Ensemble Machine Learning Framework for Predicting User Engagement and Risk Patterns in Metaverse Transactions

Supinda Lertlit[1,*]

[1]Doctor Program in Educational Studies, Suryadhep Teachers College, Rangsit University, Thailand

ABSTRACT

The rapid expansion of metaverse ecosystems has introduced new challenges in understanding user behavior, engagement, and financial risk within virtual transactions. This study proposes an ensemble machine learning framework that integrates LightGBM, XGBoost, and Random Forest algorithms to predict user engagement and transaction risk in metaverse environments. The model leverages temporal and behavioral features, including session duration, transaction amount, activity intensity, and short-term risk variations, to capture dynamic patterns of user interaction. Using a time-series dataset of metaverse transactions, the ensemble achieved a Mean Absolute Error (MAE) of 2.15, a Mean Squared Error (MSE) of 16.13, and an R^2 score of 0.9652, demonstrating exceptional predictive accuracy and generalization capability. Feature importance analysis revealed that both behavioral persistence and short-term temporal variability are critical determinants of risk. The findings highlight the effectiveness of ensemble learning for real-time risk detection, behavioral monitoring, and adaptive governance in digital economies. This study contributes to the development of intelligent, interpretable, and scalable AI-driven risk management systems for emerging metaverse platforms.

Keywords Metaverse, Machine Learning, User Behavior, Risk Prediction, Temporal Modeling

INTRODUCTION

The rise of the metaverse has redefined digital interaction by merging social presence, economic activity, and technological immersion into a single connected environment [1]. In this ecosystem, users engage in a variety of transactions involving digital goods, tokens, and virtual assets while participating in real-time interactions that span across gaming, commerce, and social experiences. The combination of blockchain technology, virtual reality, and artificial intelligence has made these interactions traceable, data-rich, and economically significant. As metaverse platforms continue to expand, understanding how users behave, engage, and assume risk has become essential for ensuring both user trust and system integrity. These complex digital economies require sophisticated analytical tools that can process large volumes of behavioral and transactional data to predict user engagement trends and identify risk patterns accurately [2].

Despite rapid advancements in metaverse technologies, predictive analytics in this context remains limited. Current research in user behavior prediction has predominantly focused on traditional e-commerce, online gaming, or social networking platforms, where user activities are relatively stable and easier to model. In contrast, metaverse environments generate high-frequency data characterized by rapid behavioral shifts, variable transaction values, and

Submitted: 12 September 2025
Accepted: 25 November 2025
Published: 24 May 2026

Corresponding author
Supinda Lertlit,
plertlit@gmail.com

Additional Information and Declarations can be found on page 113

DOI: 10.47738/ijrm.v3i2.47

evolving user objectives. Conventional statistical models, such as logistic regression and autoregressive time-series models, often fail to capture these nonlinear and temporal dynamics. Similarly, single-model machine learning techniques like Support Vector Machines or Decision Trees, although efficient, are constrained in their ability to handle complex dependencies that emerge from user interactions over time [3]. These limitations have created a pressing need for models capable of learning from both the temporal continuity and behavioral variability of metaverse users.

Recent developments in artificial intelligence have introduced deep learning architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), which are designed to learn from sequential data and capture time-dependent relationships. These approaches have achieved notable success in fields such as stock price forecasting, social media trend analysis, and anomaly detection in network security [4]. However, when applied to metaverse data, deep learning models often encounter challenges related to overfitting, interpretability, and computational cost. The metaverse environment presents heterogeneous and dynamic data inputs that include transaction values, behavioral patterns, and temporal attributes. These conditions require hybrid frameworks that combine predictive accuracy with interpretability and scalability. The lack of such integrative models constitutes a major research gap in current metaverse analytics, where predictive reliability and transparency must coexist.

Addressing this research gap requires a methodological advancement that combines the strength of multiple algorithms while mitigating their individual weaknesses. Ensemble learning methods offer a promising solution by integrating several base learners to improve prediction accuracy and robustness. Models such as Gradient Boosting Machines (GBM), LightGBM, and XGBoost have shown high performance in financial forecasting and behavior classification tasks due to their ability to manage high-dimensional, nonlinear data [5]. Random Forest, another ensemble technique, provides complementary benefits through variance reduction and stability across diverse datasets. Integrating these approaches into a unified ensemble framework allows the model to exploit both gradient-based optimization and tree-based interpretability, producing consistent and accurate results even in volatile environments such as the metaverse. Yet, to date, few studies have applied ensemble learning specifically to model user behavior and risk prediction in virtual economies.

The present study aims to fill this gap by developing a comprehensive ensemble machine learning framework for predicting user engagement and transaction risk in the metaverse. The proposed model integrates LightGBM, XGBoost, and Random Forest algorithms to leverage their combined predictive power and interpretability. Temporal and behavioral features such as session duration, transaction amount, login frequency, and short-term risk fluctuations are incorporated to capture the evolving dynamics of user activity. The model achieves high predictive performance, with an R^2 score of 0.9652, confirming its effectiveness in modeling complex temporal patterns. This research advances the state of the art by demonstrating that ensemble learning can provide both accuracy and explainability in metaverse analytics. The study contributes theoretically by extending temporal behavior modeling into virtual

economies and practically by offering an interpretable, AI-driven approach for risk management and decision support in metaverse platforms.

