationally efficient and works well with high-dimensional data, making it ideal for real-time systems. To address the issue of changing fraud patterns, the system incorporates a drift detection mechanism using Adaptive Windowing (ADWIN). Concept drift occurs when the statistical properties of transaction data change over time, which can degrade model performance. ADWIN monitors the data stream and detects changes in distribution, allowing the system to adapt accordingly. The system also includes a preprocessing module that performs feature extraction and normalization, ensuring consistency between training and real-time data. A web-based interface built using Django allows users to interact with the system and analyze transactions in real time. Overall, the proposed system aims to provide a scalable, efficient, and adaptive solution for payment fraud detection. By combining anomaly detection with drift awareness, the system offers improved accuracy and robustness in dynamic environments.

II. LITERATURE SURVEY (WITH EXISTING METHODS)

Payment fraud detection has been widely studied in recent years, with various machine learning and statistical approaches proposed to improve detection accuracy. Traditional methods include rule-based systems and statistical techniques such as logistic regression. While these approaches provide baseline performance, they often struggle with evolving fraud patterns and imbalanced datasets. Supervised learning algorithms such as Decision Trees, Support Vector Machines (SVM), and Neural Networks have been extensively

used for fraud detection. For instance, Random Forest classifiers have shown promising results due to their ability to handle large datasets and capture nonlinear relationships. However, these methods require labeled datasets, which are often scarce and imbalanced, leading to biased models. Unsupervised learning techniques, particularly anomaly detection methods, have gained popularity due to their ability to detect unknown fraud patterns. The Isolation Forest algorithm, introduced by Liu et al., is one of the most effective anomaly detection methods. It isolates anomalies by randomly partitioning data, making it computationally efficient and suitable for large datasets. Several studies have demonstrated its effectiveness in fraud detection tasks. Another important aspect of fraud detection is handling concept drift. Fraud patterns evolve over time, making static models less effective. Drift detection methods such as ADWIN (Adaptive Windowing), DDM (Drift Detection Method), and EDDM (Early Drift Detection Method) have been proposed to address this issue. ADWIN, in particular, is widely used due to its ability to automatically adjust window sizes and detect changes in data distribution with statistical guarantees.Recent research has focused on hybrid approaches that combine multiple techniques for improved performance. For example, integrating anomaly detection with drift detection mechanisms has shown significant improvements in adaptability and accuracy. Deep learning models, such as autoencoders and recurrent neural networks, have also been explored for capturing complex transaction patterns. Despite these advancements, challenges remain in achieving real-time performance, scalability, and interpretability. Many existing systems are computationally expensive or lack transparency, making them difficult to deploy in real-world environments. The proposed system builds upon these existing methods by combining Isolation Forest for anomaly detection with ADWIN for drift detection. This hybrid approach ensures both accuracy and adaptability, addressing key limitations of traditional and machine learning-based systems.

III. EXISTING SYSTEM

Existing payment fraud detection systems primarily rely on rule-based and supervised learning approaches. Rule-based systems use predefined conditions to flag suspicious transactions, such as high transaction amounts or unusual locations. While these systems are easy to implement, they lack flexibility and are unable to detect new fraud patterns. Maintaining and updating rules is also a labor-intensive process. Supervised machine learning models, such as logistic regression, decision trees, and support vector machines, have been widely used to improve detection accuracy. These models require labeled datasets for training, which can be difficult to obtain due to privacy concerns and the rarity of fraudulent transactions. Additionally, these models often suffer from class imbalance issues, where the number of fraudulent transactions is significantly lower than legitimate ones. Another limitation of existing systems is their inability to handle concept drift. Fraud patterns change over time, and static models become less effective as new types of fraud emerge. Most traditional systems do not incorporate mechanisms to detect and adapt to these changes, leading to reduced performance over time.Furthermore, many existing solutions are not designed for real-time processing, making them unsuitable for modern digital payment systems where instant decision-making is required. High computational costs and lack of scalability also pose challenges for deployment in large-

scale environments. In summary, existing systems face limitations in adaptability, scalability, and real-time performance. These challenges highlight the need for advanced solutions that can dynamically learn from data and respond to evolving fraud patterns, which is addressed by the proposed system.

IV. PROPOSED METHOD

The proposed system introduces an adaptive and real-time payment fraud detection framework that combines anomaly detection with concept drift awareness. The system leverages the Isolation Forest algorithm to identify fraudulent transactions and integrates Adaptive Windowing (ADWIN) for detecting changes in transaction behavior over time. The architecture is designed to process streaming transaction data efficiently. Each transaction is analyzed based on multiple features, including transaction amount, location score, device trust score, transaction frequency, and temporal attributes such as the hour of the transaction. These features provide both behavioral and contextual insights, enabling the system to distinguish between legitimate and suspicious activities. The Isolation Forest model is used due to its ability to isolate anomalies without requiring labeled data. This is particularly useful in fraud detection, where fraudulent transactions are rare and difficult to label. The model assigns an anomaly score to each transaction, which is used to determine whether the transaction is fraudulent.

