Scholar Publishing Group
International Journal of Social Sciences and Economic Management
https://doi.org/ 10.38007/IJSSEM.2026.070102
ISSN 2789-715X Vol. 7, Issue 1: 10-18

Research on Abnormal Detection and Transaction Risk Management Based on Machine Learning

Shuang Yuan

American Airlines, Technology Risk Management, Fort Worth, Texas, 76155, United States

Keywords: Machine learning; Anomaly detection; Transaction risk management; Risk prediction; Dynamic adjustment

Abstract: In the modern complex trading environment, trading risk management is facing unprecedented challenges. The research of anomaly detection and transaction risk management technology based on machine learning, by optimizing the traditional risk management model and introducing machine learning algorithm, improves the accuracy of risk prediction, the intelligence of trading strategy and the real-time monitoring and early warning. This paper discusses the optimal design of transaction risk management model, including the construction of risk prediction, evaluation model, trading strategy and risk control. This paper analyzes the anomaly detection methods based on machine learning, focusing on statistics-driven anomaly detection, machine learning algorithm and anomaly recognition technology based on neighborhood relationship. The results show that machine learning techniques can significantly improve the efficiency and precision of risk management, helping to deal with uncertainty in complex markets.

Introduction

With the increasingly complex market trading environment, traditional risk management methods have gradually exposed their limitations, especially in the face of massive data and real-time decision-making needs, often unable to provide timely and accurate risk prediction and control. The rapid development of machine learning technology provides a new way to improve the level of transaction risk management. Through machine learning, we can use large amounts of historical data for self-learning and model optimization, so as to achieve more accurate risk prediction and intelligent risk control. This paper discusses the anomaly detection and transaction risk management technology based on machine learning, and studies its application in the identification, assessment, control and dynamic adjustment of transaction risk, so as to provide more efficient and reliable technical support for transaction risk management.

1. Optimal design of transaction risk management model

1.1. Risk prediction and evaluation model

In the trading risk management, accurate risk prediction and evaluation is very important to avoid loss. Traditional risk forecasting methods tend to adopt static models based on historical data, which makes them less stable in the face of market fluctuations, so these methods need to be optimized. The core of the optimization is the introduction of machine learning technology to build a more detailed dynamic risk assessment model by deeply analyzing multidimensional information such as historical trading data, market sentiment, and macroeconomic indicators. That is, advanced technologies such as integrated learning and deep neural networks are used to enhance the accuracy of risk prediction. For example, using time series models such as Long short-term memory Network (LSTM) can effectively capture the long-term trends and short-term volatility characteristics of the market. In addition, the model needs to be adaptive, able to flexibly adjust to market changes based on real-time data, and identify new risk signals. By increasing the integration of data dimensions and model optimization, trading risk prediction can be more accurate and timely, thus providing more powerful support for strategy adjustment and decision-making, and effectively preventing potential market risks (see Figure 1).

