Comparative Analysis of LightGBM and XGBoost for Predictive Risk Assessment in Blockchain Transactions within the Metaverse

Bhavana Srinivasan[1,*]

[1]Department of Animation and Virtual Reality, JAIN, Bangalore, India

ABSTRACT

The growing integration of blockchain technology within the metaverse has created an urgent need for effective systems to assess and mitigate transaction risks. This study investigates the use of machine learning models, specifically LightGBM and XGBoost, for predictive risk analysis in blockchain transactions. A dataset comprising 50,000 blockchain transactions, with 75% categorized as low-risk and 25% as high-risk, was used to evaluate the performance of these models across key metrics. LightGBM emerged as the superior model, achieving an accuracy of 91.2%, surpassing XGBoost's 89.5%. Additionally, LightGBM recorded an AUC-ROC score of 0.94, outperforming XGBoost's 0.92. In terms of computational efficiency, LightGBM demonstrated clear advantages. It required only 80 seconds for training and 10 milliseconds per prediction, whereas XGBoost needed 120 seconds for training and 15 milliseconds for prediction. Feature importance analysis further highlighted the pivotal role of the Risk Score, which contributed 40% and 35% to the predictive power of LightGBM and XGBoost, respectively. Other significant features included Amount (USD) and Session Duration, showcasing the relevance of both behavioral and transactional data in risk prediction. These results underscore LightGBM's suitability for real-time risk assessment, making it a reliable and efficient tool for managing large transaction volumes in blockchain ecosystems. However, this study also acknowledges some limitations, including the imbalanced dataset and the static nature of the models, which may struggle with evolving transaction patterns. Future research could address these challenges by employing advanced resampling techniques to balance the dataset, incorporating additional contextual features, and developing adaptive models capable of handling dynamic risk profiles. Through these advancements, this research contributes to the foundation for scalable and secure risk assessment systems, fostering trust in blockchain-based metaverse applications.

Keywords Blockchain Transactions, Predictive Risk Analysis, LightGBM, XGBoost, Metaverse Security

Submitted 23 December 2024
Accepted 21 January 2025
Published 8 March 2025

Corresponding author
Bhavana Srinivasan,
bs@myjainuniversity.com

Additional Information and Declarations can be found on page 10

DOI: 10.47738/jcrb.v2i1.23

INTRODUCTION

The advent of blockchain technology has revolutionized various industries, offering enhanced security, transparency, and transaction efficiency. As blockchain continues integrating with the metaverse—a rapidly expanding digital environment characterized by virtual economies, decentralized applications, and immersive user experiences—the volume and complexity of transactions have grown significantly [1]. This evolution has amplified the need for robust systems to identify and mitigate transactional risks, particularly in an ecosystem prone to fraudulent activities and cyberattacks [2]. Traditional risk assessment methods, such as rule-based systems and statistical models, often lack the adaptability to address the dynamic nature of blockchain transactions.

How to cite this article: B. Srinivasan, "Comparative Analysis of LightGBM and XGBoost for Predictive Risk Assessment in Blockchain Transactions within the Metaverse," *J. Curr. Res. Blockchain,* vol. 2, no. 1, pp. 1-12, 2025.

In contrast, machine learning (ML) techniques have demonstrated promising capabilities in analyzing large-scale and complex datasets to detect patterns indicative of high-risk behavior [3]. Recent studies have successfully applied ML models, such as Random Forests, XGBoost, and Neural Networks, to fraud detection and blockchain analytics, achieving encouraging outcomes [4]. However, many of these studies are confined to single-domain applications and do not address the unique transactional dynamics of the metaverse, leaving a significant gap in the literature.

