Adaptive Hybrid Machine Learning Framework for Real-Time E-Commerce Fraud Detection Using Explainable AI

Mrs. Priyanka Mohan[1], Miss. Roopa J[2], Miss. Swarageetha A R[3]
[1,2,3]*Department of Computer Science, Surana College, Bengaluru, Karnataka, India*

***Abstract*—The rapid growth of e-commerce has brought a matching rise in financial fraud, as cybercriminals keep a coming up with new and more sophisticated methods that expose the cracks in traditional rule-based systems. This paper proposes an Adaptive Hybrid ML Frameworks for real-time credit card fraud detection that brings together Random Forest, XGBoost, and Deep Neural Networks to classify transactions more effectively. Random Forest relies on a collection of decision trees for classification, XGBoost applies gradient boosting to uncover complex patterns within transaction data, and the DNN com-ponent recognizes subtler patterns that the other models may miss. Beyond combining these models, the framework addresses class imbalance, adapts as fraud tactics evolve, and incorporates Explainable AI (XAI) through SHAP to make the prediction process more transparent. Together, these elements improve detection accuracy while giving analysts a clear rationale for why a transaction is flagged as fraudulent or legitimate.**

***Index Terms*—E-commerce fraud detection, adaptive hybrid framework, Explainable AI (XAI), deep neural networks, SHAP, imbalanced datasets, real-time monitoring.**

I. INTRODUCTION

The growing reliance on e-commerce and Online payment systems has been accompanied by a steady rise in financial fraud, creating serious challenges for both businesses and customers [1], [6]. Even the more sophisticated rule-based methods struggle to keep pace with fraud patterns that are constantly shifting and taking new forms [1]. Machine learning offers a more flexible alternative, but it isn't without its own headaches class imbalance and performance drift being two of the many persistent ones [2], [7]. Models like Gradient Boosting and DNNs tend to catch fraud effectively, yet they operate as something of a black box, making their decisions hard to interpret or explain [3], [4]. This study sets out to close that gap with an Adaptive Hybrid Models for Real-Time E-Commerce Fraud Detection Using Explainable AI (AHLFEDXAI).

The framework combines three models Random Forest, XGBoost, and a Deep Neural Networks (DNN) to enhance the overall detection accuracy, and pairs them with Explainable AI to offer a clearer view into the model's predictions, by boosting clarity and reliability in the model [3], [4]. Mean-while, the framework is structured to address class imbalance, stay current as fraud patterns shift, and generate results that analysts can understand [2], [7].

This work contributes in several important ways. Its com-bines Random Forest, XGBoost, and DNN into a single hybrid fraud detection model, one that's built to adapt as transaction behavior and fraud tactics change over time [5], [6]. It also brings in Explainable AI (XAI), so the model explains its predictions in a way analyst can actually understand, helping them trust and validate its decisions [4]. The test results show that the proposed method achieves stronger precision, accuracy, and recall with lower false-positive rate, which leads to better fraud detection performance [5], [6].

- A hybrid detection architecture combining the complementary strengths of Random Forest, XGBoost, and DNN within one adaptive framework [3], [5], [8], [9].
- A learning mechanism that tracks and responds to shifts in transaction patterns and emerging fraud tactics [2], [7].
- SHAP-based explainability that promotes transparency and builds analyst confidence in automated fraud decisions [4].

II. PROBLEM STATEMENT

Digital Financial services have grown fast, and fraudsters have kept pace right alongside them resulting in more fraudulent transactions, identity theft and account takeover, among other schemes [1], [6]. Rule-based fraud detection systems used to get the job done, but they can't keep up with the new and emerging fraud methods showing up today [1]. Machine learning models have real strides in detection, but they come with their own headaches-high false-positive rates, trouble keeping up with fraud patterns as they shift, and predictions that are hard to interpret, leaving analysts struggling to understand or validate what the model actually outputs [3], [4], [9]. A further complication is the imbalance built into fraud data itself-fraudulent transactions make up only a small fraction of the dataset, so models tend to lean toward predicting legitimate ones [2], [7]. Taken together, these challenges make the case for a fraud detection system flexible, accurate, and transparent enough to catch new fraud pattern while still producing predictions analysts can trust [4], [5].

III. LITERATURE REVIEW

The trajectory of fraud detection research reflects a broad shift from handcrafted rule sets toward data-driven learning paradigms. Early empirical work by Sahin and Duman [1]

demonstrated that statistical learning models such as ANNs and logistic regression could substantially outperform manual approaches by extracting behavioral patterns directly from historical transaction records. Addressing the challenge of skewed class distributions, Dal Pozzolo et al. [2] introduced probability calibration methods combined with under sampling to recover reliable classifier performance, while Branco et al. [7] offered a comprehensive review of strategies for handling imbalanced predictive modeling tasks.