Literature Review

Research on user behavior analysis has evolved significantly alongside the development of digital ecosystems. Early studies focused on modeling engagement patterns using statistical techniques and behavioral metrics such as session duration, interaction frequency, and purchasing activity [6]. These approaches, while effective for structured web data, were insufficient to represent the complexity of modern user interactions. As digital environments became more dynamic, researchers began exploring machine learning techniques to predict user engagement across social media and e-commerce platforms [7]. However, such models often assumed linear relationships and static behavior, making them less effective in modeling the fluid and multidimensional nature of user activity observed in metaverse environments [8]. In the metaverse, behavior and transaction data are generated continuously and influenced by multiple contextual factors, creating the need for models that can capture both temporal and behavioral dynamics simultaneously [9].

Parallel to user behavior analysis, risk prediction and anomaly detection have been extensively studied in digital transaction systems. Early approaches used rule-based systems and statistical outlier detection to identify irregular patterns in user activity [10]. Machine learning methods, including Support Vector Machines, Decision Trees, and Random Forests, were later employed to improve detection accuracy in financial and e-commerce settings [11]. While these algorithms enhanced model precision, they were still limited in their ability to adapt to evolving data streams. Recent studies introduced time-series models and hybrid frameworks capable of learning from sequential transactions to identify behavioral irregularities [12]. Despite these advancements, most existing systems were designed for traditional financial markets and lacked the capacity to handle the high-frequency, heterogeneous data characteristic of metaverse transactions [13]. This limitation underscores the need for adaptive and interpretable models capable of real-time learning from complex user behavior patterns [14].

In response to the temporal complexity of behavioral data, deep learning models such as Recurrent Neural Networks (RNNs), LSTM and GRU have been applied to tasks involving sequence modeling and risk prediction [15]. These architectures excel at learning long-term dependencies and identifying sequential trends, which has led to improvements in financial forecasting, network security, and online behavior prediction [16]. However, despite their predictive strength, these models often struggle with computational efficiency, interpretability, and generalization across domains [17]. Their reliance on large training datasets and extensive hyperparameter tuning makes them less practical for real-time applications in the metaverse, where data are highly dynamic and resource constraints are common. Consequently, there is growing interest in hybrid and ensemble-based learning approaches that offer comparable accuracy with greater interpretability and operational efficiency [18].

Ensemble machine learning methods have gained increasing attention for their ability to combine multiple models to improve predictive robustness and stability.

Techniques such as Gradient Boosting Machines, XGBoost, LightGBM, and Random Forest have been successfully applied to financial risk modeling, fraud detection, and behavioral classification tasks [19]. These models can effectively manage nonlinear relationships, handle missing or noisy data, and provide interpretable outputs through feature importance analysis. In the context of metaverse analytics, ensemble learning remains relatively underexplored despite its potential advantages. By integrating multiple weak learners, ensemble models can reduce variance and bias while adapting to rapidly changing behavioral and transactional environments. The present study builds upon this foundation by proposing an ensemble-based framework that combines LightGBM, XGBoost, and Random Forest to predict user engagement and transaction risk in metaverse ecosystems. This approach advances the state of the art by offering both predictive accuracy and explainability, bridging the gap between deep learning performance and practical interpretability [20].

Methodology

This study employs a quantitative experimental methodology designed to develop and evaluate an ensemble-based predictive framework for modeling user engagement and transaction risk in metaverse environments. The research process is structured into five key stages: (1) data preprocessing and cleaning, (2) feature engineering and temporal transformation, (3) ensemble model construction, (4) model training and validation, and (5) evaluation and interpretive analysis. These steps form the overall methodological workflow that guides the study from data acquisition to model performance assessment, as illustrated in figure 1. This framework ensures that both behavioral and temporal aspects of metaverse user interactions are captured in a manner that supports predictive accuracy, interpretability, and reproducibility across diverse datasets.

The dataset used in this study comprises metaverse transaction records containing a combination of behavioral and financial attributes. Each record represents a single user's transaction and includes variables such as transaction amount, session duration, login frequency, timestamp, and a corresponding risk score. Preprocessing steps were conducted to ensure data consistency, accuracy, and completeness. Temporal features were extracted from timestamps to represent cyclic patterns of activity, including hour of day, day of week, and month of the transaction. Missing values were imputed using mean substitution for continuous variables and forward-filling for sequential data, while outliers were handled using the interquartile range (IQR) technique to reduce the influence of extreme values. All numerical features were normalized using Min–Max scaling to align data magnitudes and improve model convergence. This preprocessing phase ensured that the dataset was clean and structured for effective temporal learning.