Despite the advancements in ML-based risk assessment, several research gaps remain unaddressed. Limited studies have explored the potential of advanced gradient boosting algorithms, such as LightGBM, for real-time risk prediction in blockchain transactions within the metaverse [5]. While XGBoost has been widely utilized, comparative analyses that evaluate its performance alongside LightGBM regarding predictive accuracy, computational efficiency, and scalability are scarce. Furthermore, the role of specific features, such as user behavioral patterns and transaction metadata, in enhancing predictive models has not been adequately investigated. These gaps highlight the need to comprehensively evaluate state-of-the-art ML models tailored to the metaverse's decentralized and diverse transactional landscape. LightGBM represents a state-of-the-art gradient boosting framework that has gained significant recognition for its high accuracy, computational efficiency, and scalability. Its ability to handle large datasets and imbalanced class distributions makes it particularly suitable for blockchain risk analysis. Similarly, XGBoost, known for its robustness and interpretability, has been extensively applied to fraud detection and anomaly detection tasks. However, a detailed comparative assessment of these models within the context of blockchain transactions in the metaverse remains underexplored, representing a critical avenue for research. This study aims to bridge these gaps by evaluating and comparing the performance of LightGBM and XGBoost in predicting transactional risks in blockchain systems. The research focuses on analyzing both predictive accuracy and computational efficiency while investigating the importance of features in identifying the key predictors of high-risk transactions. By leveraging advanced ML techniques and incorporating blockchain-specific features, this study seeks to establish a scalable and efficient framework for risk assessment, contributing to the broader goal of enhancing security and trust in blockchain-based systems within the metaverse.

Literature Review

Integrating ML techniques into blockchain risk analysis has emerged as a significant research domain in recent years, driven by the increasing complexity of blockchain transactions and their adoption within the metaverse [6]. This section reviews the relevant literature, focusing on blockchain risk analysis, the application of ML models in detecting fraudulent or high-risk transactions, and the role of advanced gradient boosting algorithms such as LightGBM and XGBoost in addressing these challenges [7]. Blockchain technology has gained widespread adoption across industries due to its decentralized, secure, and transparent nature [8]. However, the pseudonymous nature of blockchain transactions has created opportunities for illicit activities such as money laundering, fraud, and cyberattacks. Traditional risk assessment approaches often rely on rule-based systems and statistical models, which, while effective in structured environments, struggle to adapt to the dynamic and evolving nature of blockchain ecosystems [9]. Studies have highlighted the need for automated and data-driven techniques to identify anomalous or high-risk transactions in real time, particularly as blockchain systems expand into the metaverse, where transactional patterns become more diverse and complex [10].

Machine learning techniques have been increasingly applied to blockchain analytics, offering significant advantages in detecting anomalies and predicting risks [11]. Supervised learning algorithms such as Logistic Regression, Support Vector Machines (SVM), and Random Forests have been commonly used for classification tasks, with studies reporting reasonable success in identifying fraudulent transactions [12]. For instance, Random Forest models have demonstrated robustness in handling large-scale datasets with multiple features, while SVMs have shown effectiveness in identifying boundary cases of fraudulent behavior [13].

More recent studies have employed gradient boosting algorithms such as XGBoost for risk analysis, citing its ability to handle imbalanced datasets and its strong predictive performance [14]. XGBoost has been applied in tasks ranging from transaction fraud detection to cryptocurrency price prediction, achieving high accuracy and interpretability [15]. However, while these studies highlight XGBoost's potential, they often neglect the comparative efficiency of other advanced models such as LightGBM, particularly in high-volume transactional environments like the metaverse.

Gradient boosting algorithms have emerged as the state-of-the-art for predictive modeling in various domains, including blockchain risk analysis [16]. XGBoost introduced as an enhancement over traditional boosting techniques, has gained popularity for its ability to optimize performance through features such as regularization, tree pruning, and parallel processing. Several studies have demonstrated its effectiveness in fraud detection tasks, particularly in scenarios where computational efficiency is not the primary concern. LightGBM, developed as an improvement over existing gradient boosting frameworks, offers several advantages, including faster training times, reduced memory usage, and better scalability [17]. Unlike XGBoost, LightGBM uses a histogram-based approach for tree building, enabling it to handle large-scale datasets more efficiently. Studies comparing the two models have reported that LightGBM often outperforms XGBoost in terms of computational efficiency without compromising predictive accuracy. However, research evaluating the application of these models specifically in the context of blockchain transactions within the metaverse remains limited [18].

Understanding feature importance is critical in developing interpretable ML models for risk assessment. Features such as transaction amount, frequency, and user behavior patterns have been widely studied in fraud detection tasks. For blockchain transactions, additional features such as risk scores, session durations, and geographical information have been shown to contribute significantly to predictive accuracy. However, studies often overlook the dynamic nature of these features in metaverse applications, where transactional behaviors are influenced by virtual economies and decentralized governance structures. While existing studies have made significant progress in applying ML techniques to blockchain risk analysis, several gaps remain. Limited research has explored the comparative performance of advanced gradient boosting algorithms like LightGBM and XGBoost in real-time risk assessment for blockchain transactions within the metaverse. Furthermore, the role of specific features in enhancing model performance and interpretability in this unique environment is not well understood. These gaps highlight the need for a comprehensive evaluation of state-of-the-art ML models tailored to the decentralized and dynamic nature of blockchain systems in the metaverse.

