Machine Learning-Based Risk Classification of Blockchain Cryptocurrency Transactions Using Ensemble Algorithms

R. Karthick Manoj[1,*], Gokulnath Anandakumar[2]

[1,2]Department of Electrical and Electronics Engineering, AMET University, Chennai, India

ABSTRACT

Blockchain cryptocurrency transactions are increasingly exposed to financial risks, fraudulent activities, and suspicious transaction behaviors due to the decentralized nature of blockchain networks. This study proposes a machine learning-based framework for blockchain cryptocurrency transaction risk classification using ensemble learning algorithms. The dataset consists of 9,000 blockchain transactions with multiple behavioral and transaction-related attributes, including transaction amount, transaction fee, wallet age, transaction velocity, darknet association score, and unusual transaction patterns. To address the class imbalance problem, the SMOTEENN resampling technique was applied, transforming the original class distribution of 6,086 normal transactions and 1,114 risky transactions into a more balanced distribution of 5,866 risky transactions and 2,588 normal transactions. Several ensemble learning algorithms, including Random Forest, Gradient Boosting, XGBoost, and LightGBM, were evaluated using confusion matrix analysis and ROC-AUC metrics. Experimental results demonstrate that XGBoost achieved the highest risky transaction detection capability with 169 True Positive predictions and 110 False Negatives, while LightGBM generated the highest True Negative value of 1,422 and the lowest False Positive value of 99. Random Forest achieved the highest ROC-AUC score of 0.553531, followed by LightGBM with 0.542258. Feature importance analysis revealed that output_count and input_count were the most influential attributes with importance scores of 1289 and 1224, respectively. Although the obtained ROC-AUC values indicate that blockchain transaction risk classification remains a challenging problem, the findings demonstrate that ensemble learning algorithms can support blockchain risk analytics and cryptocurrency fraud detection systems.

Keywords Blockchain, Cryptocurrency Risk Classification, Ensemble Learning, Fraud Detection, Machine Learning

INTRODUCTION

The rapid development of blockchain technology and cryptocurrency systems has significantly transformed modern digital financial transactions by enabling decentralized, transparent, and borderless financial activities [1]. Cryptocurrencies such as Bitcoin and Ethereum have gained widespread global adoption due to their efficiency, security, and anonymity features [2]. However, the increasing growth of cryptocurrency transactions has also introduced various cybersecurity threats, including fraud, money laundering, phishing attacks, darknet-related activities, and illegal financial transactions [3].

Unlike conventional banking systems, blockchain networks operate without centralized supervision, making transaction monitoring and risk analysis considerably more challenging [4]. The pseudonymous nature of blockchain wallet addresses further complicates the identification of suspicious transaction behavior and malicious financial activities [5]. As a result, blockchain transaction risk analysis and cryptocurrency fraud detection have become important

Submitted 28 May 2026
Accepted 5 August 2026
Published 17 November 2026

Corresponding author
R. Karthick Manoj,
karthickmanoj.r@gmail.com

Additional Information and Declarations can be found on page 276

DOI: 10.47738/jcrb.v3i4.84

Copyright
2026 Manoj and Anandakumar

Distributed under
Creative Commons CC-BY 4.0

How to cite this article: R. K. Manoj and G. Anandakumar, "Machine Learning-Based Risk Classification of Blockchain Cryptocurrency Transactions Using Ensemble Algorithms," *J. Curr. Res. Blockchain*, vol. 3, no. 4, pp. 263-278, 2026.

research topics in cybersecurity and financial technology research [6].

Machine learning techniques have been widely utilized to improve blockchain transaction monitoring systems because of their ability to identify hidden transaction patterns within large-scale financial datasets [7]. Previous studies have applied various machine learning algorithms such as Logistic Regression, Decision Tree, Support Vector Machine, Random Forest, and Neural Networks for cryptocurrency fraud detection and blockchain risk prediction [8]. Recently, ensemble learning algorithms have attracted significant attention due to their capability to improve classification stability and predictive performance by combining multiple machine learning models [9].

Several studies reported that Random Forest and XGBoost achieved promising results in fraud detection and anomaly detection tasks involving financial transaction data [10]. Other studies also demonstrated that Gradient Boosting and LightGBM were effective for handling high-dimensional transaction datasets and complex classification problems [11].

Despite the growing number of studies related to blockchain fraud detection, several limitations remain unresolved. Many previous studies focused primarily on traditional machine learning approaches without extensively evaluating multiple ensembles learning algorithms on imbalanced blockchain transaction datasets. In addition, several existing studies relied heavily on accuracy metrics, which may produce misleading performance evaluations when applied to fraud detection problems with severe class imbalance conditions. Furthermore, previous research often neglected detailed feature importance analysis and transaction behavioral interpretation in blockchain risk prediction systems.

Another important challenge is the overlapping behavioral characteristics between normal and risky blockchain transactions, which reduce the discriminative capability of machine learning algorithms. These limitations indicate the need for more comprehensive research integrating ensemble learning algorithms, imbalance handling techniques, and feature importance analysis for blockchain cryptocurrency transaction risk classification.

Based on these challenges, this study proposes a machine learning-based framework for blockchain cryptocurrency transaction risk classification using ensemble learning algorithms. The proposed framework evaluates Random Forest, Gradient Boosting, XGBoost, and LightGBM combined with the SMOTEENN resampling technique to address the class imbalance problem. The study utilizes a dataset consisting of 9,000 blockchain cryptocurrency transactions with multiple behavioral and transaction-related attributes.

