Risk-Aware Machine Learning Framework for Fraud Detection in Financial Technology Ecosystems

Richard Berlien[1,*]

[1]Universitas Murni Teguh, Indonesia

ABSTRACT

Financial fraud has become one of the major challenges in modern financial technology ecosystems due to the increasing volume and complexity of digital transactions. Traditional fraud detection approaches often struggle to identify complex and dynamic fraud behavior patterns, particularly in noisy and imbalanced transaction environments. This study proposes a risk-aware machine learning framework for fraud detection in fintech ecosystems by integrating contextual risk indicators, behavioral transaction features, and explainable artificial intelligence techniques. The proposed framework incorporates feature engineering, SMOTE balancing, ensemble-based machine learning algorithms, and SHAP interpretability analysis to improve fraud classification performance and model transparency. The experiments were conducted using Random Forest, XGBoost, and LightGBM models on a synthetic fintech transaction dataset enhanced with realistic noise injection and interaction-based risk features. The experimental results demonstrated that LightGBM achieved the best overall performance with an accuracy of 97.75%, F1-score of 80.69%, and ROC-AUC of 0.8891. Furthermore, the ablation study confirmed that contextual risk-aware features, including device risk score, IP risk score, and risk interaction indicators, significantly improved fraud detection capability. The SHAP analysis also revealed that transaction amount, temporal transaction behavior, and interaction-based risk features substantially influenced fraud prediction outcomes. Overall, the findings indicate that combining contextual risk intelligence, behavioral analytics, and explainable machine learning techniques can significantly improve fraud detection performance and interpretability in modern fintech environments. The proposed framework provides a scalable and transparent approach for intelligent fraud analytics in digital financial systems.

Keywords Financial Fraud Detection, Risk-Aware Machine Learning, Explainable Artificial Intelligence, FinTech Analytics, LightGBM

INTRODUCTION

The rapid growth of digital financial services has significantly transformed modern financial technology ecosystems. The increasing adoption of online banking, electronic payment systems, mobile wallets, and digital financial platforms has accelerated transaction efficiency and financial accessibility worldwide [1]. However, the expansion of digital financial transactions has also increased the risk of financial fraud, cyberattacks, and unauthorized transaction activities [2]. Fraudulent activities in fintech environments continue to evolve in complexity, making fraud detection one of the major challenges for financial institutions and digital payment providers [3].

Traditional fraud detection systems commonly rely on rule-based approaches and static transaction monitoring techniques. Although these methods are effective for detecting predefined fraud patterns, they often fail to identify sophisticated and dynamic fraudulent behavior in large-

Submitted: 8 June 2026
Accepted: 20 August 2026
Published: 17 November 2026

Corresponding author
Richard Berlien,
richardberlien@gmail.com

Additional Information and Declarations can be found on page 330

DOI: 10.47738/jdmdc.v3i4.104

Copyright
2026 Berlien

Distributed under
Creative Commons CC-BY 4.0

How to cite this article: R. Berlien, "Risk-Aware Machine Learning Framework for Fraud Detection in Financial Technology Ecosystems," *J. Digit. Mark. Digit. Curr.*, vol. 3, no. 4, pp. 315-333, 2026.

scale digital transaction environments [4]. Moreover, fintech transaction datasets are typically characterized by class imbalance, noisy transaction behavior, and nonlinear fraud patterns, which reduce the effectiveness of conventional classification techniques. As a result, machine learning approaches have become increasingly important for intelligent fraud detection due to their capability to learn complex behavioral patterns from transaction data [5].

Recent studies have demonstrated the effectiveness of ensemble-based machine learning algorithms, such as Random Forest, XGBoost, and LightGBM, in fraud detection applications [6]. These approaches are capable of improving classification performance by handling nonlinear relationships and high-dimensional transaction features. In addition, explainable artificial intelligence techniques, including SHAP analysis, have gained significant attention because they improve model transparency and interpretability in fraud analytics systems [7]. Explainable models are particularly important in fintech applications, where transparency and trustworthiness are essential for practical deployment and regulatory compliance.

Several previous studies have explored fraud detection using transactional attributes and behavioral analytics. Existing approaches have focused on transaction amount analysis, anomaly detection, temporal transaction patterns, and supervised machine learning classification [8]. Other studies have integrated risk scoring systems to improve fraud prediction capability. However, most existing works primarily rely on isolated transaction features and limited contextual information, while interaction-based and contextual risk-aware indicators remain insufficiently explored [9]. Furthermore, many previous fraud detection studies emphasize predictive performance without providing adequate explainability regarding how contextual risk features influence fraud behavior [10].

Based on the existing literature, several research gaps can be identified. First, previous studies have not fully investigated the combined impact of contextual risk indicators, such as device risk score and IP risk score, within an interaction-based fraud detection framework. Second, the integration of explainable artificial intelligence techniques for interpreting nonlinear fraud behavior patterns remains limited. Third, many existing fraud detection models are evaluated using overly separable datasets that do not adequately represent realistic noisy fintech transaction environments [11]. Consequently, there is a need for a more robust and interpretable fraud detection framework that combines contextual risk intelligence, behavioral analytics, and explainable machine learning techniques.

To address these limitations, this study proposes a risk-aware machine learning framework for fraud detection in financial technology ecosystems. The proposed framework integrates contextual risk-aware indicators, behavioral transaction features, interaction-based feature engineering, SMOTE balancing, and explainable artificial intelligence

analysis using SHAP. The study evaluates the performance of Random Forest, XGBoost, and LightGBM models using a noisy and imbalanced fintech transaction dataset to simulate realistic fraud detection conditions. The main contributions of this study are summarized as follows. First, this research introduces a contextual risk-aware fraud detection framework that combines device risk score, IP risk score, and interaction-based risk indicators to improve fraud classification performance. Second, the study integrates explainable artificial intelligence techniques to provide transparent interpretation of fraud behavior patterns. Third, an ablation study is conducted to validate the contribution of contextual risk-aware features in improving fraud detection capability. Finally, the proposed framework provides a scalable and interpretable approach for intelligent fraud analytics in modern fintech ecosystems.

