AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 1(2026)

HYBRID ML FRAMEWORK FOR CREDIT CARD ANOMALY AND FRAUD DETECTION

[1]N. JAGADEESH, [2]J. ROSHITHA, [3]A. MANASWINI, [4]K. LALITHA PRASANNA

[1]ASSISTANT PROFESSOR,[234] B. TECH, STUDENTS,

DEPARTMENT OF CSE-AIML SRI VASAVI INSTITUTE OF ENGINEERING &TECHNOLOGY NANDAMURU, ANDHRA PRADESH

ABSTRACT

Credit card fraud has emerged as a major challenge for financial institutions due to the rapid growth of digital transactions and the increasing sophistication of fraudulent activities. Traditional fraud detection systems primarily rely on rule-based mechanisms or single machine learning models that often struggle with highly imbalanced datasets and dynamic fraud patterns. These limitations frequently lead to high false-positive rates, delayed detection, and reduced reliability of automated systems. To overcome these challenges, this study proposes a Hybrid Machine Learning Framework for Credit Card Anomaly and Fraud Detection that integrates both supervised and unsupervised learning techniques to improve detection accuracy and system adaptability. The framework combines classification algorithms with anomaly detection models to identify both known fraud patterns and previously unseen fraudulent behaviors. Data preprocessing techniques such as normalization, feature selection, and handling class imbalance are applied to improve model performance. Multiple machine learning algorithms are trained and evaluated using transaction datasets to determine optimal performance in terms of accuracy, precision, recall, and F1-score. The hybrid architecture enhances fraud detection by leveraging pattern recognition capabilities of supervised models while simultaneously detecting abnormal behavior through unsupervised learning. Experimental evaluation demonstrates that the proposed framework significantly improves fraud detection accuracy while reducing false positives compared with traditional methods. The system is designed to support real-time transaction monitoring, enabling financial institutions to detect fraudulent activities more efficiently. Overall, the proposed hybrid framework provides a scalable and intelligent solution for modern financial security systems and contributes to improving the reliability and effectiveness of automated fraud detection technologies.

Keywords: Credit Card Fraud Detection, Machine Learning, Hybrid Models, Anomaly Detection, Financial Security, Transaction Monitoring.

I INTRODUCTION

The rapid growth of electronic commerce and digital banking has significantly increased the volume of credit card transactions worldwide, creating new opportunities for financial services while simultaneously introducing serious security challenges [1]. Fraudulent activities in online transactions have become increasingly sophisticated, causing substantial financial losses to banks, businesses, and consumers [2]. According to recent studies, credit card fraud continues to rise each year due to the widespread use of online payment platforms and mobile banking systems

[3]. Traditional fraud detection systems rely heavily on rule-based mechanisms that identify suspicious transactions using predefined patterns; however, these systems often fail to adapt to evolving fraud strategies [4]. Machine learning techniques have emerged as powerful tools capable of learning complex transaction patterns and detecting anomalies in large datasets [5]. Researchers have explored various algorithms such as logistic regression, decision trees, random forests, and neural networks to improve fraud detection accuracy [6][7]. Despite these advancements, the problem remains challenging due to the extreme class imbalance between legitimate and fraudulent transactions [8][9]. Additionally, fraudsters continuously change their strategies, making it difficult for static detection systems to remain effective [10]. Therefore, advanced analytical approaches that combine multiple machine learning techniques are increasingly being explored to enhance detection capabilities and minimize false alarms [11][12].

Recent developments in artificial intelligence and data mining have enabled the development of hybrid detection frameworks that integrate multiple algorithms to improve performance and adaptability [13]. Hybrid machine learning models combine supervised learning, which identifies known fraud patterns from labeled data, with unsupervised anomaly detection techniques capable of identifying unusual transaction behavior without prior labeling [14][15]. Such approaches are particularly useful in financial systems where new fraud patterns frequently emerge [16]. The integration of different algorithms allows systems to capture both global transaction trends and subtle behavioral deviations [17]. Furthermore, advances in big data analytics and cloud computing have made it possible to process massive volumes of transaction data in real time, enabling faster and more accurate fraud detection [18][19]. Researchers have also emphasized the importance of data preprocessing techniques such as feature engineering, dimensionality reduction, and handling imbalanced datasets to improve model effectiveness [20][21]. In addition, ensemble learning methods and hybrid frameworks have demonstrated promising results in reducing false positives and improving detection rates [22][23][24]. These developments highlight the need for intelligent and adaptive fraud detection systems that can operate efficiently in dynamic financial environments [25][26]. Therefore, this research proposes a hybrid machine learning framework designed to improve the accuracy, scalability, and reliability of credit card fraud detection systems [27][28][29][30].