The experimental results demonstrate that XGBoost achieved the highest risky transaction detection capability with 169 True Positive predictions, while LightGBM generated the highest True Negative value of 1422 and the lowest False Positive value of 99. In addition, Random Forest achieved the highest ROC-AUC score of 0.553531, indicating relatively stable classification performance. Feature importance analysis also revealed that transaction structure and behavioral attributes such as output_count, input_count, and darknet_association_score significantly influenced blockchain transaction risk prediction.

The contributions of this study are summarized as follows. First, this study evaluates the performance of multiple ensemble learning algorithms for blockchain cryptocurrency transaction risk classification. Second, the study addresses the class imbalance problem using the SMOTEENN resampling technique. Third, feature importance analysis is conducted to identify the most influential transaction attributes affecting blockchain risk prediction. Finally, this

study provides insights into the challenges and trade-offs of applying ensemble learning algorithms in blockchain cybersecurity and cryptocurrency fraud detection systems.

Literature Review

Blockchain technology has become one of the most important innovations in digital financial systems because it enables decentralized and transparent financial transactions without centralized intermediaries [12]. Cryptocurrencies such as Bitcoin and Ethereum have experienced rapid adoption in recent years, leading to increased transaction volumes and more complex blockchain ecosystems [13]. However, the rapid growth of cryptocurrency transactions has also increased the risk of fraudulent activities, money laundering, phishing attacks, and darknet-related financial crimes [14].

Machine learning techniques have been widely applied in blockchain risk analysis and cryptocurrency fraud detection because of their capability to identify hidden behavioral patterns in transaction data [15]. Previous studies utilized algorithms such as Logistic Regression, Decision Tree, Support Vector Machine, and Neural Networks for fraud detection and anomaly classification tasks [16]. However, traditional machine learning models often experience limitations when handling large-scale and highly imbalanced blockchain transaction datasets.

Ensemble learning algorithms have recently gained significant attention due to their ability to improve classification performance and predictive robustness [17]. Random Forest and XGBoost have demonstrated promising results in fraud detection and cybersecurity classification tasks because of their capability to handle nonlinear transaction behavior and high-dimensional features [18]. In addition, Gradient Boosting and LightGBM have also shown effective performance in financial transaction analysis and anomaly detection systems [19].

Class imbalance remains one of the major challenges in blockchain transaction risk classification because risky transactions typically represent only a small portion of the dataset [20]. Imbalanced data conditions may cause machine learning models to become biased toward majority class predictions, reducing their capability to identify suspicious blockchain activities. To address this issue, several resampling methods such as SMOTE and SMOTEENN have been proposed to improve minority class representation during model training [21].

Feature importance analysis has also become an important aspect of blockchain transaction analytics. Previous studies revealed that transaction amount, wallet activity, transaction velocity, and darknet-related indicators significantly influence cryptocurrency fraud detection performance [22]. These behavioral transaction characteristics are considered important predictors for identifying suspicious blockchain activities.

Despite the growing number of studies related to blockchain fraud detection, several research gaps remain unresolved. Many previous studies focused primarily on single machine learning models without comprehensively evaluating multiple ensemble learning algorithms on imbalanced blockchain cryptocurrency transaction datasets. Furthermore, several studies relied mainly on accuracy metrics without analyzing the trade-off between risky transaction detection capability and classification stability.

Based on these limitations, this study evaluates the performance of multiple ensemble learning algorithms, including Random Forest, Gradient Boosting, XGBoost, and LightGBM, for blockchain cryptocurrency transaction risk classification. In addition, the SMOTEENN resampling technique is applied to address the class imbalance problem, while feature importance analysis is conducted to identify the most influential transaction attributes affecting blockchain transaction risk prediction.

Research Methodology

This study proposes a machine learning-based framework for blockchain cryptocurrency transaction risk classification using ensemble learning algorithms. The proposed methodology consists of dataset collection, data preprocessing, imbalance handling, ensemble model training, and performance evaluation. Figure 1 illustrates the overall research framework used in this study.

Research Framework

The research framework begins with blockchain cryptocurrency transaction dataset collection followed by preprocessing procedures, including identifier removal, feature scaling, and imbalance handling using the SMOTEENN technique. After preprocessing, several ensemble learning algorithms were trained and evaluated using confusion matrix analysis and ROC-AUC metrics.

