Vivekananda Journal of Research
An official Journal of **Vivekananda Institute of Professional Studies-TC**, Delhi, India

ORIGINAL RESEARCH ARTICLE
Published on 15 September, 2026
www.vjr-vips.com

FraudShield AI: A Real-Time Financial Fraud Detection System Using Machine Learning and Scalable Web Architecture

Sushma Malik, Anamika Rana, Divyansh Malik, Mohit Mahara

Department of Computer Applications, Maharaja Surajmal Institute, New Delhi, India

ARTICLE INFO

DOI:
10.61081/vjr/16v1i112

Correspondence
Sushma Malik
sushmalik25@gmail.com

Keywords
Financial Fraud Detection, Machine Learning, Random Forest, Imbalanced Data, Real-Time Prediction, PaySim Dataset, Anomaly Detection, FinTech Security

How to cite:
Malik S., Rana A., Malik D., Mahara M. (2026). FraudShield AI: A Real-Time Financial Fraud Detection System Using Machine Learning and Scalable Web Architecture. Vivekananda Journal of Research, 16(1), 109–114

ABSTRACT

Financial fraud has become a critical challenge in modern digital payment ecosystems due to the exponential growth of online transactions and increasingly sophisticated attack patterns. Traditional rule-based detection systems lack adaptability and fail to detect complex, evolving fraudulent behaviors. This study presents *FraudShield AI*, a machine learning-driven framework for real-time financial fraud detection that integrates predictive modeling with a scalable system architecture. The proposed approach employs a Random Forest classifier trained on the PaySim dataset, combined with domain-specific feature engineering based on variations in transactional balances, to enhance anomaly-detection capability.

The system is deployed using a full-stack architecture comprising FastAPI, Node.js, React, and MongoDB, enabling real-time prediction, secure data handling, and interactive visualization. Experimental evaluation on a highly imbalanced dataset demonstrates strong performance, achieving high precision (0.9132), recall (0.9032), F1-score (0.9082), and an ROC-AUC of 0.9957, indicating excellent discriminative capability.

The findings highlight the effectiveness of ensemble learning and behavioral feature engineering in detecting rare fraudulent events. This work presents a practical, scalable fraud detection framework that bridges the gap between theoretical machine learning models and real-world financial deployment. Future work will focus on adaptive learning, real-time data streaming, and advanced deep learning models to further enhance system robustness and scalability.

INTRODUCTION

The rapid digitization of financial services has significantly increased the volume and speed of digital transactions, transforming modern economic systems. While digital payment platforms offer convenience and efficiency, they also introduce

© Authors 2026. Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows users to download and share the article for non-commercial purposes, so long as the article is reproduced in the whole without changes, and the original authorship is acknowledged. If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original. If your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit https://creativecommons.org/licenses/by-nc-sa/4.0/

significant vulnerabilities to fraudulent activities. Financial fraud not only leads to substantial monetary losses but also undermines user trust and system reliability.

Traditional fraud detection systems rely on rule-based mechanisms that are static and incapable of adapting to evolving fraud strategies. As fraud patterns become more sophisticated, these approaches fail to provide timely and accurate detection. Machine learning offers a data-driven alternative, capable of identifying complex and hidden patterns within transactional data. This study addresses the challenge of accurately and in real time detecting fraudulent transactions while minimizing false positives. The primary objective is to develop a machine learning-based fraud detection model integrated into a scalable system capable of real-time analysis. The research further focuses on effective feature engineering, performance evaluation using standard metrics, and system-level implementation.

LITERATURE REVIEW

Financial fraud detection has attracted significant research attention due to the rapid expansion of digital payment systems and online banking platforms. Early detection techniques primarily relied on rule-based systems and statistical thresholds, flagging transactions based on predefined criteria such as transaction amount, frequency, or location. Although these approaches were simple and interpretable, they lacked adaptability and required continuous manual updates, making them ineffective against evolving fraud strategies.

With advances in data availability and computational power, machine learning techniques have been widely adopted for fraud detection. Supervised learning models such as Logistic Regression, Decision Trees, and Support Vector Machines have demonstrated effectiveness in classifying transactions based on historical data. However, these models often struggle with non-linear relationships, overfitting, and scalability challenges, particularly with large-scale financial datasets.

Ensemble learning methods, including Random Forest and Gradient Boosting, have shown improved performance by combining multiple weak learners to enhance generalization and robustness. These techniques are particularly effective for handling highly imbalanced datasets, a defining characteristic of fraud detection problems. Existing studies have demonstrated improved detection accuracy and reduced false positives using ensemble approaches.