To address the issue of evolving fraud patterns, the system incorporates ADWIN, a drift detection algorithm that monitors the statistical properties of incoming data. ADWIN detects significant changes in data distribution by comparing the means of different segments of a data window. When drift is detected, the system can trigger retraining or adjustment mechanisms to maintain detection accuracy. The system also includes an integrated preprocessing module that performs feature engineering and normalization to ensure consistency between training and real-time data. A Django-based web interface allows users to input transaction details and receive instant fraud predictions along with confidence scores. Overall, the proposed system provides a scalable, adaptive, and efficient solution for real-time fraud detection in dynamic financial environments.

V. IMPLEMENTATION

The implementation of the proposed payment fraud detection system is carried out using Python, integrating machine learning libraries and a web framework for real-time interaction. The system is divided into multiple components, including data preprocessing, model training, anomaly detection, drift detection, and web interface integration. The core of the system is the PaymentFraudEngine, which encapsulates all functionalities required for fraud detection. The system begins with the initialization of the preprocessing module and the Isolation Forest model. The preprocessing module uses Standard Scaler to normalize input features, ensuring that all variables contribute equally to the model. During the initial training phase, synthetic transaction data is generated to simulate normal transaction behavior. This dataset includes features such as transaction amount, location score, device score, transaction frequency, and transaction time. The data is scaled and used to train the Isolation Forest model. The trained model and scaler

are stored using joblib for reuse. When a new transaction is received, it is first passed through the preprocessing module. Feature engineering is performed to extract the hour from the timestamp, and the data is transformed into a format suitable for the model. The scaled data is then passed to the Isolation Forest model, which predicts whether the transaction is normal or anomalous.The model also computes a decision function score, which indicates the degree of anomaly. A threshold is applied to determine whether the transaction is fraudulent. Additionally, a confidence score is calculated based on the anomaly score, providing a measure of certainty in the prediction. To handle concept drift, the system integrates a simplified implementation of ADWIN. This module continuously monitors transaction values and detects changes in their statistical distribution. If a significant change is detected, it indicates a shift in transaction patterns, prompting the system to adapt accordingly. Drift detection is essential for maintaining model performance in dynamic environments where fraud patterns evolve over time. The user interface is built using the Django framework. Users can input transaction details through a web form, and the system processes the data in real time. The results, including fraud status, confidence score, and drift detection status, are displayed to the user. The modular design ensures scalability and flexibility, allowing future integration of advanced models or additional features.

VI. ALGORITHMS

1. Isolation Forest Algorithm

Isolation Forest is an unsupervised anomaly detection algorithm that isolates anomalies by randomly partitioning data. Unlike distance-based methods, it focuses on how easily a data point can be separated from others.

Steps:

1. Randomly select a feature.
2. Randomly select a split value between min and max.
3. Partition the data recursively.
4. Compute path length for each data point.
5. Anomalies have shorter path lengths.

This method is efficient and works well with high-dimensional data, making it suitable for fraud detection.

2. ADWIN (Adaptive Windowing) Algorithm

ADWIN is used for detecting concept drift in streaming data.

Steps:

1. Maintain a sliding window of recent data.
2. Split the window into two sub-windows.

3. Compare their means statistically.
4. If the difference exceeds a threshold:
    - Drift is detected
    - Older data is discarded

ADWIN dynamically adjusts window size and ensures accurate drift detection in evolving data streams.

3. Fraud Decision Logic

1. Input transaction data
2. Preprocess and scale features
3. Predict anomaly using Isolation Forest
4. Compute anomaly score
5. If score < threshold → Fraud
6. Else → Normal
7. Update ADWIN with transaction value
8. Output fraud status and confidence score