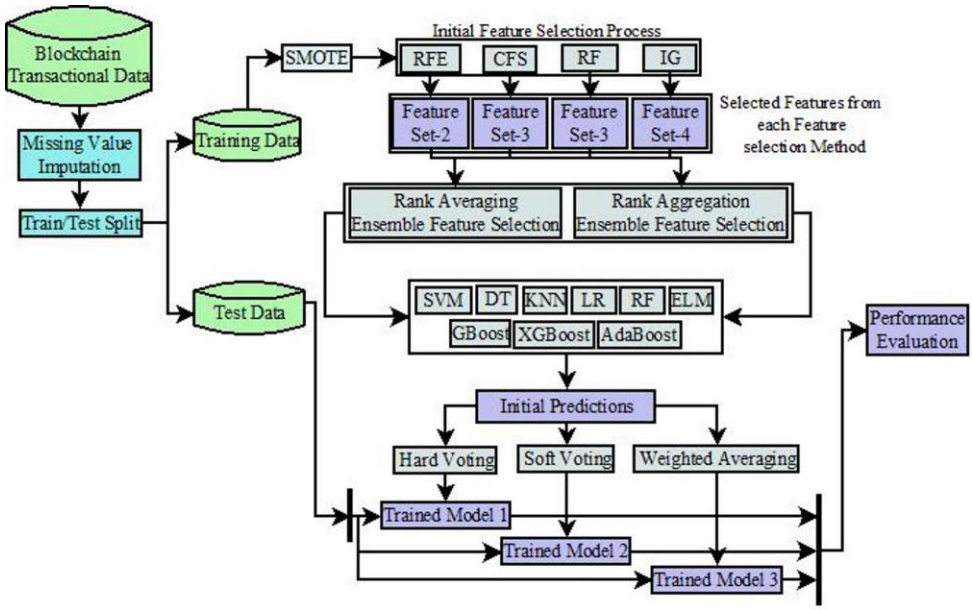


Figure 1 Overall Research Framework

The proposed framework was designed to improve risky transaction detection capability while maintaining classification stability for normal blockchain transactions.

Dataset Description

The dataset used in this study consists of 9,000 blockchain cryptocurrency transactions containing multiple behavioral and transaction-related attributes. The dataset includes features associated with transaction structure, wallet activity, transaction frequency, and suspicious transaction indicators.

Table 1 presents the attributes used in the blockchain cryptocurrency transaction dataset. The dataset contains transaction-related and behavioral features, including transaction amount, wallet activity, darknet association indicators, and suspicious transaction behavior scores.

Table 1 Dataset Description

Attribute	Description
transaction_amount	Cryptocurrency transaction value
transaction_fee	Blockchain transaction fee
block_confirmation_time	Time required for block confirmation
input_count	Number of transaction inputs
output_count	Number of transaction outputs
wallet_age_days	Wallet age in days
total_wallet_transactions	Total wallet transaction history
avg_wallet_transaction_value	Average wallet transaction value
cross_border_flag	Cross-border transaction indicator
mixing_service_used	Indicator of cryptocurrency mixing service usage
darknet_association_score	Darknet activity association score
transaction_velocity_24h	Transaction frequency over 24 hours
unusual_pattern_score	Suspicious transaction behavior score
risk_label	Transaction risk classification label

Dataset Statistics

The dataset statistics were analyzed to understand the characteristics of the blockchain transaction data before model training.

Table 2 shows that the dataset exhibits a class imbalance condition where normal transactions dominate risky transactions. Such imbalance conditions may reduce the capability of machine learning algorithms to effectively identify risky blockchain activities.

Table 2 Dataset Statistics

Metric	Value
Total Transactions	9000
Number of Features	13
Target Variable	risk_label
Normal Transactions	7607
Risky Transactions	1393

Data Preprocessing

As illustrated in figure 2, data preprocessing was performed to improve dataset quality before the machine learning training process. The preprocessing stage included removing identifier attributes, train-test splitting, feature scaling, and class balancing.

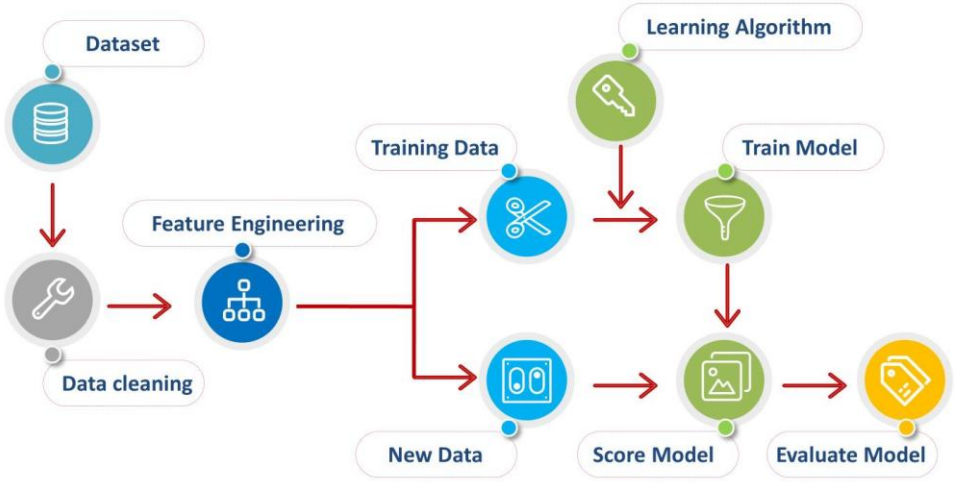


Figure 2 Data Preprocessing Workflow

The preprocessing workflow begins by removing transaction_id and wallet_id because these attributes do not contribute to transaction risk prediction. The dataset was then divided into training and testing sets using an 80:20 ratio. Numerical features were standardized using StandardScaler to improve model learning stability.

StandardScaler transforms numerical features into standardized distributions using the following equation:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

x represents the original feature value, μ represents the feature mean, and σ represents the standard deviation. This process ensures that all features contribute proportionally during model training.

As presented in table 3, to address the class imbalance problem, the SMOTEENN resampling technique was applied during the training process. SMOTEENN combines Synthetic Minority Oversampling Technique (SMOTE) and Edited Nearest Neighbor (ENN) to improve minority class representation while reducing noisy samples.