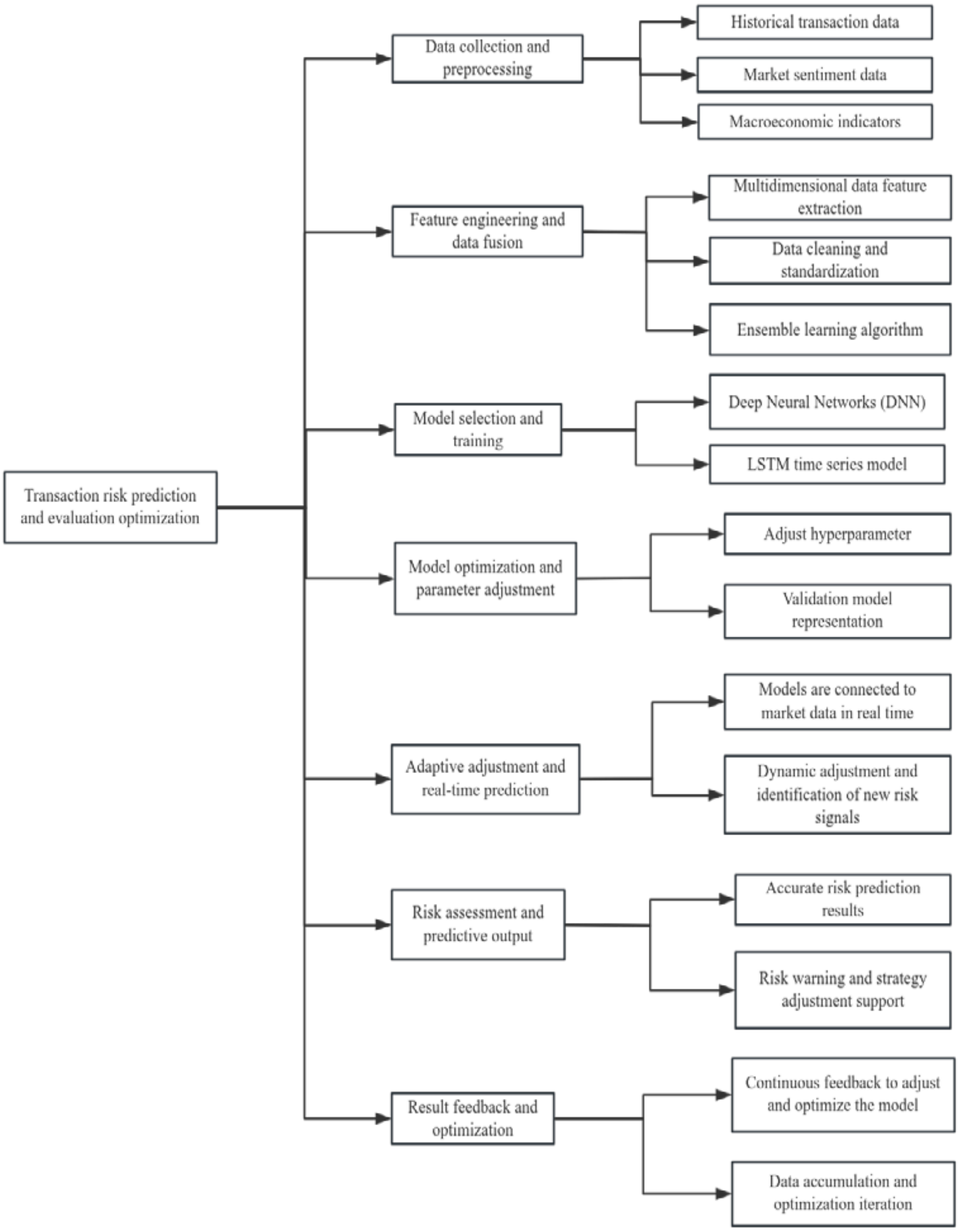

Figure 1 Each link of risk prediction optimization by machine learning technology

1.2. Trading strategy and risk control

Optimizing trading strategy and risk control is the core task of trading risk management. Compared with the traditional trading strategies that rely on fixed rules and manual experience, the optimized strategies are more dynamic, intelligent and automated. Through reinforcement learning methods in machine learning, trading decisions can automatically adapt to the market environment, self-adjusting through constant interaction with the market. In the process of optimization, comprehensive factors including stop loss setting, position management, fund allocation, and trading frequency should be considered in multiple dimensions. The algorithm can quickly identify potential opportunities in the market, while controlling risks and reducing losses caused by violent market fluctuations. The machine learning model is able to adjust itself to market changes, such as dynamically setting key parameters such as stop-loss points and retracement thresholds based on historical data and real-time market movements. Through backtesting techniques, the historical performance of different strategies can be verified and optimized to select the most stable trading strategy with the highest earning potential. The optimized trading strategy can improve profitability, effectively reduce potential risks, and form a comprehensive risk control mechanism.

1.3. Transaction monitoring and early warning system

The optimization of the transaction monitoring and early warning system is not only limited to improving the abnormal recognition ability, but also to enhance the real-time response and intelligent decision-making function of the system. Through anomaly detection techniques in machine learning, the optimized monitoring system can identify subtle fluctuations and potential risks in the market, not just relying on static rules or thresholds. For example, the use of deep learning models to analyze a large number of historical trading data can capture complex rules in trading, timely detect violations such as price manipulation and insider trading, and quickly launch an early warning mechanism. In addition, the optimized early warning system can not only monitor trading activities in real time, but also analyze market trends through data mining technology to anticipate potential risks in advance. The system will adjust monitoring strategies based on real-time data to enable more accurate risk classification and response. With the introduction of an intelligent decision-making layer, the early warning system can not only automatically adjust the risk level, but also provide countermeasures, such as automatically suspending certain transactions, notifying relevant decision makers, or directly implementing risk prevention and control measures. These optimizations enable the transaction monitoring and early warning system to identify and deal with various transaction risks more efficiently under the changing market environment, thus improving the security and stability of the transaction process.