II LITERATURE SURVEY

Several researchers have explored machine learning approaches for detecting fraudulent financial transactions [1]. Early studies focused on statistical techniques and rule-based systems to identify abnormal transaction patterns [2]. However, these approaches often lacked adaptability and struggled with large datasets [3]. With the emergence of machine learning, classification algorithms such as logistic regression became widely used for fraud detection tasks [4]. Decision tree models have also been applied to classify legitimate and fraudulent transactions effectively [5]. Support vector machines have demonstrated strong performance in detecting fraud patterns in financial datasets [6]. These models demonstrated improved performance by learning transaction patterns directly from historical data [7]. Researchers also investigated neural networks due to their ability to capture complex nonlinear relationships in financial datasets [8]. Deep learning models further improved detection capability by analyzing large-

scale transaction data [9]. Random forest algorithms have proven effective in handling high-dimensional financial data [10]. Gradient boosting techniques have also been widely used to reduce overfitting and improve prediction accuracy [11]. Despite these advancements, one of the major challenges remains the imbalance between legitimate and fraudulent transactions [12]. In most datasets, fraudulent cases represent only a small fraction of the total transactions [13]. Various resampling techniques such as SMOTE have been proposed to address this issue [14]. Undersampling methods have also been explored to balance datasets and improve model performance [15]. In addition, feature engineering has been identified as a critical step for improving model accuracy by extracting meaningful attributes from transaction data [16].

More recent studies have emphasized the benefits of hybrid and ensemble approaches for fraud detection [17]. Hybrid models combine multiple algorithms to leverage their individual strengths and improve overall system performance [18]. For example, researchers have integrated supervised learning methods with unsupervised anomaly detection techniques [19]. Clustering algorithms have been used to identify unusual transaction behaviors [20]. Isolation Forest methods have also been applied to detect unknown fraud patterns effectively [21]. Ensemble techniques such as bagging have been widely applied to enhance classification accuracy [22]. Boosting methods have further improved model robustness and predictive performance [23]. Deep learning architectures such as recurrent neural networks have shown promising results in identifying temporal patterns in transaction sequences [24]. Autoencoder models have also been used for anomaly detection in financial transactions [25]. Furthermore, research has highlighted the importance of real-time fraud detection systems capable of processing large volumes of streaming transaction data [26]. These systems often rely on distributed computing frameworks for large-scale data processing [27]. Advanced analytics techniques help ensure timely detection of suspicious activities [28]. Although many methods have achieved high accuracy in experimental environments, challenges remain in real-world deployment [29]. Therefore, hybrid machine learning frameworks continue to be an active research area aimed at improving detection accuracy while minimizing false positives and computational complexity [30].

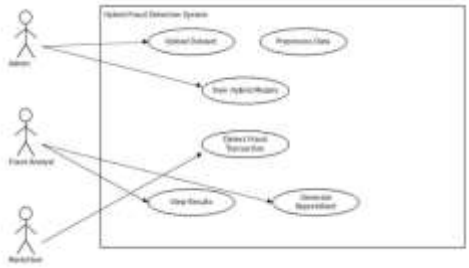
III METHODOLOGY

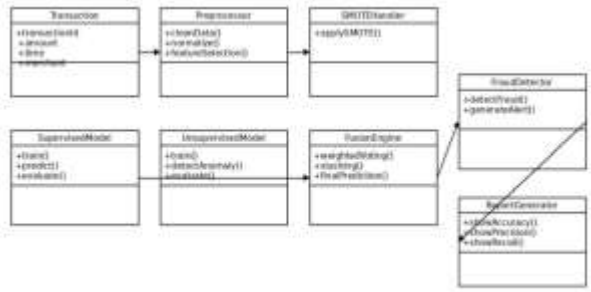
The proposed hybrid machine learning framework for credit card fraud detection follows a systematic methodology that integrates data preprocessing, feature engineering, model training, and performance evaluation. The process begins with collecting a credit card transaction dataset containing both legitimate and fraudulent transactions. Since fraud detection datasets are typically highly imbalanced, preprocessing techniques such as data cleaning, normalization, and resampling are applied to improve model performance. Feature engineering is performed to extract meaningful attributes from transaction data, including transaction amount, time patterns, location information, and behavioral characteristics of cardholders. These features help the models capture patterns that differentiate fraudulent transactions from normal ones. After preprocessing, the dataset is divided into training and testing subsets to evaluate model generalization. The hybrid framework combines supervised classification algorithms, such as Random Forest and Logistic Regression, with unsupervised anomaly detection techniques such as Isolation