Table 3 Data Preprocessing Steps

Step	Description
Remove ID Features	Remove transaction_id and wallet_id
Train-Test Split	Split dataset into training and testing sets
Feature Scaling	Standardize numerical features using StandardScaler
Class Balancing	Apply SMOTEENN resampling
Model Training	Train ensemble learning algorithms

Ensemble Learning Algorithms

This study utilized four ensemble learning algorithms, namely Random Forest, Gradient Boosting, XGBoost, and LightGBM. These algorithms were selected because of their capability to improve classification performance and predictive

robustness for complex financial transaction datasets. Random Forest utilizes multiple decision trees to improve classification stability and reduce overfitting. The prediction process of Random Forest can be represented as:

$$\hat{f}(x) = \frac{1}{B}\sum_{b=1}^{B} f_b\,(x) \tag{2}$$

B represents the total number of decision trees and $f_b(x)$ represents the prediction generated by each tree.

Gradient Boosting iteratively improves prediction performance by minimizing classification errors from weak learners. The boosting process can be represented as:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{3}$$

As illustrated in figure 3, $F_m(x)$ represents the updated prediction model, $h_m(x)$ represents the weak learner, and γ_m represents the learning rate contribution. XGBoost extends gradient boosting optimization with regularization to improve prediction stability and reduce overfitting. Meanwhile, LightGBM utilizes histogram-based learning and leaf-wise tree growth to improve computational efficiency for large-scale datasets.

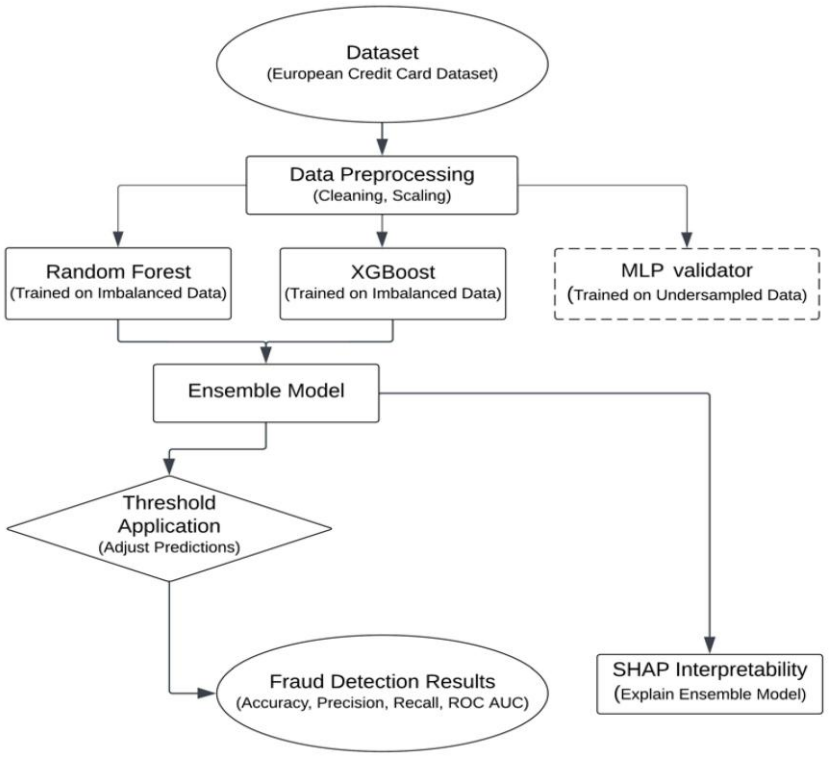


Figure 3 Proposed Machine Learning Architecture

As summarized in table 3, the proposed architecture integrates preprocessing, resampling, ensemble learning algorithms, and performance evaluation into a unified blockchain transaction risk classification framework..

Table 4 Hyperparameter Configuration

Model	Hyperparameters
Random Forest	n_estimators=300, class_weight=balanced, random_state=42
Gradient Boosting	n_estimators=300, learning_rate=0.05, random_state=42
XGBoost	n_estimators=300, max_depth=6, learning_rate=0.05, scale_pos_weight=5
LightGBM	n_estimators=300, learning_rate=0.05, class_weight=balanced

Performance Evaluation

The performance of the ensemble learning algorithms was evaluated using confusion matrix analysis and ROC-AUC metrics. Confusion matrix analysis was utilized to measure True Positive, True Negative, False Positive, and False Negative predictions generated by each model.

Accuracy was used to evaluate the overall classification performance of the model in distinguishing normal and risky blockchain transactions. Precision was used to measure the proportion of correctly predicted risky transactions among all transactions predicted as risky. Recall was utilized to evaluate the capability of the model to correctly identify risky blockchain transactions, while the F1-Score was used to measure the balance between Precision and Recall performance.

In addition, ROC-AUC was employed to evaluate the discrimination capability of the ensemble learning algorithms in distinguishing risky and normal blockchain transactions. A higher ROC-AUC value indicates better classification capability and stronger model discrimination performance.

Feature importance analysis was also conducted to identify the most influential transaction attributes affecting blockchain cryptocurrency transaction risk prediction. This analysis was used to interpret the contribution of transaction behavioral features toward the classification process.

Results

This section presents the experimental results obtained from implementing ensemble machine learning algorithms for the classification of blockchain cryptocurrency transaction risk. The experiments utilized Random Forest, Gradient Boosting, XGBoost, and LightGBM models. To address class imbalance, the SMOTEENN resampling technique was applied during preprocessing. The evaluation process focused on confusion matrix analysis, ROC-AUC performance, and feature importance interpretation.