2. Analysis of anomaly detection methods based on machine learning

2.1. Driving statistical anomaly detection methods

The anomaly detection methods that drive statistics mainly rely on the statistical characteristics of the data to identify anomalies. Traditional statistical methods can identify potential anomalies by modeling the distribution of data and judging whether data points deviate from the normal range according to specific statistics (such as mean value, standard difference, etc.). The advantage of this method is that the calculation is simple and easy to understand, especially for the data with obvious distribution law. Common statistical anomaly detection methods include threshold detection based on Z-score, mean and standard deviation, and methods based on normal distribution hypothesis. In practical applications, Z-score detection is one of the most common statistical methods. By calculating the Z-score of the data point, that is, the ratio of the deviation between the data point and the mean and the standard deviation, the abnormal degree of the data point is determined. If the Z-score of a data point exceeds a preset threshold limit, it is marked as an anomaly. Specifically, the Z-score method is suitable for situations where the data follows a normal distribution. If the Z-score of a data point is outside the ± 3 range, it is considered an anomaly. In addition, the threshold method based on mean value and standard deviation also determines whether the data point exceeds the preset range by setting the upper and lower limits of the data, so as to detect anomalies. This method is suitable for cases where the data volume is small or the data distribution is relatively simple. However, in the face of complex data sets, it may produce a high error rate. Statistical anomaly detection also includes the normal distribution hypothesis method, which assumes that the data follows a normal distribution and calculates the degree of anomaly of a data point by estimating the mean and standard deviation of the data. However, the hypothesis of normal distribution is often difficult to match the complex data structure, especially in the actual scene, the data is often multi-modal distribution or affected by multiple factors. The following table describes the main anomaly detection methods based on driving statistics and their usage (see Table 1).

Table 1 Anomaly detection methods based on driving statistics

Detection method	Application situation
Z-score	It is suitable for data following normal distribution and Z-score exceeding ± 3 is abnormal.
Mean and standard deviation threshold detection	It is suitable for small data volume or simple data distribution.
Normal distribution hypothesis method	It is suitable for data following normal distribution, but it is less effective for complex distribution.

2.2. Anomaly detection algorithm using machine learning

Anomaly detection algorithm using machine learning identifies anomalies through automatic learning of data features. Compared with traditional statistical methods, machine learning algorithm has higher adaptability and accuracy. Among common machine learning algorithms, support vector machines (SVMS), clustering methods, and neural networks are common algorithms for anomaly detection, capable of handling nonlinear data, high-dimensional data, and complex anomaly patterns. SVM is a supervised learning algorithm that distinguishes between normal and abnormal data by constructing an optimal hyperplane. By maximizing the interval between the two types of data points, SVM can achieve effective anomaly recognition in complex data Spaces. Clustering algorithms such as K-means and DBSCAN are equally important in detecting anomalies. K-means marks data points that cannot be classified into any cluster as anomalies through the clustering process, while DBSCAN uses density changes to identify anomalies. It can automatically identify noise points and outliers in data without pre-setting the number of clusters. Strong adaptability when dealing with complex data. In addition, the autoencoder in the neural network reconstructs data in the low-dimensional space by learning the low-dimensional representation of the data. If the reconstruction error is large, it indicates that the data point may have a large difference in distribution from the training data, which may be an anomaly. Autoencoders show good adaptive ability when dealing with high and complex data sets. By learning the potential structure and characteristics of data, machine learning methods can realize efficient identification of anomalies, especially in the face of complex and high-dimensional data, which has advantages that traditional statistical methods cannot match.