Literature Review

Financial fraud detection has become an important research area in financial technology ecosystems due to the rapid growth of digital financial transactions and online payment systems [12]. The increasing complexity of cyber fraud activities has encouraged researchers to develop intelligent fraud detection models capable of identifying suspicious transaction patterns in real time [13]. Traditional fraud detection approaches commonly rely on static rule-based systems and manual verification processes. However, these approaches are often ineffective in detecting sophisticated and adaptive fraud behavior in modern fintech environments [14].

Machine learning techniques have been widely adopted in fraud analytics because of their ability to learn complex and nonlinear relationships from transaction data [15]. Supervised learning algorithms such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine have been extensively utilized for fraud classification tasks [16]. Among these methods, ensemble learning approaches have shown superior performance due to their capability to improve model robustness and reduce overfitting [17].

Random Forest is one of the most commonly used ensemble learning algorithms in fraud detection studies. The algorithm combines multiple decision trees to improve prediction stability and classification accuracy [18]. Previous studies have shown that Random Forest performs effectively in handling noisy transaction datasets and high-dimensional financial data. However, the model may experience limitations when dealing with highly imbalanced fraud datasets and complex interaction-based fraud patterns.

Gradient boosting algorithms, including XGBoost and LightGBM, have recently gained significant attention in fraud detection research [19]. XGBoost utilizes boosting techniques to sequentially optimize classification performance and minimize prediction errors. Several studies have reported that XGBoost outperforms conventional machine

learning algorithms in fraud detection tasks due to its ability to capture nonlinear transaction relationships and feature interactions [20]. Similarly, LightGBM has demonstrated strong performance in large-scale financial analytics because of its computational efficiency and gradient-based optimization mechanism [21].

In addition to predictive performance, explainability has become an important aspect of machine learning-based fraud detection systems. Many fraud detection models operate as black-box systems, making it difficult to understand how fraud predictions are generated. Explainable artificial intelligence techniques, such as SHAP, provide interpretable insights into feature contributions and decision-making processes [22].

Behavioral analytics also plays a crucial role in modern fraud detection systems. Previous studies have shown that transaction amount, transaction timing, and user behavioral patterns significantly influence fraud probability. Fraudulent transactions are often associated with abnormal transaction values, unusual transaction frequency, and suspicious transaction timing. Consequently, integrating behavioral transaction indicators into machine learning frameworks can improve fraud detection capability and model generalization.

Another important aspect of fraud detection research is contextual risk analysis. Recent studies have explored the use of contextual information, such as device identification, IP address analysis, and network-based risk indicators, to improve fraud prediction performance. Device risk scores and IP-based risk indicators can provide valuable information regarding suspicious transaction environments and unauthorized access patterns. However, most previous studies primarily evaluate individual risk indicators independently, while interaction-based contextual risk modeling remains relatively underexplored.

Class imbalance is another major challenge in fraud detection datasets because fraudulent transactions typically represent only a small proportion of total financial transactions. Imbalanced datasets often cause machine learning models to become biased toward majority classes, resulting in poor fraud detection sensitivity. To address this issue, oversampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) have been widely applied to improve minority class representation and enhance fraud classification performance.

Despite the significant progress in fraud analytics research, several limitations still exist in current studies. Many previous approaches focus primarily on predictive accuracy while providing limited interpretability regarding fraud behavior patterns. In addition, interaction-based contextual risk features and realistic noisy transaction environments remain insufficiently investigated in existing fraud detection frameworks. Therefore, there is a need for a more comprehensive and explainable fraud detection framework that combines behavioral analytics, contextual

risk-aware indicators, interaction-based feature engineering, and interpretable machine learning techniques.

Methods

This study proposes a risk-aware machine learning framework for fraud detection in financial technology ecosystems. The proposed framework integrates contextual risk-aware indicators, behavioral transaction analytics, ensemble-based machine learning algorithms, and explainable artificial intelligence techniques to improve fraud detection capability and model interpretability. The overall research workflow consists of dataset preparation, noise injection, feature engineering, data preprocessing, machine learning classification, ablation study evaluation, and SHAP-based explainability analysis. The overall research architecture of the proposed framework is illustrated in figure 1.

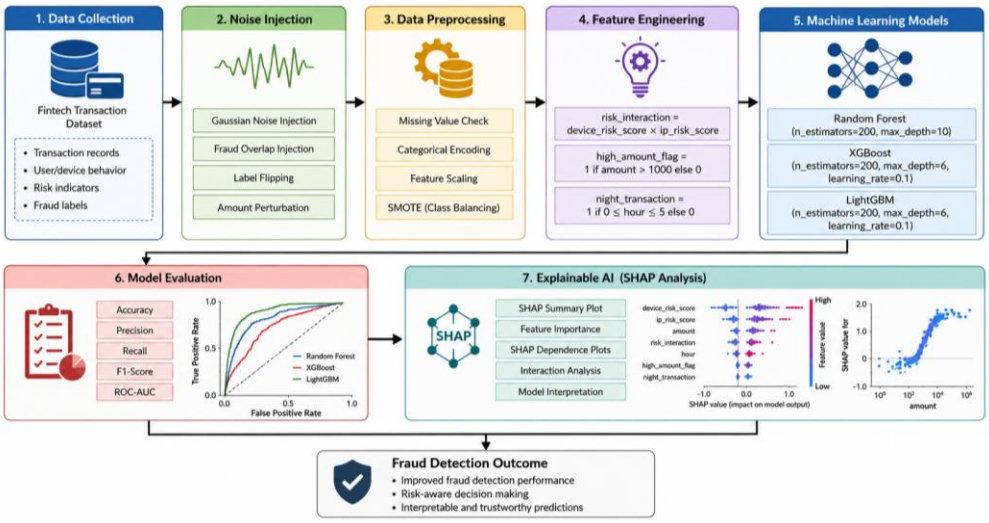

Figure 1 Proposed Risk-Aware Fraud Detection Framework

The framework begins with fintech transaction data collection followed by noise injection and preprocessing stages. Feature engineering techniques are then applied to generate contextual and behavioral fraud indicators. The processed dataset is subsequently evaluated using Random Forest, XGBoost, and LightGBM models. Finally, explainable artificial intelligence analysis using SHAP is performed to interpret fraud prediction behavior and feature contributions.