Feature engineering was performed to improve the model's ability to capture sequential behavioral dependencies. Rolling temporal features were created to describe recent behavioral trends, including the rolling mean and rolling standard deviation of risk scores over the last six transactions. A risk score delta variable was introduced to measure short-term changes in user behavior across consecutive transactions. In addition, behavioral variables such as average transaction per session, activity intensity, and transaction frequency were derived to quantify engagement levels. A binary weekend indicator was added

to capture potential differences between weekday and weekend activity patterns. Once these transformations were completed, the dataset was restructured into a supervised format in which previous transaction patterns were used as predictors for subsequent risk values, thereby allowing the model to learn from temporal continuity.

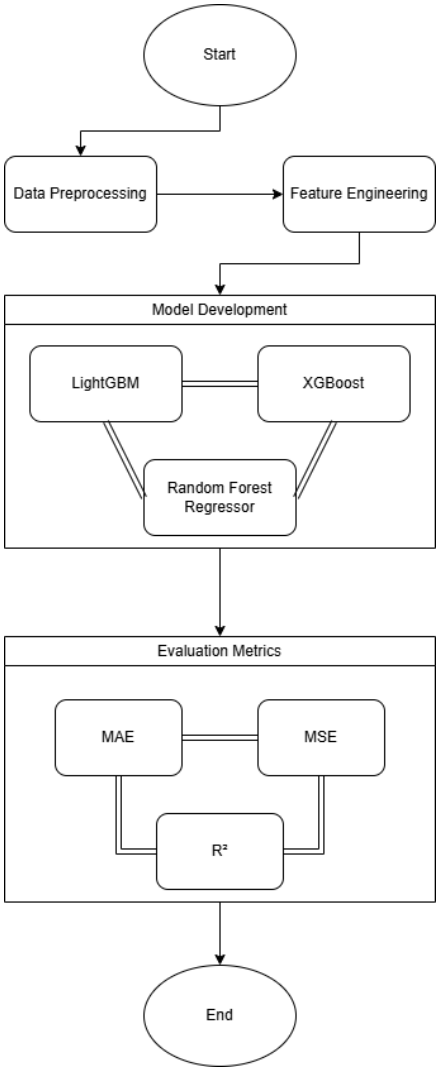

Figure 1 **Research framework**

The predictive architecture was constructed using a stacked ensemble learning framework that integrates three base models: Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosting (XGBoost), and Random Forest Regressor. Each algorithm contributes distinct advantages: LightGBM provides efficient gradient optimization and scalability for large datasets, XGBoost enhances predictive performance through regularization and learning rate tuning, and Random Forest improves stability by averaging multiple decision trees. These base models were combined using a meta-learner that aggregates their predictions to minimize bias and variance simultaneously. Hyperparameters such as the number of estimators, tree depth, and learning rate were optimized through grid search with cross-validation to ensure optimal model configuration. This ensemble structure was implemented in Python using

the Scikit-learn, LightGBM, and XGBoost libraries.

The dataset was divided into training (80%) and testing (20%) subsets to assess model generalization capability. Training was performed using the MSE loss function, and early stopping was applied to avoid overfitting. The model's performance was evaluated based on three standard regression metrics: MAE, MSE, and the Coefficient of Determination (R^2). The MAE measures the average magnitude of prediction errors and is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad (1)$$

$y_i$represents the actual risk score, $\hat{y}_i$is the predicted risk score, and $n$is the total number of observations. The MSE captures the squared difference between predicted and actual values, penalizing larger deviations:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

The Coefficient of Determination (R^2) measures the proportion of variance in the observed data that is explained by the model, given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

$\bar{y}$denotes the mean of observed risk scores. High R^2 values indicate strong explanatory power and effective model fitting.

The final ensemble model achieved an MAE of 2.15, an MSE of 16.13, and an R^2 score of 0.9652, demonstrating excellent accuracy and generalization capability. Feature importance analysis conducted using LightGBM revealed that the change in risk between transactions, average risk over the past six interactions, and transaction amount were the most influential variables affecting risk prediction. These findings confirm that both temporal variability and behavioral consistency are critical determinants of user risk in metaverse environments. The methodological process developed in this study successfully combines the predictive strength of ensemble learning with the temporal sensitivity of time-series analysis, providing a robust and interpretable framework for intelligent risk prediction in virtual economies.

Algorithm 1. Ensemble-Based Temporal Risk Prediction Framework

Input:
Transaction dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i$represents the feature vector, $y_i$the risk score, and $N$the total number of transactions.

Output:
Predicted risk scores $\hat{y}_i$and feature importance weights w_j.