2.3. Anomaly recognition technology based on neighborhood relation

The anomaly recognition technology based on neighborhood relationship identifies the anomaly by analyzing the relationship between data points and their neighbors. It is assumed that normal data points have similar properties in their neighborhood, while outliers are significantly different from other points in their neighborhood. The method is based on unsupervised learning and does not require pre-labeling of the data, so it shows a high degree of adaptability in practical operation. K-nearest neighbor (KNN) is the most commonly used anomaly detection method based on neighborhood relations. Its core idea is that by calculating the spatial distance between each data point and its K neighboring points, if a point is too far away from other points, then that point is judged to be an anomaly. KNN algorithm is simple and clear, can effectively identify isolated data points, and is especially suitable for anomaly detection in low-dimensional data. However, the efficiency of KNN algorithm is low, and its calculation cost is relatively high when facing a large number of data sets. Local outlier factor (LOF) algorithm also plays an important role in anomaly detection based on neighborhood relation. LOF determines the anomaly degree of a data point by calculating the density difference between the data point and other points in its neighborhood. If a point has a much lower density than other points in its neighborhood, it is considered an outlier (see Figure 2). LOF algorithm can deal with density changes in data well and has strong adaptability. Especially for data sets with large density differences, LOF algorithm can effectively identify local anomalies. These neighbor-based anomaly recognition techniques show high accuracy and adaptability when dealing with local density variations and irregular data sets, and are especially suitable for unsupervised anomaly detection tasks.

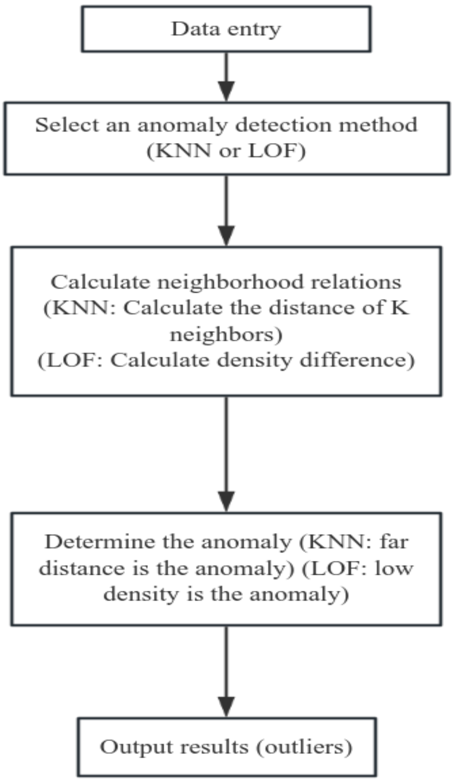

Figure 2 Anomaly recognition technology design based on neighborhood relationship

3. Research on transaction risk management technology based on machine learning

3.1. Transaction risk identification and monitoring

In financial transactions, transaction risk identification and monitoring is an important link to ensure the accuracy and timeliness of investment decisions. Market instability, liquidity problems of funds, credit risk and other factors may threaten the accuracy and timeliness of trading decisions, so accurate and flexible risk monitoring technology is needed. Faced with the explosive growth of data volume, traditional manual monitoring methods have been unable to meet the needs of rapid identification. At this point, risk identification and monitoring methods based on machine learning show their advantages, which can automate the processing of massive data and discover potential market risks. The application of machine learning to trading risk mainly analyzes historical data to learn risk patterns. For example, support vector machines (SVM) can be used to effectively distinguish regular transactions from abnormal transactions by constructing decision boundaries. When SVM recognizes that a data point has become too far away from other trading points, it will mark it as an anomaly, and then issue a market risk warning. This approach not only identifies a single point of risk, but also provides traders with a clear risk range to help them make timely adjustments. In addition to support vector machines, ensemble learning methods are also widely used in risk identification. Through the combination of multiple models, ensemble learning can improve the classification accuracy and reduce the possible errors caused by a single model. For example, Random Forest algorithm can judge the normality of transaction data through the voting mechanism of multiple decision trees, which improves the robustness and anti-interference ability of recognition results. In specific applications, the combination of multiple strategies and real-time market data can detect and report abnormal fluctuations in a timely manner, helping traders to avoid risk. At the same time, anomaly detection methods are also commonly used for risk identification of unlabeled data. These methods automatically identify data points that differ significantly from standard data, helping trading systems identify potential market anomalies and provide early warning of possible market volatility. The decision function formula of support vector machine is as follows:

$$f(x)=\sum_{i=1}^{n}\alpha_i y_i K(x_i,x)+b \tag{1}$$

x_i is the sample point, y_i is the category label of the sample, α_i is a Lagrange multiplier of the support vector, $K(x_i,x)$ is kernel function, b is the offset term。The formula can be used to calculate the probability that a data point falls into a certain category, thereby identifying potential risks.