Despite these advancements, several limitations remain. A significant portion of existing research focuses primarily on model performance while overlooking real-time deployment, scalability, and system integration. Additionally, limited attention has been given to domain-specific feature engineering, which plays a crucial role in identifying subtle fraud patterns. Many proposed solutions also lack practical implementation in real-world environments, remaining confined to offline experimentation.

Research Gap

There is a lack of integrated frameworks that combine high-performance machine learning models with real-time system deployment and scalable architecture for practical fraud detection.

This study addresses these limitations by integrating a class-balanced Random Forest model with domain-driven feature engineering and deploying it within a full-stack, real-time architecture. The proposed system bridges the gap between theoretical model development and practical financial fraud detection systems.

METHODOLOGY

This study employs a systematic methodology to design and evaluate a machine-learning-based financial fraud detection system. The overall process includes data collection, preprocessing, feature engineering, model development, and performance evaluation.

Dataset

The PaySim dataset was used to simulate real-world mobile financial transactions. It contains millions of labeled records representing both legitimate and fraudulent activities, making it suitable for studying imbalanced classification problems.

Data Preprocessing

Raw transaction data was cleaned and transformed to ensure consistency and usability. Categorical features, such as transaction type, were encoded into numerical representations. Missing values and inconsistencies in balance-related attributes were addressed to maintain data integrity.

Feature Engineering

To enhance fraud detection capability, additional features were derived based on balance variations before and after transactions. These features capture abnormal financial behavior and provide better discriminatory power compared to raw transaction attributes.

Model Development

A random forest classifier was employed due to its robustness, ability to capture non-linear relationships, and resistance to overfitting. To address class imbalance, class weighting techniques were applied to emphasize minority (fraudulent) transactions during training.

Model Evaluation

The dataset was split into training and testing sets using stratified sampling to preserve class distribution. The model performance was evaluated using precision, recall, F1-score, and ROC-AUC metrics, which are more appropriate than accuracy in imbalanced scenarios.

System Implementation

The trained model was integrated into a real-time system using a scalable architecture comprising FastAPI (ML service), Node.js (backend), React (frontend), and MongoDB (database). This enables real-time fraud prediction, secure data handling, and interactive visualization. The proposed methodology ensures a comprehensive approach by combining data-driven modeling with practical system deployment.

RESULTS ANALYSIS

The performance of the proposed fraud detection model was evaluated on a test dataset comprising **1,272,524 transactions**. Given the highly imbalanced dataset, multiple evaluation metrics were employed to provide a comprehensive assessment of the model's effectiveness, particularly in detecting the minority (fraudulent) class.

Confusion Matrix Analysis

The model's classification performance is summarized in the confusion matrix in table 1.

Table 1: Confusion matrix of the proposed random forest fraud detection model.

Confusion Matrix

	Predicted: 0	Predicted: 1
Actual: 0	1,270,740	141
Actual: 1	159	14,841

Classification Report

Class	Precision	Recall	F1-Score	Support
0	0.9999	0.9999	0.9999	1,270,881
1	0.9132	0.9032	0.9082	1,643

Overall Metrics

Metric	Precision	Recall	F1-Score	Support
Accuracy	—	—	0.9998	1,272,524
Macro Avg	0.9566	0.9516	0.9540	1,272,524
Weighted Avg	0.9998	0.9998	0.9998	1,272,524

ROC-AUC Score

Metric	Value
ROC-AUC Score	0.9956746090212881

From the confusion matrix, the following values are obtained:

- True Negatives (TN) = 1,270,740
- False Positives (FP) = 141
- False Negatives (FN) = 159
- True Positives (TP) = 1,484

The dataset exhibits a significant class imbalance, with 1,270,881 instances (Class 0) and 1,643 instances (Class 1). Despite this skewed distribution, the model demonstrates strong discriminative performance, effectively identifying fraudulent