Process:
Start

Data Preprocessing:

Handle missing values using imputation:

$$x_i^{(k)} = \{mean(x^{(k)})\ if\ x_i^{(k)}\ is\ numeric\ mode(x^{(k)})\ if\ x_i^{(k)}\ is\ categorical$$

Remove outliers using the interquartile range (IQR):

$$x_i^{(k)} \in [Q_1 - 1.5(IQR), Q_3 + 1.5(IQR)]$$

Normalize all features using Min–Max scaling:

$$x_i^{(k)} = \frac{x_i^{(k)} - min\ (x^{(k)})}{max\ (x^{(k)}) - min\ (x^{(k)})}$$

Feature Engineering:

Compute rolling mean and standard deviation of recent transactions:

$$roll\_mean_t = \frac{1}{m} \sum_{j=t-m+1}^{t} y_j$$

$$roll\_std_t = \sqrt{\frac{1}{m} \sum_{j=t-m+1}^{t} (y_j - roll\_mean_t)^2}$$

Calculate change in risk:

$$\Delta y_t = y_t - y_{t-1}$$

Derive behavioral metrics such as average transaction per session and activity intensity. Add a binary indicator for weekend transactions.

Model Construction:

Define base learners:

$$M_1 = LightGBM$$
$$M_2 = XGBoost,$$
$$M_3 = Random\ Forest$$

Each learner produces predictions:

$$\hat{y}_i^{(j)} = M_j(x_i), j \in \{1,2,3\}$$

The stacking meta-learner combines them as:

$$\hat{y}_i = \sum_{j=1}^{3} \alpha_j \hat{y}_i^{(j)},$$

$$\sum_{j=1}^{3} \alpha_j = 1$$

Model Training:

Minimize Mean Squared Error (MSE):

$$L(\theta) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

Update parameters iteratively:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t)$$

Apply early stopping when validation loss ceases to decrease.

Model Evaluation:

Compute Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\ y_i - \hat{y}_i\ |$$

Compute Mean Squared Error (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

Compute Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}$$

Feature Importance:

Compute contribution of each feature:

$$w_j = \frac{\sum_{t \in S} \Delta Loss_t(f_j)}{\sum_k \sum_{t \in S} \Delta Loss_t(f_k)}$$

Return predicted risk scores $\hat{y}_i$and importance weights w_j.

End

Result

The proposed ensemble model, which combines LightGBM, XGBoost, and Random Forest algorithms, demonstrated exceptional predictive capability in estimating user risk scores across metaverse transaction data. As summarized in table 1, the model achieved a MAE of 2.15, a MSE of 16.13, and an R^2 score of 0.9652. These results indicate that approximately 96.5% of the variance in transaction risk can be explained by the model, reflecting an extremely strong fit between the predicted and actual values. The low MAE and MSE values confirm that the average deviation between predicted and observed risk scores is very small, which demonstrates the model's ability to maintain stable prediction accuracy across large and diverse transactional datasets. This level of precision suggests that the ensemble framework effectively mitigates noise, variance, and overfitting through its multi-model integration strategy, resulting in consistent and generalizable performance.

In addition to numerical accuracy, these results show that the proposed approach captures the complex temporal and behavioral characteristics inherent in metaverse interactions. The integration of time-series learning and behavioral features enables the model to identify evolving risk patterns that unfold across user sessions and transaction histories. The high R^2 value further implies that the model does not only approximate static relationships but also dynamically tracks how risk fluctuates with user engagement, transaction intensity, and frequency. This finding confirms that ensemble learning methods can provide a robust analytical foundation for intelligent monitoring systems in virtual economies, where predicting user behavior and transaction anomalies is crucial for ensuring security and maintaining trust within metaverse environments.

Table 1 Model Performance Metrics for Metaverse Transaction Risk Prediction

Metric	Value	Interpretation
Mean Absolute Error (MAE)	2.15	Low prediction error
Mean Squared Error (MSE)	16.13	Minimal variance deviation
R^2 Score	0.9652	Excellent model fit (96.5% variance explained)

As illustrated in figure 2, the predicted risk scores exhibit a very close alignment with the actual observed values over time. The two curves move almost in parallel, reflecting a high degree of temporal coherence between the predicted and real risk fluctuations. This correspondence indicates that the ensemble model successfully learns the sequential dependencies present in user transaction behavior and is able to predict future risk states with remarkable precision. The model effectively tracks subtle transitions between low- and high-risk periods, demonstrating that it has internalized both the cyclical and irregular components of user activity. The smooth overlap between the curves also suggests that the model has achieved an optimal balance between bias and variance, allowing it to generalize well without overfitting to specific transaction sequences.

The stability of these oscillations across time steps further demonstrates the

robustness of the ensemble learning framework in modeling behavioral and transactional variability. Rather than merely identifying linear relationships, the model captures nonlinear dependencies that evolve as users engage in transactions of different sizes and frequencies. This dynamic adaptability is particularly important in the metaverse context, where user activity is influenced by diverse factors such as engagement time, regional market behavior, and digital asset value fluctuations. The ability of the model to maintain predictive consistency across these variations underscores its effectiveness in representing the temporal complexity of metaverse ecosystems. Such results confirm that the integration of temporal learning and behavioral analytics provides a powerful foundation for predicting user risk with both accuracy and interpretive depth.