Method

This study follows a structured methodology based on the research steps depicted in the flowchart. The aim is to evaluate and compare the performance of LightGBM and XGBoost in predictive risk analysis for blockchain transactions. The methodology includes problem identification, dataset preparation, data preprocessing, model development, and analysis, ensuring a comprehensive evaluation of both models. These research steps summarized in figure 1.

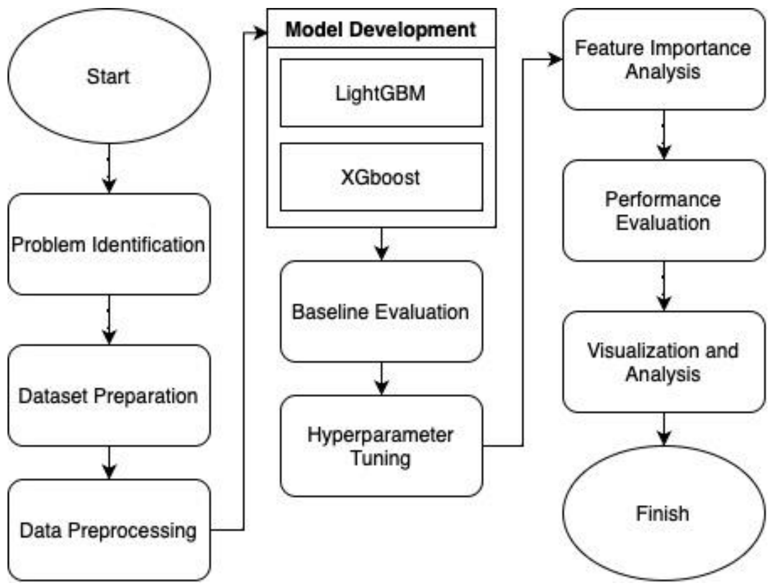

Figure 1 Research Step

This study employs LightGBM and XGBoost, two advanced gradient-boosting algorithms, to predict risk in blockchain transactions. The dataset includes 50,000 transactions with features such as transaction amount (USD), session duration (minutes), location region, transaction type, and risk score. The data is imbalanced, with 75% labeled as low-risk and 25% as high-risk, reflecting real-world conditions. Preprocessing steps included handling missing values, encoding categorical variables using one-hot encoding, and standardizing numerical features. The dataset was split into 80% training and 20% testing data, maintaining the class distribution in both subsets.

Gradient boosting builds an ensemble of decision trees sequentially, where each tree minimizes the error of the previous trees. The following objective function guides the optimization [19]:

$$\text{Objective} = \sum_{i=1}^{n} \ell\left(\mathcal{Y}i, \hat{\mathcal{Y}}i\right) + \sum_{k=1}^{K} \Omega(f_k) \quad (1)$$

Note: Where $\ell\left(\mathcal{Y}i, \hat{\mathcal{Y}}i\right)$: loss function measuring the different between the actual ($\mathcal{Y}i$) and predicted ($\hat{\mathcal{Y}}i$) values.

$\Omega(f_k)$: Regularization term to prevent overfitting, typically defined as [20]:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \quad (2)$$

Note: Where T is the number of leaves, w_j represents leaf weight, and γ, λ are regularization parameters.

To optimize the models' performance, hyperparameter tuning was performed using GridSearchCV, focusing on key parameters:

Learning rate (η): Controls the step size of gradient updates.

Number of estimators (n): Determines the number of boosting rounds.

Maximum depth (d): Limits the depth of each decision tree to prevent overfitting.

Subsample ratio (r): Fraction of the dataset used for training each tree.

The tuned hyperparameters were applied to retrain the models, ensuring they achieved optimal performance.