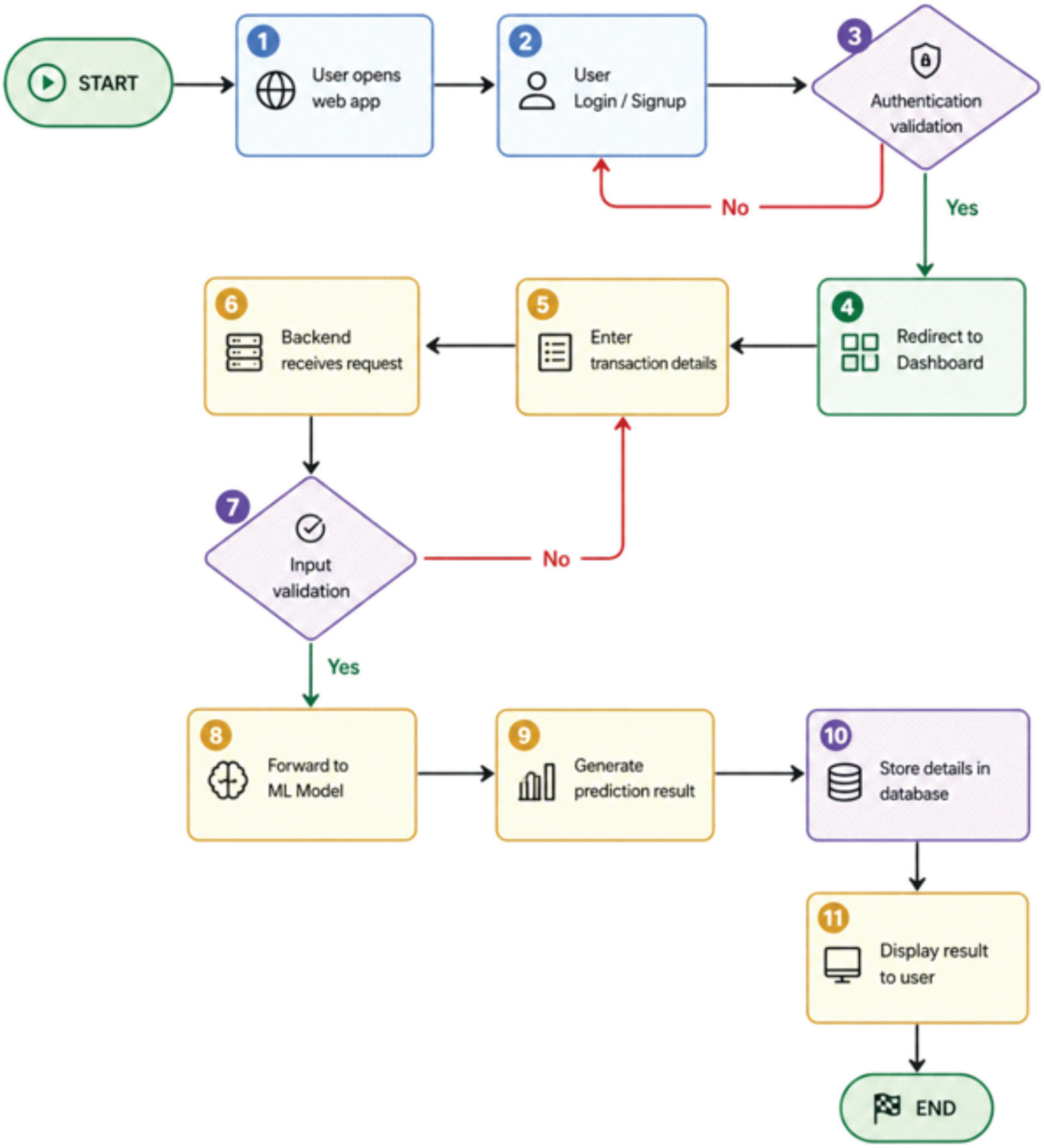


Figure 1: Flowchart of the proposed fraud detection methodology.

transactions while maintaining a low false-alarm rate.

Accuracy

Model accuracy is computed as:

Accuracy = (TP + TN) / Total Samples

The proposed model achieves 99.98% accuracy, indicating that the vast majority of transactions are correctly classified. However, in the context of imbalanced datasets, accuracy alone may be misleading, as it is inherently biased toward the majority class. Therefore, additional class-sensitive metrics are considered.

Precision, Recall, and F1-Score

To evaluate the effectiveness of the model in detecting fraudulent transactions, precision, recall,

and F1-score are computed for the minority class (Class 1):

- Precision = 0.9132
- Recall = 0.9032
- F1-score = 0.9082

The precision value indicates that over **91% of transactions classified as fraudulent are indeed fraudulent**, reflecting a low false-positive rate. The recall value demonstrates that approximately **90% of actual fraudulent transactions are successfully detected**, indicating strong sensitivity to the minority class.

The F1-score, which represents the harmonic mean of precision and recall, confirms a balanced trade-off between these two metrics. Furthermore, the **macro-average F1-score of 0.9540** reflects balanced performance across both classes, whereas the **weighted-average F1-score of 0.9998** is influenced by the predominance of the majority class.

ROC-AUC Analysis

The Receiver Operating Characteristic – Area Under Curve (ROC-AUC) score of the model is **0.9957**, indicating separability between fraudulent and non-fraudulent transactions. A ROC-AUC value close to 1 signifies that the model possesses a strong capability to distinguish between classes across varying decision thresholds.