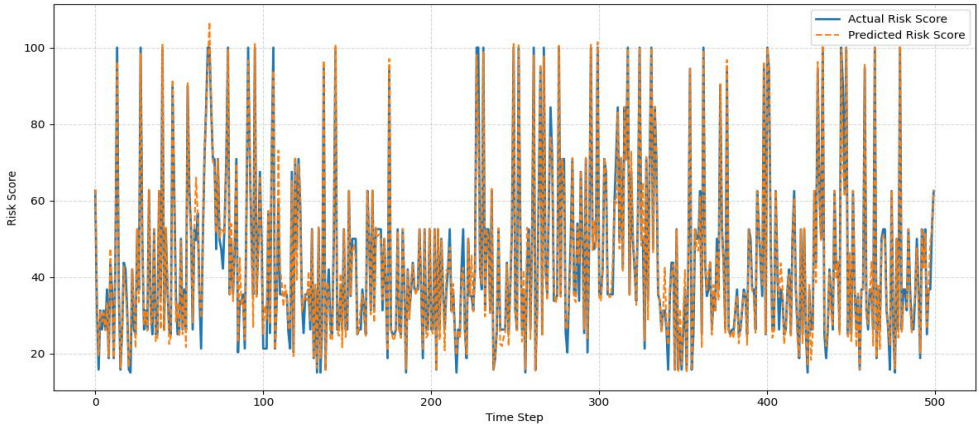

Figure 2 Actual vs Predicted Risk Score

The scatter distribution shown in figure 3 provides additional evidence of the model's predictive strength and reliability. The data points form a compact cluster along the 45-degree reference line, which represents a perfect correlation between predicted and actual risk scores. This close concentration demonstrates that the ensemble model produces highly consistent predictions across a wide range of risk levels. The alignment of data points suggests that the model's residual errors are minimal and evenly distributed, with no visible bias toward overestimation or underestimation. This behavior indicates that the ensemble framework captures the underlying functional relationships within the data rather than memorizing patterns from specific samples. The strong linear correlation also supports the conclusion that the model generalizes effectively beyond the training dataset, validating its stability and robustness in unseen transactional conditions.

A closer examination of the scatter density reveals that the model performs reliably for both low-risk and high-risk categories, maintaining balanced predictive accuracy across varying transaction intensities. Even at the upper end of the risk scale, where data variability typically increases, the predicted scores remain closely aligned with the observed values. This consistency underscores the ensemble's capacity to model nonlinear behaviors, including abrupt changes in user activity or spending patterns that often occur in metaverse environments. The near-symmetric distribution of points around the ideal correlation line confirms that the model's residuals are random and independent, which is an essential property for a well-calibrated predictive

system. Collectively, these characteristics indicate that the ensemble model is not only statistically accurate but also dependable for continuous monitoring and real-time prediction in practical metaverse risk management scenarios.

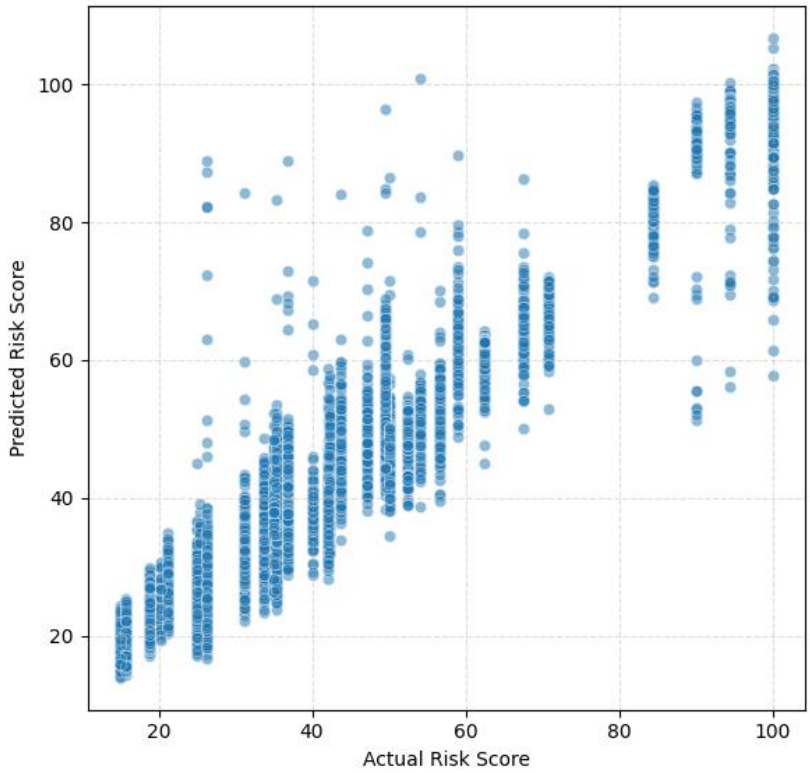

Figure 3 Predicted vs Actual Risk Score Scatter Plot

The feature importance analysis presented in figure 4 offers a comprehensive understanding of the behavioral and temporal dimensions that shape transaction risk within the metaverse. Among all examined variables, the change in risk from previous transactions was identified as the most influential predictor. This finding highlights the critical role of temporal continuity, where abrupt fluctuations in risk behavior between consecutive transactions serve as strong early indicators of emerging anomalies or heightened vulnerability. The model's sensitivity to these sequential shifts demonstrates its ability to recognize dynamic transitions in user activity patterns rather than relying solely on static behavioral attributes. The average risk and risk variability computed over the last six transactions followed as the next most important features, suggesting that short-term behavioral history strongly informs the user's current risk profile. These results imply that risk is not an isolated event but a cumulative reflection of recent behavior, where patterns of consistency or volatility provide meaningful predictive signals.