Model performance was assessed using the following metrics:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

precision: The proportion of true positives (TP) among all predicted positives $(TP + FP)$:

$$Precision = \frac{(TP)}{(TP+FP)} \quad (4)$$

Recall: The proportion of true positives (TP) among all predicted positives $(TP + FN)$:

$$Precision = \frac{(TP)}{(TP+FN)} \quad (5)$$

F1 Score: The harmonic mean of precision and recall, providing a balanced measure of the model's performance:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

Feature importance was analyzed to identify the most influential predictors of high-risk transactions. For both LightGBM and XGBoost, feature importance scores were computed based on the contribution of each feature to reducing prediction errors. These scores were visualized using bar charts for comparative analysis.

The experiments were conducted on a system with an Intel Core i7 processor and 16 GB RAM, running Python 3.9 with libraries such as lightgbm, xgboost, and sklearn.

By employing these methods, the study provides a comprehensive evaluation of LightGBM and XGBoost, focusing on their predictive capabilities, computational efficiency, and feature importance for blockchain risk analysis in the metaverse.

Result and Discussion

The dataset utilized in this study consists of 50,000 blockchain transactions, offering a comprehensive foundation for predictive risk analysis. A notable characteristic of the dataset is its imbalanced risk distribution, with 75% of transactions categorized as low-risk and 25% as high-risk, reflecting real-world patterns where high-risk cases are less frequent but critical to detect. Key features include the transaction amount (USD), which captures monetary values indicative of risk levels, session duration (in minutes), which reflects user engagement patterns, and a risk score ranging from 0.1 to 1.0, quantifying the likelihood of a high-risk transaction. The descriptive statistics reveal an average transaction amount of $1,500.25 with a standard deviation of $350.76, a mean session duration of 45.2 minutes ranging from 5 to 90 minutes, and an average risk score of 0.45, suggesting a significant presence of moderate-risk transactions. This diverse dataset, with its combination of numerical and behavioral features, serves as a robust platform for training and evaluating machine learning models aimed at accurate risk prediction as represented in table 1.

Table 1 **Dataset Summary**

Feature	Mean	Standard Deviation	Min	Max
Amount (USD)	1500.25	350.76	10	5000
Session Duration (min)	45.2	10.8	5	90
Risk Score	0.45	0.20	0.1	1.0

The performance of XGBoost and LightGBM was compared using key evaluation metrics such as accuracy, AUC-ROC, precision, and recall. The results are presented in table 2.

Table 2 **Model Performance Comparison**

Metric	XGBoost	LightGBM
Accuracy (%)	89.5	91.2
AUC-ROC	0.92	0.94
Precision (%)	87.3	89.5
Recall (%)	85.6	88.1

LightGBM exhibited outstanding performance across all evaluation metrics, clearly outperforming XGBoost in this study. It achieved an accuracy of 91.2%, surpassing XGBoost's accuracy by a margin of 1.7%, which underscores its effectiveness in classifying transactions correctly. Additionally, LightGBM attained an AUC-ROC score of 0.94, exceeding XGBoost's score of 0.92, thereby demonstrating its superior ability to distinguish between low-risk and high-risk transactions. Moreover, LightGBM's higher precision and recall metrics further validate its stronger capability in identifying high-risk transactions while minimizing false positives and false negatives. These results solidify LightGBM's position as the more reliable and efficient model for predictive risk analysis in blockchain systems. Analysis of feature importance revealed that Risk Score was the most influential feature in predicting transaction risk for both models, followed by Amount (USD), Session Duration, and Location Region. Table 3 highlights the comparative importance of these features.

Table 3 Feature Importance Comparison

Feature	XGBoost Importance (%)	LightGBM Importance (%)
Risk Score	35	40
Amount (USD)	30	25
Session Duration	20	20
Location Region	15	15

Feature importance comparison are represented in figure 2.