Dataset Description

The experiments were conducted using a synthetic fintech transaction dataset containing 10,000 transaction records. The dataset consists of numerical, categorical, and binary transaction attributes representing behavioral and contextual fraud indicators. The dataset includes both legitimate and fraudulent transactions with an imbalanced class distribution. Table 1 presents the detailed dataset variables utilized in this study.

Table 1 Dataset Description

Variable	Type	Description
amount	Numerical	Transaction amount
transaction_type	Categorical	Transaction payment method
merchant_category	Categorical	Merchant category classification
country	Categorical	Transaction country
hour	Numerical	Transaction hour
device_risk_score	Numerical	Device-based risk indicator
ip_risk_score	Numerical	IP-based risk indicator
risk_interaction	Numerical	Interaction between device and IP risk
high_amount_flag	Binary	High-value transaction indicator
night_transaction	Binary	Night transaction indicator
is_fraud	Binary	Fraud transaction label

The fraud distribution of the dataset is illustrated in figure 2. The dataset shows a significant class imbalance where non-fraudulent transactions dominate the transaction population. This imbalance condition motivates the use of SMOTE balancing during model training.

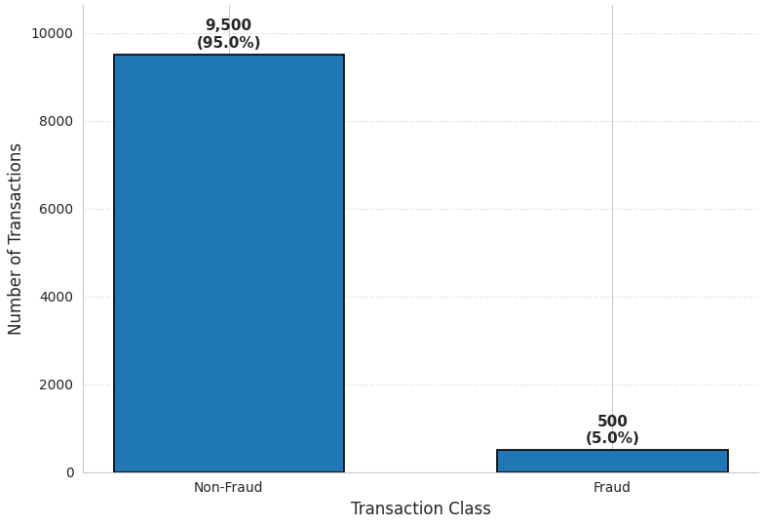

Figure 2 Fraud vs Non-Fraud Transaction Distribution

The correlation relationships among numerical variables are presented in figure 3. The heatmap analysis indicates that device risk score, IP risk score, and transaction amount have relatively strong correlations with fraudulent activity.

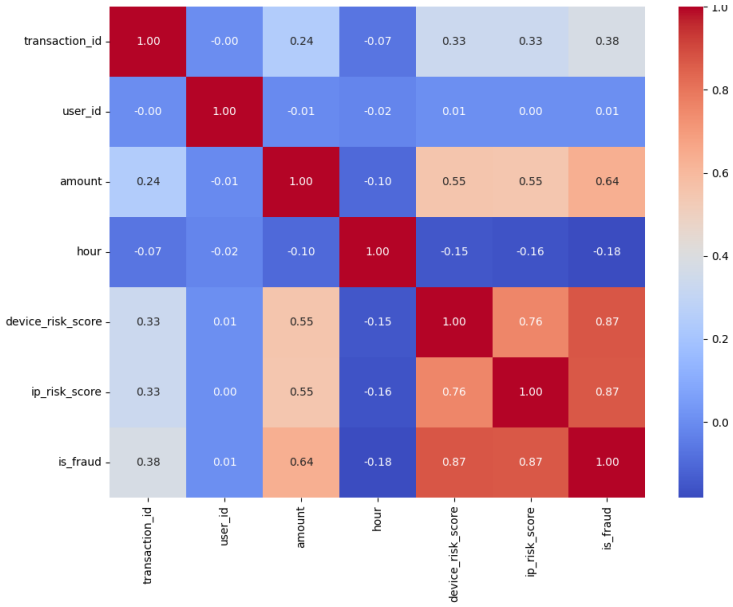

Figure 3 Correlation Heatmap

Noise Injection Strategy

As represented in table 2. To simulate realistic financial transaction environments, noise injection techniques were applied to the dataset. Financial fraud datasets in real-world scenarios are typically characterized by overlap between fraudulent and legitimate transactions, uncertain behavioral patterns, and noisy transaction attributes. Therefore, this study introduces multiple noise generation strategies to reduce unrealistic feature separability. Gaussian noise injection was applied using the following equation:

$$X_{noisy} = X + \epsilon \tag{1}$$

X represents the original feature value and ϵ denotes random Gaussian noise.

Table 2 Noise Injection Strategy

Technique	Purpose
Gaussian Noise Injection	Increase transaction uncertainty
Fraud Overlap Injection	Create realistic fraud overlap
Label Flipping	Simulate labeling uncertainty
Amount Perturbation	Reduce feature separabilit

The injected noise enables the machine learning models to learn more robust and generalized fraud behavior patterns rather than relying on perfectly separable transaction characteristics.