Class Distribution Analysis

As shown in figure 4, the original dataset exhibited a significant imbalance between normal and risky transactions. Before resampling, the dataset contained 6086 normal transactions and only 1114 risky transactions. Such imbalanced conditions may cause machine learning models to become biased toward majority class predictions and reduce their ability to identify risky blockchain activities.

After applying the SMOTEENN technique, the class distribution became more balanced, with 5866 risky transactions and 2588 normal transactions. The balancing process improved the capability of the ensemble learning algorithms to learn minority class patterns associated with risky cryptocurrency transactions..

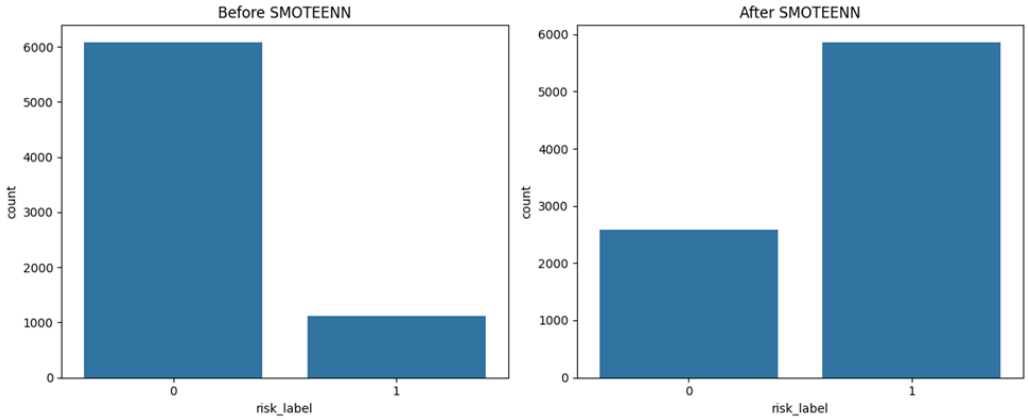
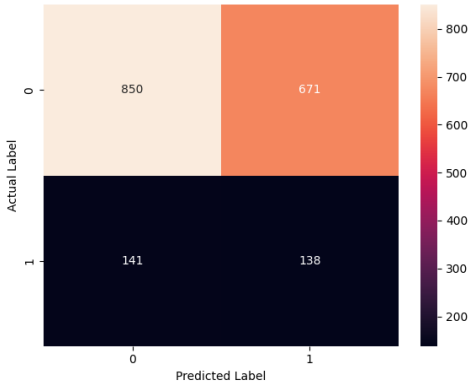


Figure 4 Class Distribution Before and After SMOTEENN

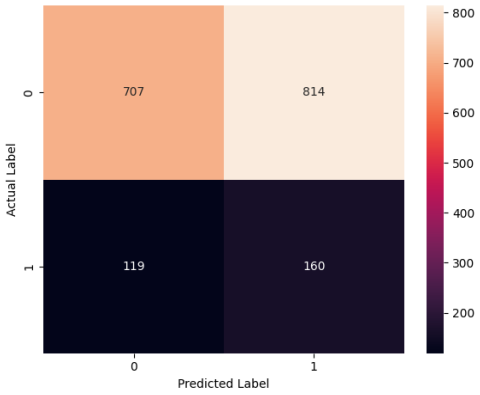
The experimental results demonstrate that SMOTEENN effectively reduced class imbalance and improved the overall training environment for the machine learning models.

Confusion Matrix Comparison

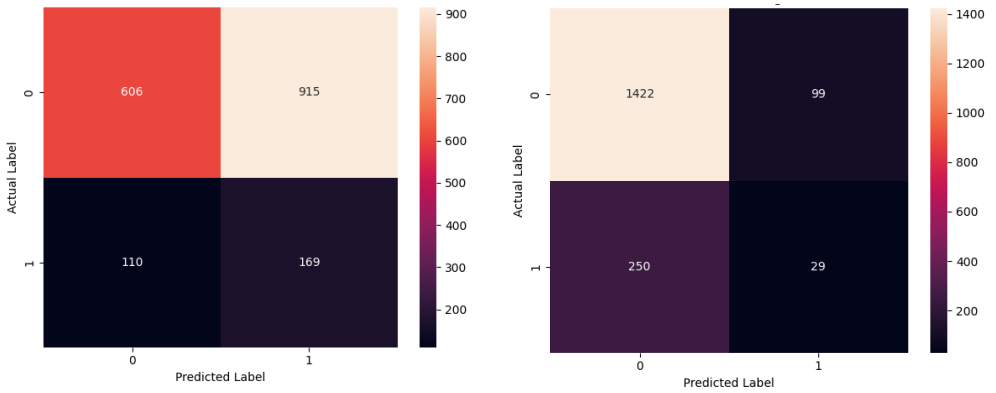
As illustrated in figure 5, the confusion matrix analysis was conducted to evaluate the classification capability of each ensemble learning algorithm in distinguishing risky and normal blockchain transactions.



Confusion Matrix-Random Forest



Confusion Matrix-Gradient Boosting



Confusion Matrix-XGBoost Confusion Matrix-LightGBM

Figure 5 Confusion Matrix Comparison

Random Forest generated 138 True Positive predictions and 850 True Negative predictions, demonstrating relatively balanced classification performance. However, the model produced 671 False Positive predictions and 141 False Negative predictions, indicating that some normal transactions were incorrectly classified as risky. The confusion matrix also shows that Random Forest maintained moderate detection capability while preserving acceptable classification stability.