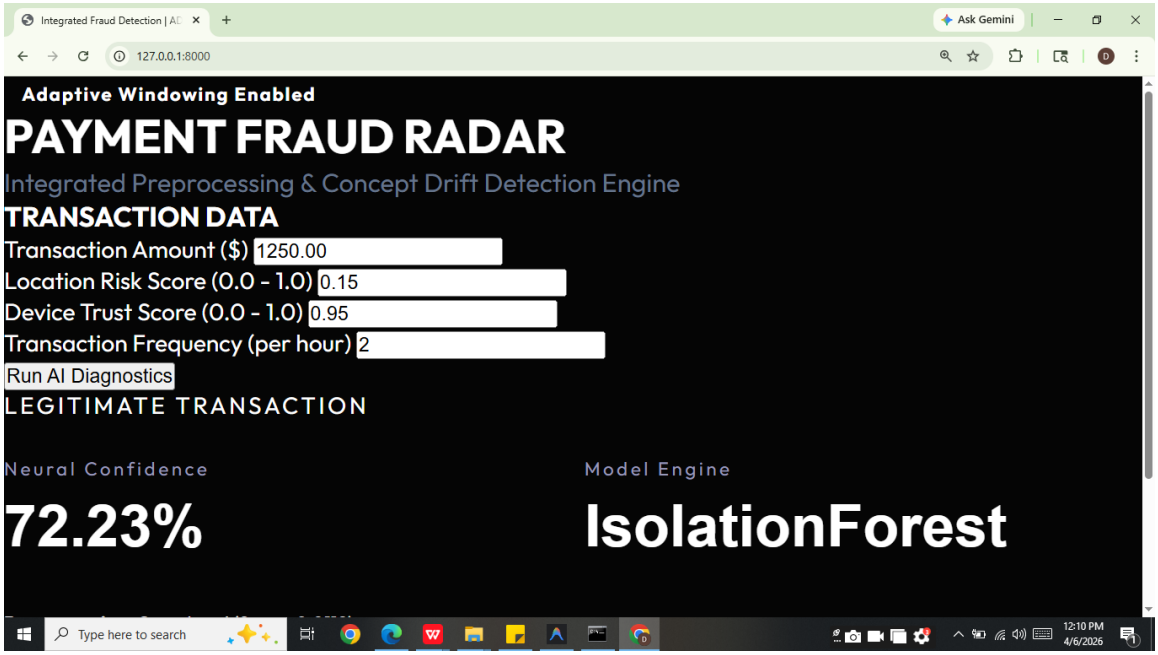
VII. SYSTEM DESIGN

The system design follows a modular and layered architecture to ensure scalability, flexibility, and real-time performance. It consists of four main layers: Data Input Layer, Processing Layer, Machine Learning Layer, and Presentation Layer. The Data Input Layer handles user inputs through a web interface built using Django. Users provide transaction details such as amount, location score, device score, and transaction frequency. The system also captures the current timestamp automatically. The Processing Layer is responsible for data preprocessing and feature engineering. This includes converting input data into a structured format, extracting time-based features, and scaling numerical values using StandardScaler. This layer ensures that the data is consistent with the format used during model training. The Machine Learning Layer is the core of the system. It includes the Isolation Forest model for anomaly detection and the ADWIN module for drift detection. The Isolation Forest model analyzes the processed data and assigns an anomaly score to each transaction. Based on this score, the system determines whether the transaction is fraudulent.

The ADWIN module continuously monitors incoming data to detect changes in data distribution. When drift is detected, it indicates that the model may need to be retrained or updated. This ensures that the system remains effective even as fraud patterns evolve over time. Handling concept drift is critical in real-time systems, as static models degrade in performance when data distributions change. The Presentation Layer displays the results to the user through the web interface. It shows whether the transaction is fraudulent, the confidence level, and whether drift has been detected. The system also includes a storage component where trained models and scalers are saved for reuse. This reduces computational overhead and ensures faster processing. Overall, the design

ensures real-time processing, adaptability, and scalability, making it suitable for deployment in modern financial systems.

SYSTEM DESIGN IMAGES



VIII. CONCLUSION

In this project, an adaptive and real-time payment fraud detection system has been developed using machine learning techniques. The system effectively combines the Isolation Forest algorithm for anomaly detection with the ADWIN algorithm for concept drift detection, providing a robust solution for identifying fraudulent transactions. The use of Isolation Forest enables the system to detect anomalies without requiring labeled data, making it highly suitable for real-world applications where fraud data is limited. The integration of ADWIN ensures that the system can adapt to changing transaction patterns, addressing one of the major challenges in fraud detection—concept drift. The system is designed with a modular architecture, allowing for scalability and easy integration with existing financial systems. The Django-based web interface provides a user-friendly platform for real-time transaction analysis, making the system practical and accessible.

Experimental analysis demonstrates that the system can accurately detect fraudulent transactions while maintaining low false-positive rates. The inclusion of confidence scores enhances interpretability, enabling better decision-making. However, there are opportunities for further improvement. Future work may include integrating deep learning models such as autoencoders, incorporating additional features like user behavior patterns, and implementing automated retraining mechanisms triggered by drift detection. In conclusion, the proposed system provides a scalable, adaptive, and efficient solution for real-time fraud detection. It addresses key limitations of traditional systems and contributes to the advancement of intelligent financial security systems.

REFERENCES

1. Bifet, A., & Gavaldà, R. (2007). Learning from time-changing data with adaptive windowing (ADWIN).
2. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation Forest.
3. Togbe, M. U., et al. (2021). Anomaly Detection Using Isolation in Concept-Drifting Data Streams
4. Al-Daoud, K., et al. (2025). Robust AI for Financial Fraud Detection
5. Gupta, A., et al. (2025). Real-Time Fraud Detection Using Adaptive Models
6. Patel, A., et al. (2026). FL-MalDrift Framework
7. Entropy-Based Concept Drift Detection (2024)
8. Isolation Forest for Data Clustering and Anomaly Detection (2024)
9. Model Retraining with Concept Drift Detection (2025)
10. ADA-ADF Drift-Aware Framework (2025)
11. Evolutionary Concept Drift Detection (2024)
12. ROSFD: Streaming Fraud Detection Framework (2025)
13. Semi-Supervised Fraud Detection (2025)
14. Deep Learning for Fraud Detection Survey (2024)
15. Hybrid Drift Detection Systems in FinTech (2025)