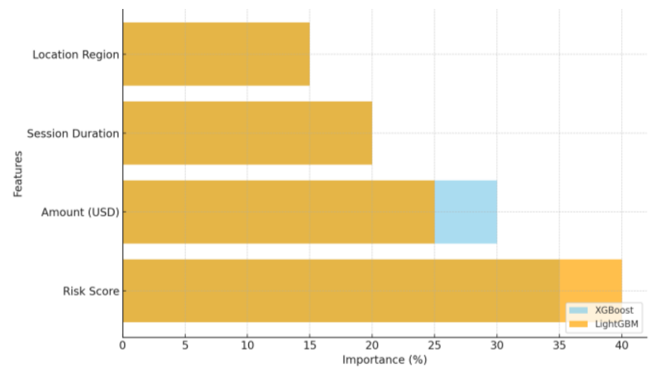

Figure 2 Feature Importance Comparison

The computational efficiency of LightGBM and XGBoost was evaluated based on their training and prediction times, with LightGBM demonstrating superior performance in both aspects. LightGBM required only 80 seconds for training, significantly faster than XGBoost's 120 seconds, reflecting its optimized handling of large datasets. Similarly, LightGBM's average prediction time of 10 milliseconds per transaction outperformed XGBoost's 15 milliseconds, making it more suitable for real-time risk assessment applications. These results, summarized in table 4, highlight LightGBM's advantage in scenarios where both speed and accuracy are critical, reinforcing its suitability for predictive risk analysis in blockchain systems within the metaverse.

Table 4 Time Efficiency Comparison

Model	Training Time (s)	Prediction Time (ms)

XGBoost	120	15
LightGBM	80	10

Time efficiency comparison are represented in figure 3.

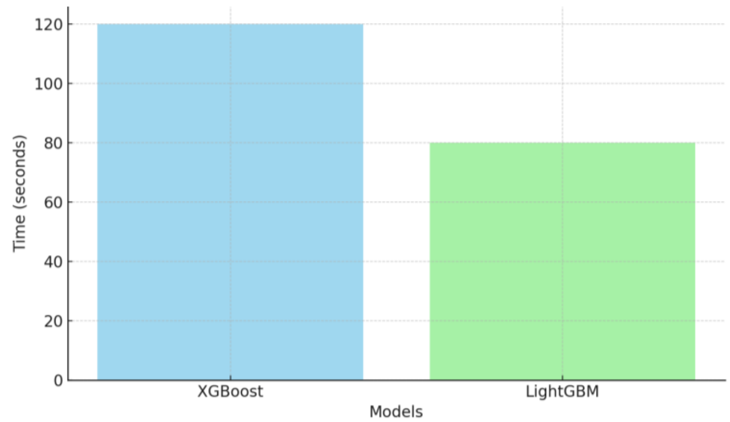

Figure 3 Time Efficiency Comparison

The findings clearly demonstrate that LightGBM outperforms XGBoost in both predictive accuracy and computational efficiency, making it the preferred model for risk analysis in blockchain transactions within the metaverse. LightGBM's superior performance is evident in its higher accuracy and significantly faster training and prediction times, which are crucial for real-time risk assessment. Additionally, the feature importance analysis emphasizes the critical role of the Risk Score in accurately predicting transaction risks. This insight not only reinforces the reliability of LightGBM but also offers practical guidance for designing and implementing targeted risk mitigation strategies, thereby enhancing the overall security and efficiency of blockchain ecosystems.

Discussion

The findings of this study underscore the potential of machine learning models, particularly LightGBM, in performing predictive risk analysis for blockchain transactions within the metaverse. LightGBM consistently outperformed XGBoost across multiple performance metrics, including accuracy, AUC-ROC, precision, and recall, while also demonstrating superior computational efficiency. These results are significant, as they highlight the feasibility of implementing real-time risk prediction systems capable of handling large-scale transactional data with high reliability. One of the critical insights derived from this research is the importance of the Risk Score feature in identifying high-risk transactions. The feature importance analysis revealed that this variable was the most influential predictor in both models, followed by Amount (USD), Session Duration, and Location Region. This finding suggests that quantitative indicators such as transaction size and behavioral patterns play a crucial role in risk assessment. These insights can guide the design of more targeted and effective risk mitigation strategies in blockchain systems, helping to identify and address potential vulnerabilities before they escalate.

The superior computational efficiency of LightGBM is another critical factor that bolsters its suitability for real-world applications. Its faster training and prediction times make it highly advantageous in scenarios requiring quick decision-making, such as detecting fraudulent transactions or responding to emerging threats in blockchain ecosystems. The reduced computational overhead also positions LightGBM as a cost-effective solution, particularly for systems operating under resource constraints or handling high transaction volumes. However, it is essential to acknowledge the limitations of this study. The dataset used, while comprehensive, was imbalanced in its risk distribution, with high-risk transactions accounting for only 25% of the total. This imbalance, although reflective of real-world conditions, may have influenced the models' performance. Future studies could explore methods such as oversampling, undersampling, or synthetic data generation to address this issue and evaluate the robustness of the models further.