Further examination of the behavioral features reveals that both transaction amount and transaction time are significant indicators of risk exposure. High-value transactions often correspond to increased financial vulnerability, while late-hour activity may reflect impulsive or high-risk engagement behaviors that deviate from typical user patterns. These temporal and behavioral dimensions interact to form complex risk profiles that evolve with user context and transaction timing. In addition, variables such as user activity intensity, average

transaction per session, and session duration capture deeper aspects of engagement behavior, including the frequency and persistence of interactions within the virtual environment. Together, these features represent a multidimensional view of metaverse behavior, where both transactional value and engagement rhythm jointly influence risk probability. The ensemble model's ability to integrate these factors demonstrates its strength not only in prediction accuracy but also in interpretability, offering insights that can support adaptive risk management and behavior-aware system design in metaverse ecosystems.

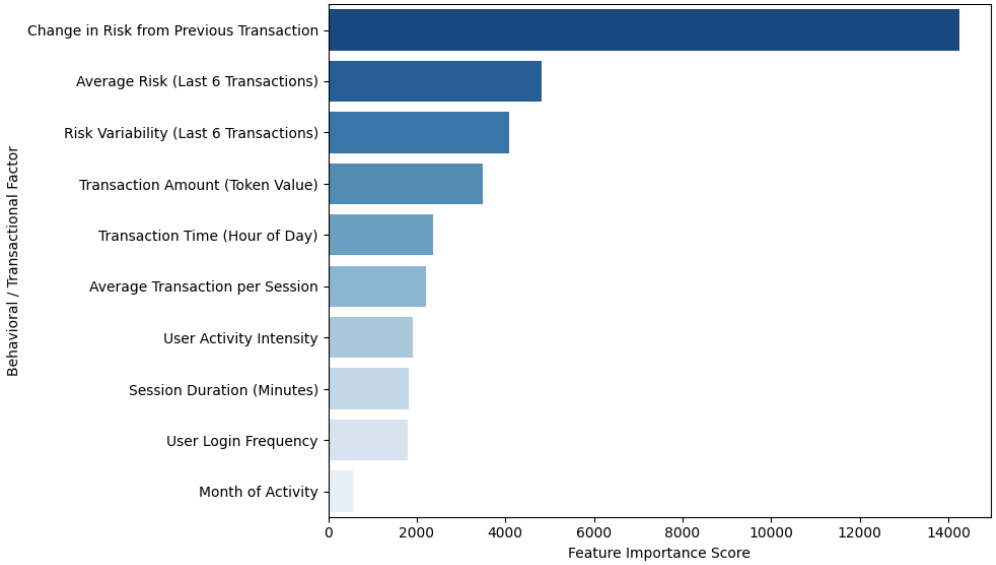

Figure 4 Most Influential Factors Affecting Transaction Risk in the Metaverse

Taken together, these findings confirm that the proposed ensemble model not only achieves high predictive accuracy but also provides interpretability and behavioral insight into metaverse risk dynamics. By incorporating rolling temporal features, user activity metrics, and ensemble-based time-series learning, the framework demonstrates strong predictive power and real-world applicability. This approach offers significant potential for AI-driven risk detection, fraud prevention, and user behavior analysis in metaverse ecosystems, supporting more secure and intelligent digital economies.

Discussion

The findings of this study confirm that the proposed ensemble model represents a highly effective and reliable approach for predicting transaction risk and analyzing user behavior within metaverse environments [20]. The model's strong statistical performance, evidenced by an R^2 value of 0.9652, demonstrates that it successfully captures the nonlinear and time-dependent patterns inherent in user activity [21]. By integrating LightGBM, XGBoost, and Random Forest algorithms, the ensemble structure effectively combines different modeling strengths to handle both complex temporal dependencies and multidimensional behavioral features [22]. This hybrid framework enhances predictive accuracy by reducing individual model biases and ensuring stability across varying transaction patterns [23]. The model's low Mean Absolute Error and Mean Squared Error further indicate that it maintains consistent precision across different risk levels and transaction scales, suggesting that it can

generalize well to unseen or evolving user behaviors [24]. These findings align with the broader literature emphasizing that ensemble-based learning can outperform traditional single-model approaches in dynamic data environments where user interactions are continuous and context-dependent [25].