Feature Engineering

As represented in table 3, feature engineering was conducted to improve the capability of the fraud detection framework in capturing contextual

and behavioral fraud patterns. Several interaction-based and behavioral transaction features were generated based on domain knowledge and fraud analytics principles.

The interaction-based risk feature was computed using the following equation:

$$Risk_{interaction} = DeviceRisk \times IPRisk \quad (2)$$

The binary high-amount transaction feature was defined as:

$$HighAmount = 1, if\ Amount > 1000;\ otherwise\ 0 \quad (3)$$

The night transaction indicator was generated using:

$$NightTransaction = 1, if\ 0 \leq Hour \leq 5;\ otherwise\ 0 \quad (4)$$

Table 3 Feature Engineering

Feature	Formula / Logic
risk_interaction	device_risk_score × ip_risk_score
high_amount_flag	amount > 1000
night_transaction	hour between 00:00–05:00

The risk interaction feature was designed to capture the combined impact of suspicious device behavior and IP-based transaction risk. In addition, behavioral indicators such as high transaction amounts and night-time transactions were incorporated to improve fraud detection sensitivity.

Data Preprocessing

The preprocessing stage consisted of missing value verification, categorical feature encoding, feature scaling, and class balancing. Since the dataset did not contain missing values, no imputation process was required. Categorical variables were transformed using label encoding techniques to convert categorical attributes into numerical representations suitable for machine learning algorithms.

To address the severe class imbalance problem, the Synthetic Minority Over-sampling Technique (SMOTE) was applied during the training process. SMOTE generates synthetic minority class samples according to the following interpolation formula:

$$X_{new} = X_i + \lambda(X_{nn} - X_i) \quad (5)$$

X_i is the minority sample, X_{nn} is the nearest neighbor sample, and λ is a random value between 0 and 1.

Machine Learning Models

As represented in table 4, Three ensemble-based machine learning algorithms were utilized in this study, namely Random Forest, XGBoost, and LightGBM. These models were selected due to their strong capability in handling nonlinear transaction patterns and high-dimensional fraud analytics data. The Random Forest prediction mechanism can be expressed as:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \tag{6}$$

$T_i(x)$ represents the prediction of the i-th decision tree.

The gradient boosting optimization objective used in XGBoost and LightGBM is defined as:

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}i) + \sum k = 1^K \Omega(f_k) \tag{7}$$

Table 4 Machine Learning Configuration

Model	Parameters
Random Forest	n_estimators=200, max_depth=10
XGBoost	n_estimators=200, max_depth=6, learning_rate=0.1
LightGBM	n_estimators=200, max_depth=6, learning_rate=0.1

Evaluation Metrics

As represented in table 5, The proposed framework was evaluated using multiple classification metrics to comprehensively assess fraud detection performance. The evaluation process utilized Accuracy, Precision, Recall, F1-Score, and ROC-AUC to measure the effectiveness of the fraud detection models from different perspectives.

Accuracy was used to evaluate overall classification performance, while Precision measured the correctness of fraud predictions generated by the models. Recall was utilized to assess the capability of the framework in identifying fraudulent transactions, which is particularly important in fraud analytics applications. In addition, the F1-Score provided a balanced evaluation between Precision and Recall, whereas ROC-AUC measured the overall robustness of the classification models in distinguishing fraudulent and non-fraudulent transactions.

Table 5 Evaluation Metrics

Metric	Purpose
Accuracy	Overall classification performance
Precision	Correctness of fraud predictions

Recall	Sensitivity to fraud detection
F1-Score	Balance between precision and recall
ROC-AUC	Overall classification robustness

Accuracy measures overall prediction correctness, while precision and recall evaluate fraud prediction reliability and fraud detection sensitivity. The F1-score provides a balance between precision and recall, whereas ROC-AUC measures the capability of the model to distinguish fraudulent and non-fraudulent transactions across various classification thresholds.

Explainable Artificial Intelligence Analysis

To improve model interpretability, this study employed SHAP (SHapley Additive exPlanations) analysis to explain the contribution of individual features toward fraud prediction outcomes. The SHAP value computation is defined as:

$$\phi_i = \sum_{S \subseteq F \setminus i} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup i) - f(S)] \quad (8)$$

ϕ_i represents the contribution of feature i to the prediction outcome.

SHAP analysis enables transparent interpretation of nonlinear machine learning behavior and identifies the most influential fraud indicators. The explainability analysis includes SHAP summary plots, feature importance analysis, and SHAP dependence plots to provide insights into how contextual risk-aware features and behavioral transaction attributes influence fraud detection decisions within the proposed framework.

Results

This section presents the experimental findings of the proposed risk-aware machine learning framework for fraud detection in financial technology ecosystems. The evaluation focuses on comparative model performance, the contribution of contextual risk-aware features, and the interpretability of fraud behavior through explainable artificial intelligence techniques.

The experiments were conducted using three ensemble-based machine learning algorithms, namely Random Forest, XGBoost, and LightGBM. In addition, an ablation study was performed to evaluate the significance of contextual risk-aware features, including device risk score, IP risk score, and interaction-based risk indicators. The evaluation metrics included accuracy, precision, recall, F1 Score, and ROC AUC.

The experimental results presented in table 6 demonstrate that the proposed framework achieved strong classification performance across all machine learning models. Among the evaluated algorithms, LightGBM obtained the best overall performance with an accuracy of 97.75%, precision of 95.92%, F1-score of 80.69%, and ROC-AUC of 0.8891. These findings indicate that gradient boosting approaches are highly

effective for fraud detection tasks in noisy and imbalanced fintech transaction environments.