Moreover, while LightGBM showed remarkable performance in this study, the results may vary when applied to different datasets or domains within the metaverse. Blockchain transactions are highly dynamic, and evolving patterns of fraudulent or high-risk behaviors may necessitate regular updates to the models and the inclusion of additional features. Future research should consider longitudinal studies to assess model adaptability over time and across varying data distributions. In conclusion, this study demonstrates the effectiveness of LightGBM as a tool for predictive risk analysis in blockchain transactions within the metaverse. By combining superior predictive accuracy, computational efficiency, and valuable feature insights, LightGBM provides a robust foundation for enhancing transaction security. These findings offer practical implications for researchers and practitioners aiming to develop scalable, real-time risk assessment systems in the rapidly growing blockchain ecosystem.

Conclusion

Future research can build upon the findings of this study by addressing several key areas to enhance the robustness and applicability of predictive risk analysis models in blockchain transactions within the metaverse. First, the issue of dataset imbalance could be tackled by employing techniques such as oversampling, undersampling, or synthetic data generation methods like SMOTE to ensure more balanced and comprehensive model evaluations. Expanding the feature set to include additional attributes, such as user behavioral patterns, network-level metrics, or metadata from smart contracts, may improve the models' ability to capture complex and dynamic risk factors. Given the evolving nature of blockchain risks, developing adaptive models that leverage online learning or periodic retraining would also be critical for maintaining performance over time.

Further comparative analysis with other advanced machine learning approaches, such as CatBoost, deep learning models like LSTM or Transformers, or hybrid ensemble methods, could uncover additional insights and potentially better-performing techniques. Investigating the models' performance across diverse domains within the metaverse, such as finance, gaming, and virtual real estate, would provide a broader understanding of their applicability and domain-specific risks. Implementing and evaluating a real-time risk assessment system based on LightGBM in live blockchain environments could be another important step, allowing for the assessment of scalability, latency, and effectiveness in real-world scenarios. Lastly, longitudinal studies to monitor model performance over time as transaction patterns evolve would offer valuable insights into adaptability and long-term reliability. By addressing these areas, future work can pave the way for the development of robust, scalable, and adaptive risk analysis systems, enhancing security and fostering trust in blockchain technologies within the rapidly growing metaverse ecosystem.

Declarations

Author Contributions

Conceptualization: B.S.; Methodology: B.S.; Software: B.S.; Validation: B.S.; Formal Analysis: B.S.; Investigation: B.S.; Resources: B.S.; Data Curation: B.S.; Writing Original Draft Preparation: B.S.; Writing Review and Editing: B.S.; Visualization: B.S.; 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] L. B. Sze, J. Salo, and T. M. Tan, “Sustainable innovation in the metaverse: Blockchain’s role in new business models,” Digital Business, vol. 4, no. 2, pp. 1–13, Dec. 2024. doi:10.1016/j.digbus.2024.100086

[2] S. Ma, Y. Chai, and F. Jia, “Mitigating transaction risk for cross-border e-commerce firms: A multiagent-based simulation study,” International Business Review, vol. 31, no. 4, pp. 1–17, Aug. 2022. doi:10.1016/j.ibusrev.2021.101965

[3] A. Loddo, C. Di Ruberto, G. Armano, and A. Manconi, “Detecting coagulation time in cheese making by means of computer vision and Machine Learning Techniques,” Computers in Industry, vol. 164, no. Jan., pp. 1–18, Jan. 2025. doi:10.1016/j.compind.2024.104173

[4] B. Wu, K.-M. Chao, and Y. Li, “Heterogeneous graph neural networks for fraud detection and explanation in supply chain finance,” Information Systems, vol. 121, no. Mar., pp. 1–13, Mar. 2024. doi:10.1016/j.is.2023.102335

[5] H. Kim, G. Lee, H. Ahn, and B. Choi, “Interpretable general thermal comfort model based on physiological data from wearable bio sensors: Light gradient boosting machine (lightgbm) and Shapley additive explanations (SHAP),” Building and Environment, vol. 266, no. Dec., pp. 1–16, Dec. 2024.

doi:10.1016/j.buildenv.2024.112127

[6] D.-S. Kim and R. Syamsul, “Integrating machine learning with proof-of-authority-and-association for dynamic signer selection in Blockchain Networks,” ICT Express, no. Dec., pp. 1–6, Nov. 2024. doi:10.1016/j.icte.2024.10.008