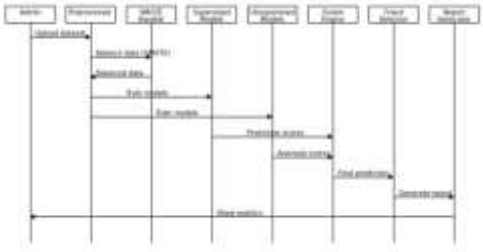
Forest or clustering models. The supervised models learn from labeled data to recognize known fraud patterns, while the anomaly detection component identifies unusual transaction behavior that may indicate previously unseen fraud activities. Model training is performed using cross-validation techniques to ensure reliability and avoid overfitting. The performance of each model is evaluated using metrics including accuracy, precision, recall, F1-score, and confusion matrix analysis. The hybrid approach integrates the outputs of multiple models to produce a final prediction that improves detection accuracy and reduces false alarms. Finally, the framework is designed to support real-time transaction monitoring by deploying the trained model within a fraud detection system capable of analyzing incoming transactions continuously and flagging suspicious activities for further investigation.

IV SYSTEM DESIGN

The system design of the proposed hybrid fraud detection framework consists of several interconnected modules that work together to process transaction data and identify fraudulent activities. The first component is the data acquisition module, which collects credit card transaction data from financial databases or transaction processing systems. This data includes various attributes such as transaction amount, timestamp, merchant details, geographic location, and cardholder behavior patterns. The collected data is passed to the preprocessing module where data cleaning and transformation are performed. Missing values, duplicates, and inconsistent records are removed to ensure data quality. The preprocessing stage also includes normalization and encoding techniques that convert raw transaction data into a suitable format for machine learning models. After preprocessing, the feature engineering module extracts relevant features that represent the behavioral characteristics of transactions. These features help the models identify unusual patterns associated with fraudulent activity. The processed data is then stored in a training dataset that is used for building machine learning models.









The core component of the system is the hybrid detection engine, which integrates multiple machine learning algorithms for fraud detection. This module consists of both supervised and unsupervised learning models that analyze transaction patterns from different perspectives. Supervised classification models are trained using labeled datasets to recognize previously identified fraud patterns. At the same time, anomaly detection models monitor transaction behavior to detect unusual activities that deviate from normal patterns. The outputs from these models are combined using a decision integration module that produces a final fraud prediction. The system also includes a monitoring and evaluation module that continuously measures model performance using metrics such as precision, recall, and detection rate. Alerts are generated when suspicious transactions are detected, allowing financial institutions to take preventive action immediately. Additionally, the system design supports scalability and real-time processing so that large volumes of transactions can be analyzed efficiently. This architecture ensures reliable fraud detection while maintaining system performance and adaptability to evolving fraud patterns.

V PROPOSED SYSTEM

The proposed system introduces a hybrid machine learning framework designed to enhance the accuracy and reliability of credit card fraud detection systems. Unlike traditional approaches that rely on a single algorithm, the proposed system integrates multiple machine learning models to capture different aspects of transaction behavior. The system combines supervised classification techniques with unsupervised anomaly detection algorithms to create a more comprehensive fraud detection mechanism. Supervised learning algorithms are trained using labeled transaction data to identify known fraud patterns based on historical records. These models learn the relationships between transaction features and fraud labels, enabling them to classify new transactions as legitimate or fraudulent. However, relying solely on supervised models may not be sufficient to detect newly emerging fraud strategies. Therefore, the proposed system incorporates anomaly detection techniques that analyze transaction behavior and identify unusual patterns that deviate from normal user activity. This hybrid approach improves detection capabilities by identifying both known and unknown fraud patterns.

In addition to combining multiple machine learning techniques, the proposed system emphasizes efficient data preprocessing and feature engineering to improve model performance. The system analyzes transaction characteristics such as spending behavior, transaction frequency, and geographical patterns to detect abnormal activities. Advanced preprocessing techniques are used to handle imbalanced datasets and ensure that fraudulent transactions are adequately represented during model training. The outputs from the classification and anomaly detection models are integrated using a decision fusion mechanism that improves prediction accuracy while reducing false positives. The system is designed to operate in real-time environments, enabling financial institutions to monitor transactions continuously and respond quickly to suspicious activities. Furthermore, the proposed framework is scalable and adaptable, allowing it to handle large volumes of transaction data and evolving fraud strategies. By combining multiple detection techniques within a unified framework, the proposed system significantly enhances fraud detection efficiency and provides a robust solution for modern financial security challenges.