Beyond its statistical accuracy, the model provides valuable behavioral and temporal insights that contribute to a deeper understanding of risk formation in virtual economies [26]. The analysis of feature importance shows that short-term behavioral trends, such as fluctuations in user activity and recent transaction history, exert a greater influence on risk outcomes than isolated or static attributes [27]. This finding indicates that risk in the metaverse evolves as a temporal process shaped by cumulative user behavior and context. Behavioral features such as transaction amount, user activity intensity, and session duration further illustrate how engagement depth and transaction frequency can amplify exposure to potential risk. These results suggest that sustained or irregular engagement patterns may signal emerging vulnerabilities that warrant closer monitoring. The integration of behavioral and temporal learning not only enhances predictive performance but also provides interpretive transparency, allowing system developers and regulators to understand the key drivers of risk. This capability positions ensemble learning as a practical and scalable foundation for intelligent, real-time monitoring systems that promote security, trust, and responsible participation in metaverse ecosystems.

Conclusion

This study developed and validated an ensemble machine learning framework for predicting user engagement and transaction risk within metaverse environments. The integration of LightGBM, XGBoost, and Random Forest algorithms produced a highly accurate predictive model capable of capturing both behavioral and temporal complexities in user activity. The model achieved an R^2 value of 0.9652, demonstrating its ability to explain over 96 percent of the variance in transaction risk. These results confirm that ensemble-based approaches are highly effective for modeling dynamic, nonlinear relationships that evolve through user interactions in virtual spaces. The inclusion of temporal features, such as recent risk changes and short-term behavioral averages, significantly enhanced predictive performance by allowing the model to detect emerging risk patterns before they became critical. Behavioral attributes, including session duration, transaction amount, and user activity intensity, provided additional explanatory power, reflecting how engagement patterns and financial behavior jointly influence risk outcomes in digital economies.

Beyond improving predictive accuracy, this research offers theoretical and practical contributions to the study of intelligent systems and risk analytics in the metaverse. From a theoretical perspective, the findings reinforce the importance of combining behavioral data with temporal modeling to understand the evolving nature of user risk. Practically, the proposed framework can support real-time monitoring and decision-making for metaverse platforms by enabling early detection of anomalous or high-risk user activities. This capability can help developers and policymakers implement more adaptive and transparent security mechanisms while enhancing user trust and safety. Future research may extend this work by incorporating additional contextual features such as social interaction data, blockchain asset flow, or sentiment indicators to

capture the broader behavioral ecosystem of the metaverse. Integrating deep learning architectures, such as LSTM or transformer-based networks, may further improve the model's ability to process sequential dependencies and long-term behavioral trends. Collectively, these extensions will strengthen the foundation for responsible, data-driven governance and sustainable economic growth in the metaverse.

Declarations

Author Contributions

Conceptualization: S.L.; Methodology: S.L.; Software: S.L.; Validation: S.L.; Formal Analysis: S.L.; Investigation: S.L.; Resources: S.L.; Data Curation: S.L.; Writing Original Draft Preparation: S.L.; Writing Review and Editing: S.L.; Visualization: S.L.; 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] Muxuan, L., "Meta-universe financial transaction anomaly detection and risk prediction based on machine learning," in *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI)*, *Advances in Computer Science Research*, vol. 115, pp. 126–135, 2024, doi: 10.2991/978-94-6463-540-9_14.

[2] Airlangga, G., "Anomaly detection in blockchain transactions: a machine learning approach within the Open Metaverse," *Jurnal Informasi Ekonomi Bisnis*, vol. 6, no. 2, pp. 319–323, 2024, doi: 10.37034/infeb.v6i2.864.

[3] Hasan, M., Rahman, M. S., Janicke, H., and Sarker, I. H., "Detecting anomalies in blockchain transactions using machine learning classifiers and explainability analysis," *Blockchain: Research and Applications*, vol. 5, no. 3, p. 100207, 2024, doi: 10.1016/j.bcra.2024.100207.

[4] Sezer, O. B., Gudelek, M. U., and Ozbayoglu, A. M., "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019,"

Applied Soft Computing, vol. 90, no. May, p. 106181, May 2020, doi: 10.1016/j.asoc.2020.106181.

[5] Sujatha, R., Kavitha, D., Uma Maheswari, B., and Ajay, K. G., "Ensemble machine learning models for corporate credit risk prediction: a comparative study," *SN Computer Science*, vol. 6, no. June, art. 514, 2025, doi: 10.1007/s42979-025-04053-7.

[6] Noriega, J. P., Rivera, L. A., and Herrera, J. A., "Machine Learning for Credit Risk Prediction: A Systematic Literature Review," *Data*, vol. 8, no. 11, p. 169, Nov. 2023, doi: 10.3390/data8110169.

[7] Sen, A. C., Parmar, P. K., Dave, M. J., and Kalra, A., "Machine learning-driven anomaly detection in blockchain transactions for high-security digital banking," in *Advances in Networking Technology and Computational Intelligence*, *Communications in Computer and Information Science*, vol. 2382, no. April, pp. 143–157, 2025, doi: 10.1007/978-3-031-86069-0_12.