Gradient Boosting improved risky transaction detection performance by generating 160 True Positive predictions and reducing False Negative predictions to 119. Nevertheless, the number of False Positive predictions increased to 814, indicating that the model became more aggressive in identifying risky transactions. This behavior suggests that Gradient Boosting prioritized minority class detection at the expense of normal transaction classification accuracy.

XGBoost achieved the highest True Positive value of 169 and the lowest False Negative value of 110 among all evaluated models. These results demonstrate that XGBoost provided the strongest capability in detecting risky blockchain transactions. However, the confusion matrix reveals that the model also produced the highest False Positive value of 915, which reduced classification specificity and increased the misclassification of normal transactions.

LightGBM demonstrated superior capability in recognizing normal blockchain transactions by generating 1422 True Negative predictions and only 99 False Positive predictions. Despite this strong normal transaction classification performance, the model produced only 29 True Positive predictions and 250 False Negative predictions, indicating limited capability in detecting risky transactions.

Overall, the confusion matrix comparison reveals that each ensemble learning algorithm exhibits different trade-offs between risky transaction detection capability and classification stability. Models such as XGBoost and Gradient Boosting achieved stronger risky transaction detection performance, while LightGBM demonstrated higher specificity for normal transaction classification.

ROC Curve Evaluation

As illustrated in figure 6, ROC curve analysis was performed to evaluate the

discrimination capability of the ensemble learning algorithms..

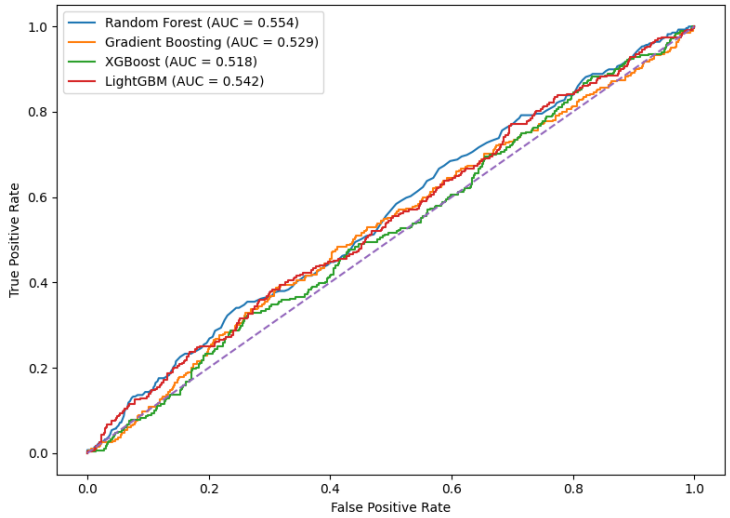


Figure 6 ROC Curve Comparison

The obtained ROC-AUC values ranged between 0.518 and 0.554. Random Forest achieved the highest ROC-AUC score of 0.553531, followed by LightGBM with 0.542258, Gradient Boosting with 0.529057, and XGBoost with 0.518189.

Although the ROC-AUC scores remain relatively modest, the results indicate that ensemble learning algorithms still provide measurable predictive capability for blockchain transaction risk classification. The relatively low ROC-AUC values suggest that the dataset contains overlapping characteristics between normal and risky transactions, making the classification process considerably challenging. The ROC curve analysis also confirms that Random Forest provided the most stable overall performance among the evaluated models.

Feature Importance Analysis

As illustrated in figure 7, feature importance analysis was conducted using the LightGBM model to identify the most influential transaction attributes affecting risk classification performance.

.

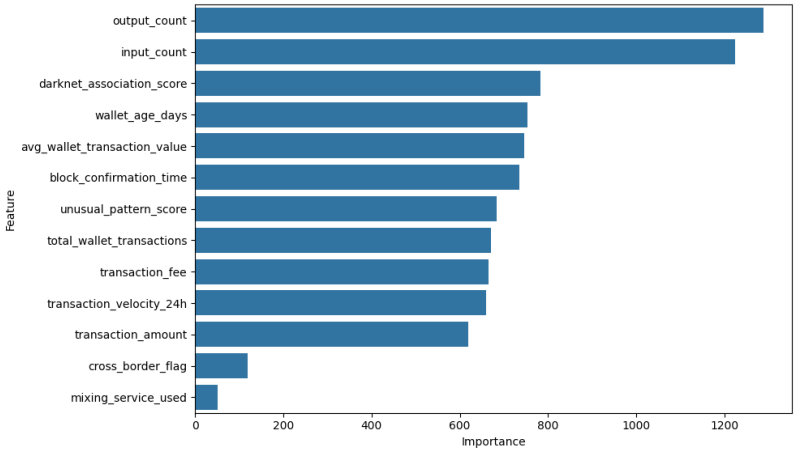


Figure 7 Feature Importance Analysis

The results indicate that output_count achieved the highest importance score of 1289, followed by input_count with 1224. These findings demonstrate that transaction structure and transaction flow characteristics strongly influence blockchain transaction risk prediction.

The darknet_association_score feature obtained an importance score of 783, indicating that darknet-related activity is an important indicator of risky cryptocurrency transactions. Other influential features included wallet_age_days with a score of 754, avg_wallet_transaction_value with 746, and block_confirmation_time with 736.

Meanwhile, cross_border_flag and mixing_service_used produced relatively low importance scores of 118 and 51, respectively. This finding suggests that transaction behavioral patterns contribute more significantly to risk classification than binary indicator features.