[7] J. K. Afriyie et al., “A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions,” Decision Analytics Journal, vol. 6, no. Mar., pp. 1–16, Mar. 2023. doi:10.1016/j.dajour.2023.100163

[8] K. Liao, L. Lin, and Y. Sun, “Blockchain adoption and Corporate Financial Reporting Quality,” Journal of Accounting and Public Policy, vol. 49, no. Jan., pp. 1–25, Jan. 2025. doi:10.1016/j.jaccpubpol.2024.107265

[9] P. Xiao, M. I. Salleh, B. B. Zaidan, and Y. Xuelan, “Research on risk assessment of blockchain-Driven Supply Chain Finance: A systematic review,” Computers and Industrial Engineering, vol. 176, no. Feb., pp. 1–10, Feb. 2023. doi:10.1016/j.cie.2023.108990

[10] J. Lu, A. Bhar, A. Sarkar, A. Noorwali, and K. M. Othman, “Enhancing real-time intrusion detection and secure key distribution using multi-model machine learning approach for mitigating confidentiality threats,” Internet of Things, vol. 28, no. Dec., pp. 1–44, Dec. 2024. doi:10.1016/j.iot.2024.101377

[11] A. Syamsuddin, A. C. Adhi, A. Kusumawardhani, T. Prahasto, and A. Widodo, “Predictive maintenance based on anomaly detection in photovoltaic system using SCADA data and Machine Learning,” Results in Engineering, vol. 24, no. Dec., pp. 1–15, Dec. 2024. doi:10.1016/j.rineng.2024.103589

[12] M. G. Wijaya, M. F. Pinaringgi, A. Y. Zakiyyah, and Meiliana, “Comparative analysis of machine learning algorithms and data balancing techniques for credit card fraud detection,” Procedia Computer Science, vol. 245, no. 245, pp. 677–688, 2024. doi:10.1016/j.procs.2024.10.294

[13] M. Liao et al., “Improving the model robustness of flood hazard mapping based on hyperparameter optimization of Random Forest,” Expert Systems with Applications, vol. 241, no. May., pp. 1–20, May 2024. doi:10.1016/j.eswa.2023.122682

[14] Q. Su, L. Chen, and L. Qian, “Optimization of Big Data Analysis Resources supported by XGBOOST algorithm: Comprehensive Analysis of Industry 5.0 and ESG performance,” Measurement: Sensors, vol. 36, no. Dec., pp. 1–10, Dec. 2024. doi:10.1016/j.measen.2024.101310

[15] Y.-C. Shih et al., “Fund transfer fraud detection: Analyzing irregular transactions and customer relationships with self-attention and graph neural networks,” Expert Systems with Applications, vol. 259, no. Jan., pp. 1–14, Jan. 2025. doi:10.1016/j.eswa.2024.125211

[16] M. Hasan, M. S. Rahman, H. Janicke, and I. H. Sarker, “Detecting anomalies in blockchain transactions using machine learning classifiers and explainability analysis,” Blockchain: Research and Applications, vol. 5, no. 3, pp. 1–17, Sep. 2024. doi:10.1016/j.bcra.2024.100207

[17] C. Ni et al., “Light gradient boosting machine (lightgbm) to forecasting data and assisting the defrosting strategy design of refrigerators,” International Journal of Refrigeration, vol. 160, no. Apr., pp. 182–196, Apr. 2024. doi:10.1016/j.ijrefrig.2024.01.025

[18] O. Alshboul, G. Almasabha, A. Shehadeh, and K. Al-Shboul, “A comparative study of lightgbm, XGBoost, and GEP models in shear strength management of SFRC-SBWS,” Structures, vol. 61, no. Mar., pp. 1–17, Mar. 2024. doi:10.1016/j.istruc.2024.106009

[19] W. Gong et al., “Gradient boosting decision tree algorithms for accelerating nanofiltration membrane design and Discovery,” Desalination, vol. 592, no. Dec., pp. 1–12, Dec. 2024. doi:10.1016/j.desal.2024.118072

[20] E. Agliari, F. Alemanno, M. Aquaro, and A. Fachechi, “Regularization, early-stopping and dreaming: A Hopfield-like setup to address generalization and overfitting,” Neural Networks, vol. 177, no. Sep., pp. 1–15, Sep. 2024. doi:10.1016/j.neunet.2024.106389