3.2. Risk control and dynamic adjustment strategy

Risk control and dynamic adjustment strategy is one of the key technologies to ensure the safety of trading. Traditional methods of risk control often rely on fixed thresholds and rules and cannot effectively respond to rapidly changing market dynamics. However, the dynamic adjustment strategy based on machine learning can dynamically adjust the risk control strategy according to real-time data, market changes and other factors. Especially in the field of Reinforcement Learning (RL), it has shown significant advantages in transaction decision-making through reward mechanism and feedback information. Reinforcement learning can optimize trading actions by adjusting key parameters such as position, stop loss, and trading frequency in real time according to market fluctuations. This process not only increases the flexibility and adaptability of the strategy, but also gradually improves the decision-making effect in the process of continuous learning. For example, when the market volatility increases, the reinforcement learning strategy automatically increases the stop loss point to avoid losses. When the market is stable, the stop loss range will be reduced accordingly to expand the profit space. Through continuous interaction with the market, reinforcement learning helps to optimize trading strategies, improve the long-term returns of trading, and effectively reduce potential risks. The strategy updating formula in reinforcement learning is as follows:

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha\left[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(s_t, a_t)\right] \tag{2}$$

$Q(s_t, a_t)$ represent current status s_t take action next a_t the value of, α is the learning rate, R_{t+1} it's a reward for a current action, γ is the discount factor。 In this way, the model is able to adjust the strategy at each step to optimize risk management.

3.3. Time series model and risk prediction

The application of time series model in trading risk prediction has become one of the mainstream technologies, especially in dealing with complex market fluctuations, which can provide accurate forecasting support. Traditional time series analysis methods, such as autoregressive model and moving average model, mainly rely on fixed historical data and basic statistical relationships to make predictions. However, these methods have some limitations when dealing with complex and interrelated market data with nonlinear characteristics. In order to overcome these problems, long short-term memory network (LSTM) has been widely used in financial time series prediction. Through its unique gating mechanism, LSTM network can effectively capture the long-term and short-term dependence relationship and automatically learn the law of market fluctuations. For example, in actual trading, LSTM can predict the future trend of the market based on historical price, volume and other information, so as to provide a basis for traders to make risk management decisions. By continuously adjusting network parameters, LSTM is able to adapt more flexibly to changes in the market, thus improving the accuracy of forecasts. In addition, combined with other machine learning techniques, multi-dimensional analysis of market risk can be achieved to improve the overall forecasting ability. The basic output formula of LSTM is as follows:

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b) \tag{3}$$

h_t is the output of the current time step, h_{t-1} is the state of the last time step, x_t is the input data of the current time step, W_h, W_x is the weight matrix, b is the offset term, σ is activation function。 Through this structure, LSTM can capture the long-term dependence of time series data during the learning process and provide accurate information for trading risk prediction.

Conclusion

Anomaly detection and transaction risk management technology based on machine learning has demonstrated its great potential in improving the efficiency, accuracy and intelligence of transaction risk management. Through the in-depth analysis of the optimal design of transaction risk management model and anomaly detection methods, this paper not only illustrates the important role of machine learning in risk prediction, assessment and control, but also reveals its challenges and development space in practical applications. Nevertheless, the application of machine learning technology in trading risk management still faces some challenges, such as the quality of data, the complexity of the model, and so on, which need to be further optimized. In the future, with the continuous progress of technology, machine learning will play a more important role in trading risk management, further promoting the intelligence and automation of the trading process, and improving the stability and security of the market.

References

[1] *James H. Managing global property portfolios, transactions and teams: A primer for corporate real estate executives — Part 4: Transactional Risk Management in Multinational Context. Corporate Real Estate Journal, 2021, 10(4):249-264.*

[2] *Lachance M. Night trading: Lower risk but higher returns?. Review of Financial Economics, 2023, 41(4):347-363.*

[3] *Esmaeil V, Ramin N, Kamran T, et al. Risk Assessment of Industrial Energy Hubs and Peer-to-Peer Heat and Power Transaction in the Presence of Electric Vehicles. Energies, 2022, 15(23):8920-8920.*

[4] *Goldwater M, Zitterbart D P, Wright D, et al. Machine-learning-based simultaneous detection and ranging of impulsive baleen whale vocalizations using a single hydrophone. The Journal of the Acoustical Society of America, 2023, 153(2):1094-1107.*

[5] *Fernando Hidalgo-Mompeán, Juan Francisco Gómez Fernández, Gonzalo Cerruela-García, et al. Dimensionality analysis in machine learning failure detection models. A case study with LNG compressors. Computers in Industry, 2021, 128(11):103434.*