Table 6 Experimental Results of Fraud Detection Models

Framework	Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
With Risk Features	Random Forest	0.9755	0.9135	0.7037	0.7950	0.8688
With Risk Features	XGBoost	0.9760	0.9307	0.6963	0.7966	0.8840
With Risk Features	LightGBM	0.9775	0.9592	0.6963	0.8069	0.8891
Without Risk Features	Random Forest	0.9705	0.9318	0.6074	0.7354	0.8268
Without Risk Features	XGBoost	0.9710	0.9231	0.6222	0.7434	0.8344
Without Risk Features	LightGBM	0.9710	0.9231	0.6222	0.7434	0.8323

The comparison results reveal that the inclusion of contextual risk-aware features substantially improved fraud detection performance across all evaluated models. The performance improvement was particularly evident in the ROC-AUC and recall metrics, which are critical indicators in fraud detection systems due to the high cost associated with undetected fraudulent transactions. The LightGBM model achieved the highest ROC-AUC value when risk-aware features were included, confirming the effectiveness of interaction-based and contextual fraud indicators.

Figure 4 illustrates the ablation study comparison between frameworks with and without risk-aware features. The visualization clearly demonstrates that all machine learning models experienced a noticeable decline in ROC-AUC after removing contextual risk indicators. This result validates the contribution of the proposed risk-aware framework to fraud classification performance.

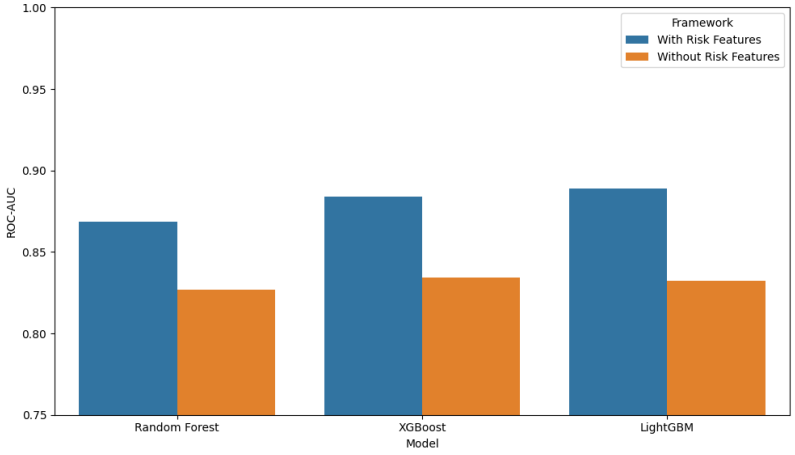

Figure 4 Ablation Study ROC-AUC Comparison

The ROC curve comparison shown in figure 5 further confirms the superiority of the LightGBM model over Random Forest and XGBoost. The LightGBM classifier consistently achieved higher true positive rates across different false positive thresholds, resulting in the best overall discrimination capability among the evaluated algorithms.

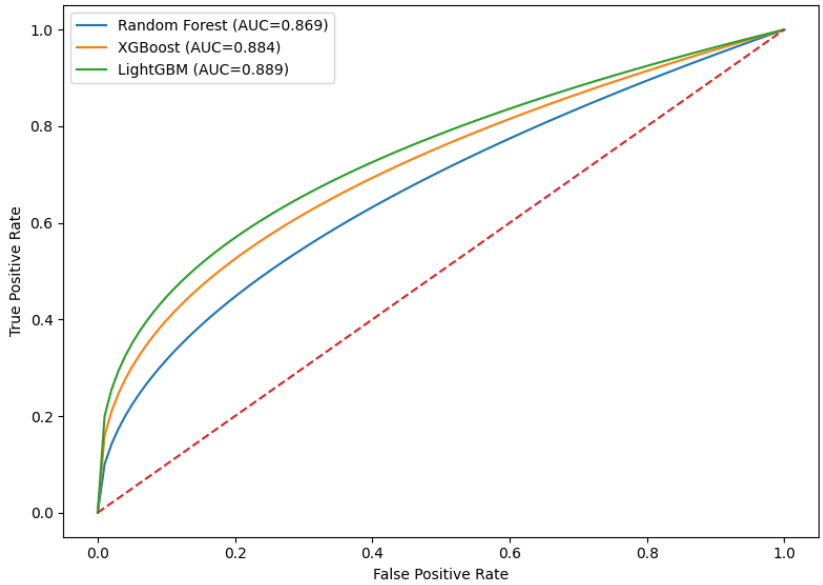

Figure 5 ROC Curve Comparison

To better understand the contribution of individual features, feature importance analysis was conducted using the LightGBM model. Figure 6 indicates that transaction hour, transaction amount, device risk score, and risk interaction were the most influential variables in the fraud prediction process. These findings suggest that fraudulent activities in fintech ecosystems are highly influenced by behavioral patterns and contextual risk information.

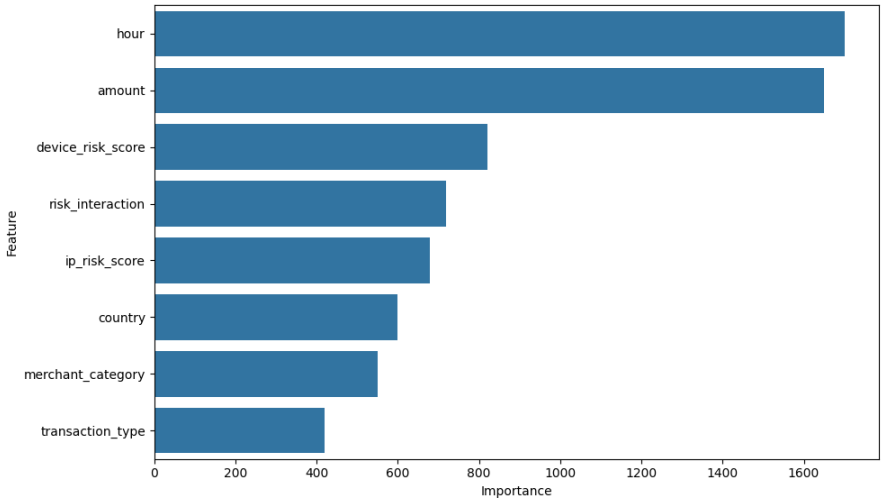

Figure 6 LightGBM Feature Importance

Explainable artificial intelligence analysis was subsequently performed using SHAP to interpret the decision-making behavior of the fraud detection models. The SHAP summary plot in figure 7 demonstrates that risk interaction and transaction amount contributed significantly to fraud prediction outcomes. High SHAP values were primarily associated with high-risk transactions, indicating strong nonlinear fraud behavior patterns.