Model Performance Comparison

Table 5 summarizes the overall performance comparison of the evaluated ensemble learning algorithms.

Table 5 Model Performance Comparison

Model	TP	TN	FP	FN	ROC-AUC
Random Forest	138	850	671	141	0.553531
Gradient Boosting	160	707	814	119	0.529057
XGBoost	169	606	915	110	0.518189
LightGBM	29	1422	99	250	0.542258

The results demonstrate that XGBoost achieved the highest risky transaction detection capability due to its highest True Positive value. However, the model also generated the largest number of False Positives. In contrast, LightGBM demonstrated the strongest capability in identifying normal transactions because it achieved the highest True Negative value and the lowest False Positive rate.

Random Forest obtained the highest ROC-AUC score, indicating relatively better classification stability compared to the other evaluated models.

Confusion Matrix Evaluation

Table 6 presents a detailed confusion matrix analysis for each ensemble learning algorithm.

Table 6 Confusion Matrix Analysis

Model	TP	TN	FP	FN	ROC-AUC
Random Forest	138	850	671	141	0.553531
Gradient Boosting	160	707	814	119	0.529057
XGBoost	169	606	915	110	0.518189
LightGBM	29	1422	99	250	0.542258

The confusion matrix evaluation confirms that aggressive models such as

XGBoost and Gradient Boosting improved risky transaction detection but significantly increased False Positive predictions. Conversely, LightGBM maintained strong normal transaction classification capability but showed limited effectiveness in detecting risky blockchain activities. These findings indicate that blockchain risk classification remains a complex machine learning problem due to overlapping transaction patterns and class imbalance characteristics.

Top Important Features

Table 7 presents the ranking of the most influential features extracted from the LightGBM feature importance analysis.

Table 7 Top Important Features

Feature	Importance
output_count	1289
input_count	1224
darknet_association_score	783
wallet_age_days	754
avg_wallet_transaction_value	746
block_confirmation_time	736
unusual_pattern_score	684
total_wallet_transactions	670
transaction_fee	666
transaction_velocity_24h	659
transaction_amount	620
cross_border_flag	118
mixing_service_used	51

The results confirm that transaction structure, transaction flow complexity, wallet behavioral characteristics, and darknet-related activity significantly influence blockchain cryptocurrency transaction risk classification. These findings demonstrate the importance of behavioral analytics in blockchain cybersecurity and fraud detection systems.

Discussion

The experimental results demonstrate that ensemble learning algorithms provide different levels of effectiveness in blockchain cryptocurrency transaction risk classification. The implementation of SMOTEENN successfully reduced the class imbalance problem and improved the capability of the models to detect risky transactions.

Among the evaluated models, XGBoost achieved the highest True Positive value, indicating stronger capability in detecting risky blockchain transactions. However, the model also generated the highest number of False Positives, which reduced classification specificity. Gradient Boosting showed similar behavior by improving risky transaction detection while increasing False Positive predictions.

Random Forest achieved the highest ROC-AUC value, indicating relatively stable overall classification performance compared to the other evaluated models. Meanwhile, LightGBM demonstrated superior capability in identifying

normal transactions due to its high True Negative value and low False Positive rate, although its capability in detecting risky transactions remained limited.

The feature importance analysis revealed that transaction structure-related features such as output_count, input_count, and darknet_association_score significantly influenced blockchain transaction risk classification. These findings indicate that transaction behavior patterns play an important role in identifying suspicious blockchain activities. Despite the improvements achieved through ensemble learning and SMOTEENN, the relatively low ROC-AUC values indicate that blockchain transaction risk classification remains a challenging problem due to overlapping characteristics between normal and risky transactions.

Conclusion

This study proposed a machine learning-based framework for blockchain cryptocurrency transaction risk classification using ensemble learning algorithms. The experiments were conducted using Random Forest, Gradient Boosting, XGBoost, and LightGBM combined with the SMOTEENN resampling technique.

The experimental results demonstrated that XGBoost achieved the strongest risky transaction detection capability, while LightGBM showed better performance in recognizing normal transactions. Random Forest achieved the highest ROC-AUC value, indicating relatively stable classification performance.

Feature importance analysis showed that transaction structure and behavioral attributes significantly influenced blockchain transaction risk prediction. Although the proposed framework improved minority class detection, the relatively low ROC-AUC values indicate that blockchain transaction risk classification remains a complex task.

Overall, the findings demonstrate that ensemble learning algorithms can support blockchain risk analytics and cryptocurrency fraud detection systems. Future studies may explore hybrid deep learning models and explainable artificial intelligence techniques to further improve classification performance.

Declarations

Author Contributions

Conceptualization: R.K.M., G.A.; Methodology: G.A.; Software: R.K.M.; Validation: R.K.M., G.A.; Formal Analysis: R.K.M., G.A.; Investigation: R.K.M.; Resources: G.A.; Data Curation: G.A.; Writing Original Draft Preparation: R.K.M., G.A.; Writing Review and Editing: G.A., R.K.M.; Visualization: R.K.M.; 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] Y. Chen, "Decentralized finance: Blockchain technology and the quest for an open financial system," *SSRN Electronic Journal*, vol. 2019, no. July, pp. 1–27, Jul. 2019, doi: 10.2139/ssrn.3418557.

[2] A. F. M. Shahen Shah *et al.*, "On the vital aspects and characteristics of cryptocurrency—A survey," *IEEE Access*, vol. 11, no. January, pp. 9451–9468, 2023, doi: 10.1109/ACCESS.2023.3240103.