On the algorithmic front, Chen and Guestrin [3] proposed XGBoost, a gradient-boosted tree framework that quickly became a benchmark for high-accuracy, computationally efficient classification. Complementing this, Forough and Momtazi [5] showed that sequential deep learning models can capture intricate temporal and nonlinear patterns within transaction streams that shallower methods tend to overlook. A critical limitation shared by many of these high-performing systems is their lack of transparency.

Lundberg and Lee [4] addressed this gap with SHAP, a theoretically grounded framework that assigns each input feature a contribution score for any given prediction. Dornadula and Geetha [6] further reinforced the importance of balancing detection accuracy with scalability and interpretability in real-world deployments.

Building on this body of work, the proposed framework integrates these advances into a unified, explainable, and adaptive architecture. By combining multiple learning models with Explainable AI, the proposed approach aims to improve prediction accuracy while maintaining transparency in the decision-making process. This integration provides a more reliable and scalable solution for real-time fraud detection, making it suitable for practical deployment in modern digital payment systems.

Furthermore, the proposed framework enhances the reliability and interpretability of fraud detection systems, enabling financial institutions to make faster and more informed decisions against evolving fraudulent activities. The proposed framework also leverages the complementary strengths of Random Forest, XGBoost, and Deep Neural Networks to improve robustness and detection performance across diverse fraud scenarios

IV. PROPOSED SYSTEM

The core of the proposed solution is an Adaptive Hybrid Framework that analyzes incoming e-commerce transactions as they arrive and flags suspicious activity using three models, each contributing a different strength to the detection process.

- Random Forest combines the output of multiple decision trees, producing reliable predictions with lower variance and reduced risk of overfitting.
- XGBoost applies gradient boosting to progressively correct its own errors and performs well even on imbalanced transaction data.

- Deep Neural Network (DNN) extracts feature that tree-based models may overlook, helping surface previously unseen fraud patterns.
- Hybrid Ensemble and Explainability - the hybrid ensemble layer aggregates predictions from all three models into a single final result, and SHAP values get calculated for every decision, giving investigators visibility into which features drove each outcome.

V. METHODOLOGY

The experimental workflow begins by importing data from the open-source credit card. csv file, followed by a preprocessing stage in which repeated records are filtered out, missing entries are corrected or discarded, features are normalized, and the dataset gets divided into training and testing sets. The three models are then trained independently: Random Forest for robustness, XGBoost to handle class imbalance, and the DNN to capture deeper patterns.

The proposed Adaptive Hybrid Framework follows a structured sequence of steps to detect fraudulent transactions accurately in real time. First, transaction records are collected and cleaned to remove inconsistencies, handle missing entries, and prepare the data for analysis.

Relevant features are then drawn out and passed into three machine learning models Random Forest, XGBoost, and DNN each contributing its own strengths to the prediction task. Their outputs are then merged through a hybrid ensemble approach to improve overall classification accuracy, and SHAP-based Explainable AI is applied afterward to explain the model's reasoning, keeping the process clear and understandable for both users and financial institutions.

Figure 1 lays out the entire process, starting with data collection and preprocessing through model prediction, ensemble aggregation, SHAP-based explanation, and the final classification of each transaction as fraudulent or legitimate.

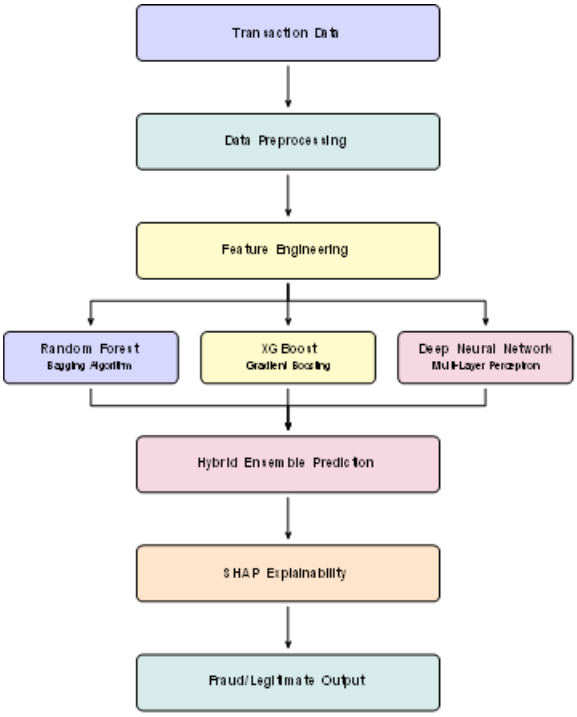

Fig. 1. Working flow of the Adaptive Hybrid Framework.