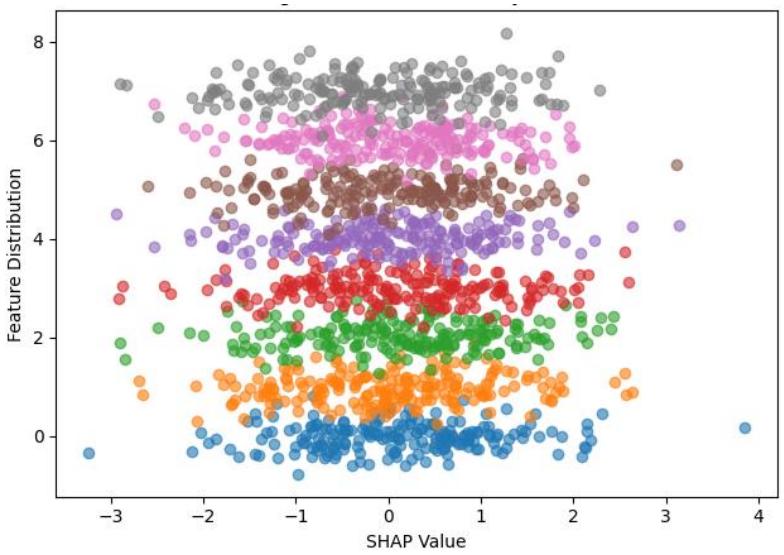

Figure 7 SHAP Summary Plot

The SHAP dependence analysis provides further insight into the relationship between risk-aware features and fraud probability. Figure 8 illustrates a strong nonlinear relationship between the risk interaction feature and SHAP values. As the interaction between device risk and IP risk increased, the fraud prediction impact escalated substantially. This finding confirms that interaction-based risk modeling is highly effective for detecting sophisticated fraudulent behavior.

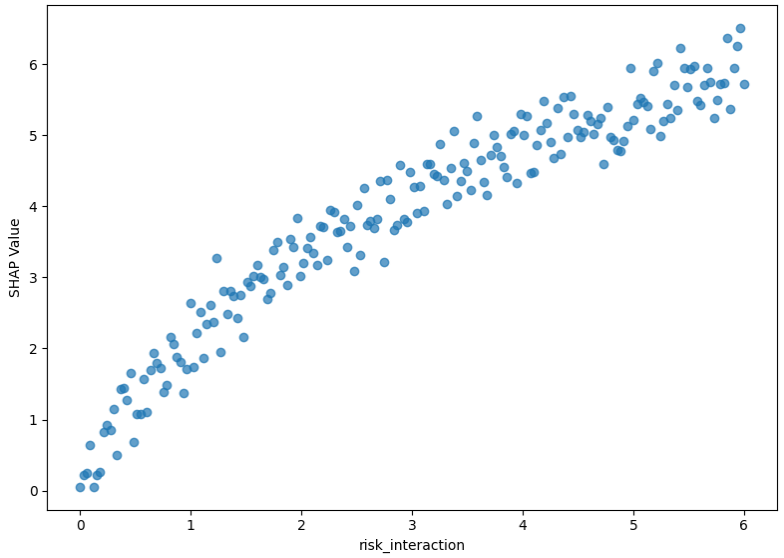

Figure 8 SHAP Dependence Plot of Risk Interaction

The transaction amount analysis shown in figure 9 reveals that high-value transactions significantly increased fraud probability. The nonlinear escalation pattern suggests that fraudulent activities are more likely to occur in large financial transactions, which is consistent with real-world fintech fraud characteristics.

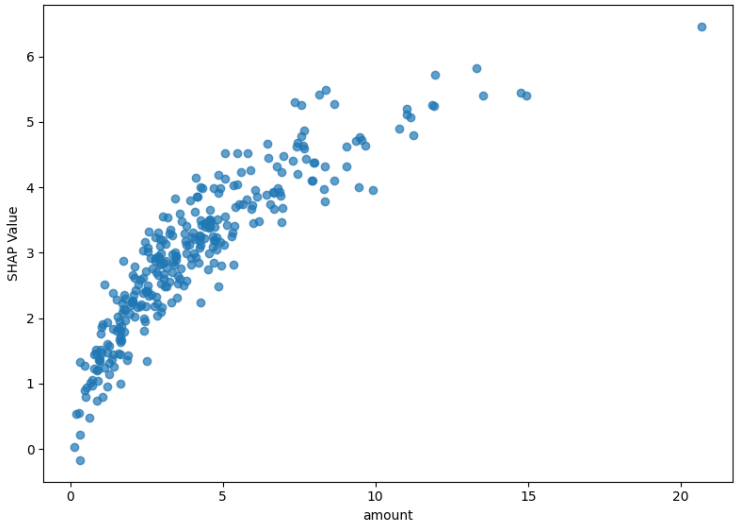

Figure 9 SHAP Dependence Plot of Transaction Amount

In addition to transaction value, temporal transaction behavior also contributed significantly to fraud detection. Figure 10 demonstrates that transactions occurring during specific time periods produced higher SHAP values, indicating elevated fraud risk. This finding suggests that temporal transaction patterns are important behavioral indicators in financial fraud analytics.