[3] A. Adel and M. Norouzifard, "Weaponization of the growing cybercrimes inside the dark net: The question of detection and application," *Big Data and Cognitive Computing*, vol. 8, no. 8, Art. no. 91, pp. 1-37, Aug. 2024, doi: 10.3390/bdcc8080091.

[4] M. Zachariadis, G. Hileman, and S. V. Scott, "Governance and control in distributed ledgers: Understanding the challenges facing blockchain technology in financial services," *Information and Organization*, vol. 29, no. 2, pp. 105–117, Jun. 2019, doi: 10.1016/j.infoandorg.2019.03.001.

[5] U. Agarwal, V. Rishiwal, S. Tanwar, and M. Yadav, "Blockchain and crypto forensics: Investigating crypto frauds," *International Journal of Network Management*, vol. 34, no. 2, p. e2255, Dec. 2023, doi: 10.1002/nem.2255.

[6] S. Saha *et al.*, "Cryptocurrency and financial crimes: A bibliometric analysis and future research agenda," *Multidisciplinary Reviews*, vol. 7, no. 8, Art. no. 2024168, May 2024, doi: 10.31893/multirev.2024168.

[7] S. Kayikci and T. M. Khoshgoftaar, "Blockchain meets machine learning: A survey," *Journal of Big Data*, vol. 11, no. 1, p. 9, Jan. 2024, doi: 10.1186/s40537-023-00852-y.

[8] E. Akyildirim, A. Goncu, and A. Sensoy, "Prediction of cryptocurrency returns using machine learning," *Annals of Operations Research*, vol. 297, nos. 1–2, pp. 3–36, Apr. 2020, doi: 10.1007/s10479-020-03575-y.

[9] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, no. September, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.

[10] R. M. Zakaria *et al.*, "Detecting financial fraud in real-time transactions using graph neural networks and anomaly detection techniques," *Journal of Economics, Finance and Accounting Studies*, vol. 7, no. 6, pp. 01–13, Sep. 2025, doi: 10.32996/jefas.2025.7.6.1.

[11] X. Ma *et al.*, "Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGBoost algorithms according to different high dimensional data cleaning," *Electronic Commerce Research and Applications*, vol. 31, no. September, pp. 24–39, Sep. 2018, doi: 10.1016/j.elerap.2018.08.002.

[12] Y. Chen and C. Bellavitis, “Blockchain disruption and decentralized finance: The rise of decentralized business models,” *Journal of Business Venturing Insights*, vol. 13, no. June, p. e00151, Jun. 2020, doi: 10.1016/j.jbvi.2019.e00151.

[13] M. H. u. Rehman, K. Salah, E. Damiani, and D. Svetinovic, “Trust in blockchain cryptocurrency ecosystem,” *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1196–1212, Nov. 2020, doi: 10.1109/TEM.2019.2948861.

[14] R. El-Kady, “Exploring the advances of artificial intelligence in cryptocurrencies and darknet forensics,” in *Advances in Computational Intelligence and Robotics*, vol. 2026, no. July, pp. 1–35, Jul. 2026, doi: 10.4018/406039.

[15] T. Ashfaq *et al*., “A machine learning and blockchain based efficient fraud detection mechanism,” *Sensors*, vol. 22, no. 19, Art. no. 7162, pp. 1-20, Sep. 2022, doi: 10.3390/s22197162.

[16] E. A. Ubom and A. E. Okoro, “Anomaly detection using artificial intelligence methods: Support Vector Machine and decision tree algorithms,” *Journal of Electronic Voltage and Application*, vol. 6, no. 2, pp. 1–12, Dec. 2025, doi: 10.30880/jeva.2025.06.02.006.

[17] Y. Zhang, J. Liu, and W. Shen, “A review of ensemble learning algorithms used in remote sensing applications,” *Applied Sciences*, vol. 12, no. 17, Art. no. 8654, pp. 1-20, Aug. 2022, doi: 10.3390/app12178654.

[18] H. T. Sukmana, “Detecting threatening content in social media: A data mining approach using random forest for classification of tweets in Cyberlaw context,” *Journal of Cyber Law*, vol. 1, no. 2, pp. 131-146, Jun. 2025, doi: 10.63913/jcl.v1i2.14

[19] M. K. Islam, P. Hridi, M. S. Hossain, and H. S. Narman, “Network anomaly detection using LightGBM: A gradient boosting classifier,” in *Proceedings of the 2020 30th International Telecommunication Networks and Applications Conference (ITNAC)*, Melbourne, VIC, Australia, 2020, vol. 2021, no. January, pp. 1–7, doi: 10.1109/ITNAC50341.2020.9315049.

[20] A. Abbas Jasim Al-Hchaimi, M. A. Khalifa, and W. El-Shafai, “Explainable AI with imbalanced learning strategies for blockchain transaction fraud detection,” *Engineering Reports*, vol. 8, no. 1, pp. 1-25, Jan. 2026, doi: 10.1002/eng2.70545.

[21] G. Husain *et al*., “SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models,” *Algorithms*, vol. 18, no. 1, Art. no. 37, pp. 1-16, Jan. 2025, doi: 10.3390/a18010037.

[22] M. Bühlmann, H.-G. Fill, and S. Curty, “Blockchain data analytics: A scoping literature review and directions for future research,” *arXiv preprint* arXiv:2505.04403, vol. 2025, no. May, pp. 1-25, 2025, doi: 10.48550/arXiv.2505.04403.