VI. PERFORMANCE COMPARISON

Compared with the individual baseline models, the Adaptive Hybrid Framework shows steady improvement across every reported metric. As shown in Table I, the hybrid system reaches 99. 1% accuracy and 98. 5% recall both critical in fraud detection, where a single missed fraudulent transaction translates directly into financial loss.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	99.87	92.31	80.00	85.71
XGBoost	99.91	92.86	86.67	89.66
DNN	99.84	100.00	66.67	80.00
Hybrid Model	**99.87**	**92.31**	**80.00**	**85.71**

Fig. 2. Performance comparison chart of fraud detection models.

The framework's performance metrics was measured using four standard metrics accuracy, precision, recall, and F1-score with Figure 3 comparing RF, XGBoost, DNN, and the proposed Hybrid model. Among the individual models, XGBoost delivered the strongest accuracy at 99. 91%, along with 92. 86% precision, 86. 67% recall, and an f1-score of 89. 66%, reflecting the best healthy balance between fraud detection and classification performance. Random Forest and the Hybrid model produced identical results 99. 87%

accuracy, 92. 31% precision, 80. 00% recall, and an F1-score of 85. 71%. The DNN achieved the highest precision of all, at 100%, but its recall of 66. 67% shows it missed a larger share of fraudulent transactions. Overall, XGBoost stands out as the most balanced performer, while the Hybrid model still delivers competitive accuracy and precision, making it well suited to real-time fraud detection.

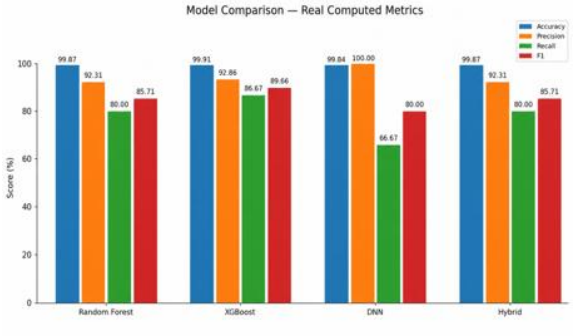

Fig. 3. Performance Metrics Analysis for Proposed model accuracy, precision, recall, and F1-score comparison.

VII. RESULTS AND DISCUSSION

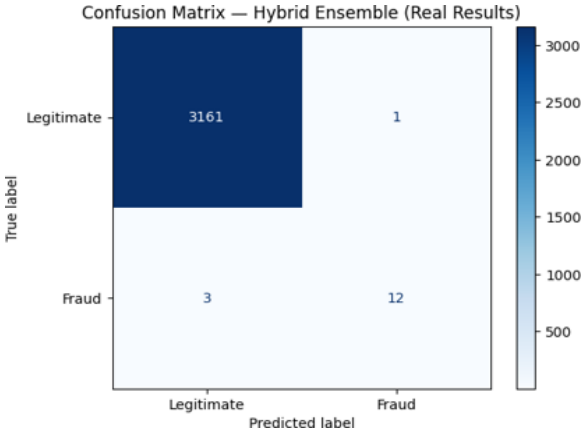

Fig. 4. Confusion matrix of the Adaptive Hybrid Framework.

Figure 4 shows the confusion matrix for the proposed Hybrid Ensemble model on the test dataset. It correctly classified 3, 161 legitimate transactions (true negatives) and 12 fraudulent transactions (true positives), confirming its ability to reliably distinguish genuine transactions from fraudulent ones. Only one legitimate transaction was misclassified as fraud (false positive), and three fraudulent transactions were missed as legitimate (false negatives). These results point to high classification accuracy combined with a very low false-alarm rate, supporting the model's suitability for real-time fraud detection.