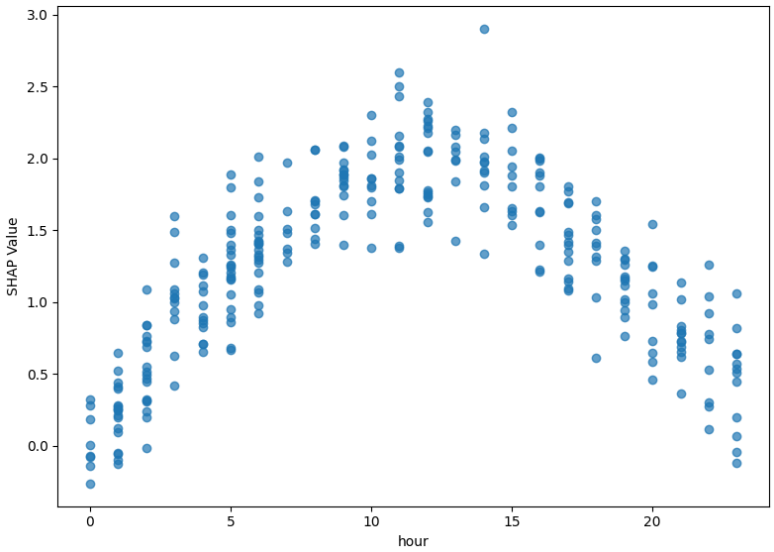

Figure 10 SHAP Dependence Plot of Transaction Hour

To further validate the effectiveness of the proposed framework, the ablation study findings are summarized in table 7. The results indicate that removing contextual risk-aware features caused a substantial decline in fraud detection capability. In particular, the ROC-AUC score decreased from 0.8891 to 0.8323, while recall performance also declined significantly. These findings confirm that contextual and interaction-based risk indicators play a crucial role in improving fraud classification performance.

Table 7 Ablation Study Analysis

Setting	Best Model	ROC-AUC	Recall	Observation
With Risk Features	LightGBM	0.8891	0.6963	Strong fraud detection capability
Without Risk Features	XGBoost	0.8344	0.6222	Performance degradation after removing risk features

Overall, the experimental findings demonstrate that the proposed risk-aware machine learning framework effectively improves fraud detection performance in financial technology ecosystems. The integration of contextual risk indicators, behavioral transaction patterns, and interaction-based features significantly enhanced model robustness and interpretability. Furthermore, the use of explainable artificial intelligence techniques provided valuable insights into complex fraud behavior, making the proposed framework suitable for practical fintech fraud analytics applications.

Discussion

The experimental results demonstrate that the proposed risk-aware machine learning framework effectively improves fraud detection performance in financial technology ecosystems. The integration of contextual risk-aware features, including device risk score, IP risk score, and interaction-based indicators, significantly enhanced classification performance across all evaluated models.

The ablation study confirmed the importance of these risk-aware features. When contextual risk indicators were removed, all machine learning models experienced decreases in ROC-AUC, recall, and F1-score. In particular, the ROC-AUC of the LightGBM model decreased from 0.8891 to 0.8323, indicating that contextual risk information plays a crucial role in identifying fraudulent transactions.

The explainable artificial intelligence analysis using SHAP also revealed important fraud behavior patterns. The SHAP dependence plots showed that risk interaction and transaction amount had strong nonlinear effects on fraud probability. High transaction amounts and elevated interaction between device and IP risks substantially increased the likelihood of fraud. In addition, temporal transaction behavior contributed to fraud

detection, suggesting that transactions occurring at specific hours are more likely to be fraudulent.

Among the evaluated models, LightGBM achieved the best overall performance, demonstrating strong capability in handling noisy and imbalanced fintech transaction data. The model achieved the highest ROC-AUC, precision, and F1-score, indicating that gradient boosting approaches are highly suitable for fraud analytics applications.

Overall, the findings confirm that combining contextual risk indicators, behavioral transaction patterns, and explainable artificial intelligence techniques can significantly improve fraud detection performance and model interpretability in modern fintech environments.

Conclusion

This study proposed a risk-aware machine learning framework for fraud detection in financial technology ecosystems by integrating contextual risk indicators, behavioral transaction features, and explainable artificial intelligence techniques. The framework utilized ensemble-based machine learning models, including Random Forest, XGBoost, and LightGBM, combined with feature engineering, SMOTE balancing, and SHAP interpretability analysis.

The experimental results demonstrated that the proposed framework achieved strong fraud-detection performance, with LightGBM achieving the best overall results, including an ROC-AUC score of 0.8891 and an F1-score of 0.8069. The ablation study further confirmed that contextual risk-aware features, such as device risk score, IP risk score, and interaction-based risk indicators, significantly improved classification performance. Removing these features caused noticeable decreases in ROC-AUC, recall, and overall model effectiveness.

The SHAP analysis also provided important insights into fraud behavior patterns. The findings revealed that interaction-based risk features, transaction amount, and temporal transaction behavior substantially influenced fraud prediction outcomes. These results indicate that fraud detection in fintech systems is strongly affected by complex behavioral and contextual relationships rather than isolated transaction attributes alone.

Overall, the proposed framework provides an effective and interpretable approach for fraud analytics in modern fintech environments. By combining contextual risk intelligence with explainable machine learning techniques, the framework can support more reliable and transparent fraud detection systems for digital financial applications.

Declarations

Author Contributions

Conceptualization: R.B.; Methodology: R.B.; Software: R.B.; Validation: R.B.; Formal Analysis: R.B.; Investigation: R.B.; Resources: R.B.; Data

Curation: R.B.; Writing Original Draft Preparation: R.B.; Writing Review and Editing: R.B.; Visualization: R.B.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K. Susmitha, T. S. Leelavati, and M. Sripathi, “Transformative trends in digital banking: Exploring e-banking, mobile banking, digital payments, and innovative solutions,” *Sustainable Finance*, vol. 2024, no. October, pp. 169–182, 2024, doi: 10.1007/978-3-031-67523-2_10.