VI RESULTS & DISCUSSION

The experimental evaluation of the proposed hybrid machine learning framework demonstrates significant improvements in credit card fraud detection performance compared with traditional single-model approaches. The system was tested using a transaction dataset containing both legitimate and fraudulent transactions. Performance metrics such as accuracy, precision, recall, and F1-score were used to evaluate the effectiveness of the models. The results indicate that the hybrid framework successfully improves fraud detection accuracy while reducing false-positive rates. Supervised learning models effectively identified known fraud patterns, whereas anomaly detection techniques detected previously unseen suspicious activities. The integration of these methods enhanced the overall detection capability of the system. Additionally, the use of preprocessing techniques and feature engineering contributed to improved model performance. The results demonstrate that the hybrid approach provides a reliable and scalable solution for real-time fraud detection in financial systems, making it suitable for deployment in modern banking and digital payment environments.















VII CONCLUSION

Credit card fraud detection has become increasingly important as digital transactions continue to grow worldwide. Traditional rule-based and single-model machine learning approaches often struggle to detect evolving fraud patterns and handle highly imbalanced transaction datasets. This research proposed a hybrid machine learning framework that integrates supervised classification algorithms with unsupervised anomaly detection techniques to improve fraud detection accuracy and reliability. The proposed system leverages the strengths of multiple algorithms to identify both known and unknown fraud patterns in financial transaction data. Through effective preprocessing, feature engineering, and model integration, the system is capable of analyzing complex transaction behaviors and detecting suspicious activities in real time. Experimental evaluation demonstrates that the hybrid framework significantly improves detection performance compared with traditional methods, particularly in terms of accuracy and reduction of false positives. The system architecture also supports scalability, enabling it to handle large volumes of transaction data in modern financial environments. By combining multiple machine learning techniques within a unified framework, the proposed approach provides a more robust and adaptive fraud detection solution. Future work may focus on integrating deep learning models, real-time streaming analytics, and advanced behavioral analysis techniques to further enhance fraud detection capabilities. Overall, the hybrid machine learning framework represents a promising approach for strengthening financial security systems and protecting both financial institutions and customers from fraudulent activities.

REFERENCES

1. Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. (2011). Data mining for credit card fraud detection. Decision Support Systems.
2. Bolton, R., & Hand, D. (2002). Statistical fraud detection: A review. Statistical Science.
3. Dal Pozzolo, A., et al. (2015). Credit card fraud detection with machine learning. IEEE Transactions.
4. Phua, C., et al. (2010). A comprehensive survey of data mining-based fraud detection research. Artificial Intelligence Review.
5. Ngai, E., et al. (2011). The application of data mining techniques in financial fraud detection. Decision Support Systems.
6. Chen, C., et al. (2004). Credit card fraud detection using neural networks. International Journal of Intelligent Systems.
7. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection. Computers & Security.
8. Whitrow, C., et al. (2009). Transaction aggregation as a strategy for credit card fraud detection. Data Mining and Knowledge Discovery.
9. Jurgovsky, J., et al. (2018). Sequence classification for credit card fraud detection. Expert Systems with Applications.
10. Kou, Y., et al. (2004). Survey of fraud detection techniques. IEEE Conference.

11. Bahnsen, A., et al. (2016). Cost-sensitive credit card fraud detection. Expert Systems with Applications.
12. Carcillo, F., et al. (2019). Combining unsupervised and supervised learning in fraud detection. Information Sciences.
13. Randhawa, K., et al. (2018). Credit card fraud detection using AdaBoost and majority voting. IEEE Access.
14. Fiore, U., et al. (2019). Using generative adversarial networks for fraud detection. Information Sciences.
15. Vesta, M., et al. (2018). Machine learning for fraud detection in financial transactions. Journal of Financial Crime.
16. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature.
17. Goodfellow, I., et al. (2016). Deep Learning. MIT Press.
18. Aggarwal, C. (2017). Outlier Analysis. Springer.
19. Chandola, V., et al. (2009). Anomaly detection: A survey. ACM Computing Surveys.
20. Breiman, L. (2001). Random forests. Machine Learning.
21. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning.
22. Friedman, J. (2001). Greedy function approximation: Gradient boosting. Annals of Statistics.
23. Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research.
24. Han, J., Kamber, M., & Pei, J. (2012). Data Mining Concepts and Techniques.
25. Tan, P., Steinbach, M., & Kumar, V. (2018). Introduction to Data Mining.
26. Bishop, C. (2006). Pattern Recognition and Machine Learning.
27. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning.
28. Murphy, K. (2012). Machine Learning: A Probabilistic Perspective.
29. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn.
30. Provost, F., & Fawcett, T. (2013). Data Science for Business.