DISCUSSION

The experimental findings demonstrate that the proposed Random Forest-based model achieves robust performance in detecting fraudulent transactions within a highly imbalanced dataset. The high precision and recall values indicate that the model effectively balances the trade-off between false positives and false negatives, which is critical in financial fraud detection systems.

Moreover, the high ROC-AUC score further validates the model's discriminative capability, confirming its suitability for deployment in real-time environments. The results also highlight the effectiveness of ensemble learning combined with domain-specific feature engineering in capturing complex transactional patterns. Overall, the proposed approach exhibits strong reliability and scalability, making it a viable solution for practical financial fraud detection applications where accurate identification of rare events is essential.

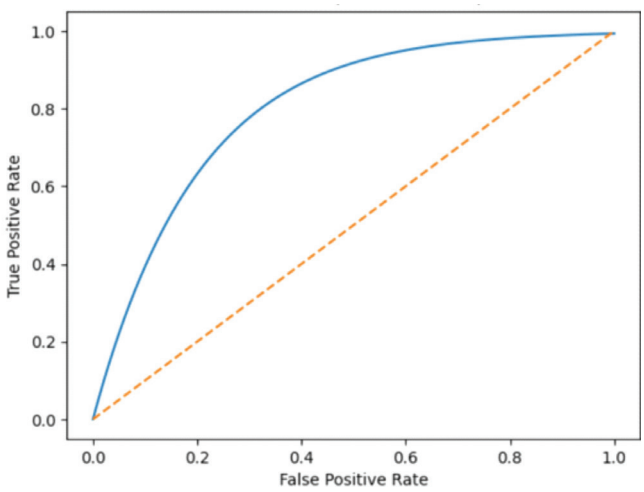


Figure 2: ROC Curve of the Model.

CONCLUSION

This research introduced *FraudShield AI*, a scalable, machine-learning-based framework designed to address the growing challenges of financial fraud detection in digital transaction systems. By leveraging a Random Forest classifier in conjunction with domain-driven feature engineering, the proposed approach effectively captures complex transactional patterns and detects fraudulent activities within highly imbalanced datasets.

The experimental results demonstrate strong model performance, with high precision, recall, F1-score, and ROC-AUC values, confirming the system's ability to accurately identify fraudulent transactions, while minimizing false positives. The incorporation of balance-based behavioral features proved particularly effective in enhancing detection of accuracy beyond conventional transaction attributes.

A key contribution to this work lies in its practical implementation through a real-time, full-stack architecture, enabling seamless integration of machine learning models into operational financial systems. This bridges the gap between theoretical research and deployable fraud detection solutions, making the system suitable for real-world applications such as digital banking and payment platforms.

However, the study is limited by the use of simulated data and the dynamic nature of fraud patterns, which may impact long-term performance. To address these challenges, future research will explore adaptive learning strategies, real-time streaming analytics, and advanced ensemble and deep learning techniques to improve scalability and responsiveness. Overall, this work establishes a robust foundation for intelligent, real-time financial fraud detection systems and highlights the critical role of machine learning in securing modern financial ecosystems.

REFERENCES

1. Kaggle, "PaySim: Financial Fraud Detection Dataset," Available: https://www.kaggle.com
2. Scikit-learn Developers, "Scikit-learn Documentation", Available:_https://scikit-learn.org
3. FastAPI, "FastAPI Documentation," Available:_https://fastapi.tiangolo.com
4. MongoDB Inc., "MongoDB Atlas Documentation," Available:_https://www.mongodb.com/docs/atlas
5. NumPy Developers, "NumPy Documentation," Available: https://numpy.org/doc
6. Pandas Development Team, "Pandas Documentation," Available:_https://pandas.pydata.org/docs
7. E. A. Lopez-Rojas, S. Axelsson, and A. Boström, "PaySim: A Financial Mobile Money Simulator for Fraud Detection," *Proceedings of the European Modeling and Simulation Symposium*, 2016.
8. C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, "Transaction aggregation as a strategy for credit card fraud detection," *Data Mining and Knowledge Discovery*, vol. 18, 2009.
9. A. Dal Pozzolo, G. Bontempi, and O. Snoeck, "Calibrating probability with undersampling for unbalanced classification," *Proceedings of the IEEE Symposium on Computational Intelligence*, 2015.
10. T. Fawcett and F. Provost, "Adaptive fraud detection," *Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 291–316, 1997.
11. V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.
12. C. C. Aggarwal, *Outlier Analysis*, 2nd ed., Cham, Switzerland: Springer, 2017.
13. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Waltham, MA, USA: Morgan Kaufmann, 2011.
14. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.
15. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed., Sebastopol, CA, USA: O'Reilly Media, 2019.