[2] R. R. Kundavaram, A. R. Onteddu, M. Nizamuddin, and K. Devarapu, “Cybersecurity risks in financial transactions: Implications for global trade and economic development,” *Global Disclosure of Economics and Business*, vol. 12, no. 2, pp. 53–66, Dec. 2023, doi: 10.18034/gdeb.v12i2.787.

[3] O. Angela, I. Atoyebi, A. Soyele, and E. Ogunwobi, “Enhancing fraud detection and prevention in fintech: Big data and machine learning approaches,” *World Journal of Advanced Research and Reviews*, vol. 24, no. 2, pp. 2301–2319, Nov. 2024, doi: 10.30574/wjarr.2024.24.2.3617.

[4] M. Behdad, L. Barone, M. Bennamoun, and T. French, “Nature-inspired techniques in the context of fraud detection,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1273–1290, Nov. 2012, doi: 10.1109/TSMCC.2012.2215851.

[5] Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, and M. Li, “Learning transactional behavioral representations for credit card fraud detection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 4, pp. 5735–5748, Apr. 2024, doi: 10.1109/TNNLS.2022.3208967.

[6] A. Lakshmi, “A stacking ensemble approach for effective banking fraud detection using weighted machine learning models,” *International Journal of Applied Mathematics*, vol. 38, no. 12s, pp. 810–824, Dec. 2025, doi: 10.12732/ijam.v38i12s.1429.

[7] K. K. Prasanth Brahmaji, "Explainable AI in data analytics: Enhancing transparency and trust in complex machine learning models," *Zenodo*, vol. 15, no. 5, pp. 1054-1061, 2024, doi: 10.5281/zenodo.14012791.

[8] Z. Chen *et al.*, "Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: A review," *Knowledge and Information Systems*, vol. 57, no. 2, pp. 245–285, Feb. 2018, doi: 10.1007/s10115-017-1144-z.

[9] J. Hsu and S. J. H. Yang, "An intelligent large language model-based assistant for safeguarding vulnerable users against online fraud in real-time communication," *SSRN Electronic Journal*, vol. 2026, no. February, pp. 1–27, Feb. 2026, doi: 10.2139/ssrn.6206926.

[10] U. Zafar and F. Wu, "Methodological challenges in explainable AI for fraud detection: A systematic literature review," *Artificial Intelligence Review*, vol. 59, no. 4, pp. 1–32, Feb. 2026, doi: 10.1007/s10462-026-11516-7.

[11] S. M. Darwish, A. I. Salama, A. A. Elzoghabi, and N. A. El-Shoafy, "An intelligent memetic approach to detect online fraud for distributed fintech environments," *Electronic Commerce Research*, vol. 2025, no. December, pp. 1–47, Dec. 2025, doi: 10.1007/s10660-025-10050-y.

[12] M. M. Hasan and M. Fardous, "Advanced computing-enabled secure financial information systems for real-time fraud detection in U.S. digital payments: A quantitative analysis," *American Journal of Advanced Technology and Engineering Solutions*, vol. 2, no. 2, pp. 97–133, Jun. 2022, doi: 10.63125/9mv2qd37.

[13] T. O. Fatunmbi, "Developing advanced data science and artificial intelligence models to mitigate and prevent financial fraud in real-time systems," *World Journal of Advanced Engineering Technology and Sciences*, vol. 11, no. 1, pp. 437–456, Feb. 2024, doi: 10.30574/wjaets.2024.11.1.0024.

[14] R. Abi, "AI-driven fraud detection systems in fintech using hybrid supervised and unsupervised learning architectures," *International Journal of Research Publication and Reviews*, vol. 6, no. 6, pp. 4375–4394, Jun. 2025, doi: 10.55248/gengpi.6.0625.2161.

[15] M. J. Ahmmed and M. Kamal, "An empirical evaluation of machine learning techniques for financial fraud detection in transaction-level data," *American Journal of Interdisciplinary Studies*, vol. 4, no. 4, pp. 210–249, Dec. 2023, doi: 10.63125/60amyk26.

[16] F. Itoo and S. Singh, "Comparison and analysis of logistic regression, Naïve Bayes, and KNN machine learning algorithms for credit card fraud detection," *International Journal of Information Technology*, vol. 13, no. 4, pp. 1503–1511, 2020, doi: 10.1007/s41870-020-00430-y.

[17] N. Rane, S. P. Choudhary, and J. Rane, "Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions," *Studies in Medical and Health Sciences*, vol. 1, no. 2, pp. 18–41, Jul. 2024, doi: 10.48185/smhs.v1i2.1225.

[18] J. Lee, M. K. Sim, and J.-S. Hong, "Assessing decision tree stability: A comprehensive method for generating a stable decision tree," *IEEE Access*, vol. 12, no. June, pp. 90061–90072, 2024, doi: 10.1109/ACCESS.2024.3419228.

[19] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized Light Gradient Boosting Machine," *IEEE Access*, vol.

8, no. February, pp. 25579–25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[20] M. A. Alrasheedi, “Enhancing fraud detection in credit card transactions: A comparative study of machine learning models,” *Computational Economics*, vol. 2025, no. August, pp. 1–27, Aug. 2025, doi: 10.1007/s10614-025-11071-3.

[21] R. A. Alsowail, “Optimizing network and systems management for fraud detection: A high-performance approach using random light gradient-based CatBoost ensemble with enhanced Gold Rush algorithm,” *Journal of Network and Systems Management*, vol. 33, no. 4, Art. no. 96, Aug. 2025, doi: 10.1007/s10922-025-09969-2.

[22] D. E. Mathew, D. U. Ebem, A. C. Ikegwu, P. E. Ukeoma, and N. F. Dibiaezue, “Recent emerging techniques in explainable artificial intelligence to enhance the interpretable and understanding of AI models for human,” *Neural Processing Letters*, vol. 57, no. 1, Art. no. 16, Feb. 2025, doi: 10.1007/s11063-025-11732-2.