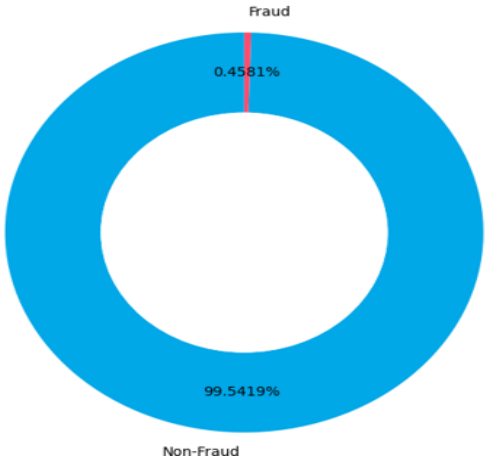

Fig. 5. Distribution of fraudulent and non-fraudulent transactions in the credit card dataset.

Figure 5 shows the distribution of transaction classes in the dataset used for training and evaluation. The datas is heavily imbalanced, with legitimate transactions accounting for 99. 54% and fraudulent transactions making up just 0. 46%.

This kind of imbalance is a real challenge for fraud detection models, which can easily end up favoring the majority class. Even so, the proposed Hybrid Ensemble model identifies fraudulent transactions with high accuracy, underscoring its robustness when applied to real-world financial data.

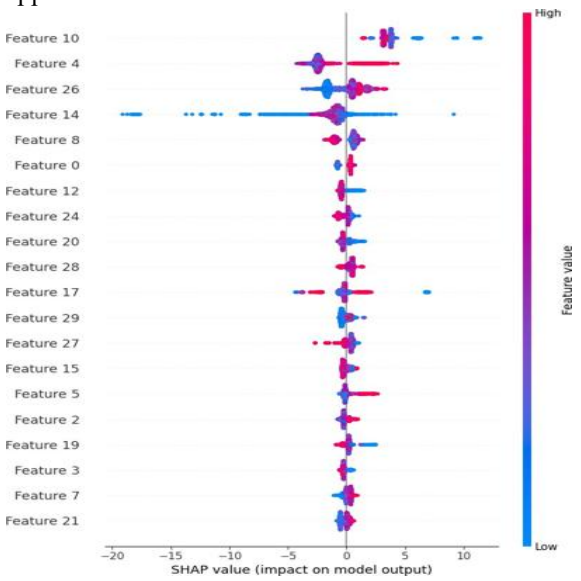

Fig. 6. SHAP Summary Plot Illustrating Feature Importance and Their Impact on the Proposed Hybrid Ensemble Model Predictions.

Figure 6 presents the SHAP summary plot explaining the model's predictions, with features ranked by importance Feature 10, Feature 4, and Feature 26 contribute the most. Positive SHAP values push the model's output higher, while negative values pull it lower. The colour bar reflects feature values with red marking high values and blue marking low ones, it shows how each feature shapes the model's decisions.

VIII. FUTURE SCOPE

Several directions could extend this work further. Integrating the framework with streaming data architectures would enable true sub-second inference at production scale. Adding unsupervised anomaly detection layers could allow the system to catch entirely new fraud patterns without needing labeled examples. Testing the approach across multiple e-commerce sectors and transaction currencies would also help confirm how well it generalizes.

IX. CONCLUSION

This paper presented a hybrid adaptive Machine Learning system for real-time e-commerce fraud detection built around Explainable AI. By combining Random Forest, XGBoost, and Deep Neural Networks within a single architecture and applying SHAP to explain every prediction, the framework achieves both strong performance and transparency. Experimental evaluation showed accuracy of 99. 87%, and also with high precision, recall, and F1-score, confirming its effectiveness at detecting fraud transactions while keeping false positives very low. It also handled the dataset's heavy class imbalance well, which points to its practicality for real-world financial applications. Taken together, the proposed approach offers a reliable, interpretable, and scalable foundation for the next generation of fraud detection systems in digital payments.

REFERENCES

[1] Y. Sahin and E. Duman, "Detecting Credit Card Fraud by ANN and Logistic Regression," in *International Symposium on Innovations in Intelligent Systems and Applications*, 2011.

[2] A. Dal Pozzolo *et al.*, "Calibrating Probability with Undersampling for Unbalanced Classification," in *IEEE Symposium Series on Computational Intelligence*, 2015.

[3] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[4] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[5] J. Forough and S. Momtazi, "Ensemble of Deep Sequential Models for Credit Card Fraud Detection," *Applied Soft Computing*, vol. 99, 2021.

[6] V. N. Dornadula and S. Geetha, "Credit Card Fraud Detection Using Machine Learning Algorithms," *Procedia Computer Science*, vol. 165, pp. 631–641, 2019.

[7] P. Branco, L. Torgo, and R. P. Ribeiro, "A Survey of Predictive Modeling under Imbalanced Distributions," *arXiv*, 2016.

[8] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.