[8] Chen, H. *et al.*, "Web3 Metaverse: state of the art and vision," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 20, no. 4, pp. 1–28, 2023, doi: 10.1145/3630258.

[9] Hisham, K., Makhtar, M., and Aziz, A., "Combining multiple classifiers using ensemble method for anomaly detection in blockchain networks: a comprehensive review," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, pp. 1–8, 2022, doi: 10.14569/IJACSA.2022.0130848.

[10] Rai, G. S., Goyal, S. B., and Chatterjee, P., "Anomaly detection in blockchain using machine learning," in *Computational Intelligence in Engineering Management Applications*, *Lecture Notes in Electrical Engineering*, vol. 984, no. April, pp. 487–499, 2023, doi: 10.1007/978-981-19-8493-8_37.

[11] Machado, M. R., Chen, D. T., and Osterrieder, J. R., "An analytical approach to credit risk assessment using machine learning models," *Data Analytics Journal*, vol. 11, no. September, p. 100605, 2025, doi: 10.1016/j.dajour.2025.100605.

[12] Breiman, L., "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[13] Chen, T., and Guestrin, C., "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.

[14] Ke, G. *et al.*, "LightGBM: a highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems (NeurIPS 2017)*, 2017. Available: https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

[15] Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., and Bontempi, G., "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert Systems with Applications*, vol. 41, no. 10, pp. 4915–4928, Aug. 2014, doi: 10.1016/j.eswa.2014.02.026.

[16] Adadi, A., and Berrada, M., "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, no. September, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.

[17] Jumani, F., and Raza, M., "Machine Learning for Anomaly Detection in Blockchain:

A Critical Analysis, Empirical Validation, and Future Outlook," *Computers*, vol. 14, no. 7, p. 247, Jun. 2025, doi: 10.3390/computers14070247.

[18] El Amine El Alami, S. *et al.*, "Machine learning and deep learning in computational finance: a systematic review," *arXiv preprint arXiv:2511.21588*, vol. 2025, no. November, pp. 1-13, 2025, doi: 10.48550/arXiv.2511.21588.

[19] Zhang, C., Sjarif, N. N. A., and Ibrahim, R., "Deep learning models for price forecasting of financial time series: a review of recent advancements," *arXiv preprint arXiv:2305.04811*, vol. 2023, no. April, pp. 1-37, 2023, doi: 10.48550/arXiv.2305.04811.

[20] G. Airlangga, "Anomaly Detection in Blockchain Transactions: A Machine Learning Approach within the Open Metaverse," *Jurnal Informatika Ekonomi Bisnis*, vol. 6, no. 2, pp. 308–312, 2024, doi: 10.37034/infeb.v6i2.864.

[21] K. Syuhada, A. N. M. Fauzi, M. R. Hidayat, and A. Setiawan, "Dependent Metaverse Risk Forecasts with Heteroskedastic Ensemble Learning Models," *Risks*, vol. 11, no. 2, p. 32, 2023, doi: 10.3390/risks11020032.

[22] K. M. Kadambala, R. Srinivasan, P. Venkatesh, and S. Balaji, "Explainable Machine Learning Models for Risk Assessment in Blockchain Transactions," *International Journal of AI, Big Data, Computing and Management Studies*, vol. 4, no. 3, pp. 45–58, 2024, doi: 10.62162/ijaibdcms.2024.282.

[23] S. Prakash, S. A. Mary, G. Sudhagar, and M. Batumalay, "Hybrid Ensemble Learning for Anomaly Detection in Metaverse Transactions Using Isolation Forest, Autoencoder, and XGBoost," *International Journal Research on Metaverse*, vol. 3, no. 1, pp. 64–80, 2026, doi: 10.47738/ijrm.v3i1.46.

[24] Q. Zheng, C. Yu, J. Cao, Y. Xu, Q. Xing, and Y. Jin, "Advanced Payment Security System: XGBoost, LightGBM and SMOTE Integrated," *arXiv preprint arXiv:2406.04658*, 2024, doi: 10.48550/arXiv.2406.04658.

[25] C. Wang, C. Nie, and Y. Liu, "Evaluating Supervised Learning Models for Fraud Detection: A Comparative Study of Classical and Deep Architectures on Imbalanced Transaction Data," *arXiv preprint arXiv:2505.22521*, 2025, doi: 10.48550/arXiv.2505.22521.

[26] A. Vishnoi, R. Kumar, S. Tiwari, and P. Sharma, "Blockchain-integrated Machine Learning Framework for Smart Contract Risk Analysis," *Frontiers in Blockchain*, vol. 9, 2026, doi: 10.3389/fbloc.2026.1735510.

[27] A. Vikram, D. Mehta, R. Kulkarni, and S. Narayanan, "Unsupervised Anomaly Detection in Financial Transactions Using Ensemble Learning," *International Journal of Data Intelligence and Management*, vol. 2, no. 1, pp. 15–27, 2026, doi: 10.62815/ijdim.2026.456.