[6] *Xiu, L. (2025, June). Research on Personalized Recommendation Algorithms in Modern Distance Education Systems. In 2025 IEEE 3rd International Conference on Image Processing and Computer Applications (ICIPCA) (pp. 2019-2024). IEEE.*

[7] *Xu, D. (2025). Integration and Optimization Strategy of Spatial Video Technology in Virtual Reality Platform. International Journal of Engineering Advances, 2(3), 131-137.*

[8] *Huang, J. (2025). Adaptive Reuse of Urban Public Space and Optimization of Urban Living Environment. International Journal of Engineering Advances, 2(4), 9-17.*

[9] *Zhou, Y. (2025). Using Big Data Analysis to Optimize the Financing Structure and Capital Allocation of Energy Enterprises. Economics and Management Innovation, 2(7), 8-15.*

[10] *Zhang, Q. (2025). Use Computer Vision and Natural Language Processing to Optimize Advertising and User Behavior Analysis. Artificial Intelligence and Digital Technology, 2(1), 148-155.*

[11] *Wang, Y. (2025). Research on Early Identification and Intervention Techniques for Neuromuscular Function Degeneration. Artificial Intelligence and Digital Technology, 2(1), 163-170.*

[12] *Wu, H. (2025). The Challenges and Opportunities of Leading an AI ML Team in a Startup. European Journal of AI, Computing & Informatics, 1(4), 66-73.*

[13] *Ren, B. (2025). Cross Modal Data Understanding Based on Visual Language Model. European Journal of AI, Computing & Informatics, 1(4), 81-88.*

[14] *Shen, D. (2025). Innovative Application of AI in Medical Decision Support System and Implementation of Precision Medicine. European Journal of AI, Computing & Informatics, 1(4), 59-65.*

[15] *Liu, X. (2025). Use Generative Al and Natural Language Processing to Improve User Interaction Design. European Journal of AI, Computing & Informatics, 1(4), 74-80.*

[16] *Liu, F. (2025). Localization Market Expansion Strategies and Practices for Global E-commerce Platforms. Strategic Management Insights, 2(1), 146-154.*

[17] Hu, Q. (2025). Research on the Combination of Intelligent Management of Tax Data and Anti-Fraud Technology. Strategic Management Insights, 2(1), 139-145.
[18] Hua, X. (2025). Key Indicators and Data-Driven Analysis Methods for Game Performance Optimization. European Journal of Engineering and Technologies, 1(2), 57-64.
[19] Ren, B. (2025). Deep Learning and Anomaly Detection in Predictive Maintenance Platform. European Journal of Engineering and Technologies, 1(2), 65-71.
[20] Hui, X. (2025). Research on the Application of Integrating Medical Data Intelligence and Machine Learning Algorithms in Cancer Diagnosis. International Journal of Engineering Advances, 2(3), 101-108.
[21] Hui, X. (2025). Utilize the Database Architecture to Enhance the Performance and Efficiency of Large-Scale Medical Data Processing. Artificial Intelligence and Digital Technology, 2(1), 156-162.
[22] Chuying Lu. Object Detection and Image Segmentation Algorithm Optimization in High-Resolution Remote Sensing Images. International Journal of Multimedia Computing (2025), Vol. 6, Issue 1: 144-151.
[23] Buqin Wang. Research on Load Balancing Technology in Distributed System Architecture. International Journal of Multimedia Computing (2025), Vol. 6, Issue 1: 152-159.
[24] Jin Li. The Impact of Distributed Data Query Optimization on Large-Scale Data Processing. International Journal of Big Data Intelligent Technology (2025), Vol. 6, Issue 2: 139-146.
[25] Lingyun Lai. Financial Modeling and Industry Insights in Investment in New Materials Industry. International Journal of Business Management and Economics and Trade (2025), Vol. 6, Issue 1: 179-187.
[26] Xia Hua. User Stickiness and Monetization Strategies in the Release of Global Game Projects. International Journal of Business Management and Economics and Trade (2025), Vol. 6, Issue 1: 188-195.
[27] Linwei Wu. The Application of Quantitative Methods in Project Management and Actual Effect Analysis. International Journal of Business Management and Economics and Trade (2025), Vol. 6, Issue 1: 204-212.
[28] Junchun Ding. Cross-Functional Team Collaboration and Project Management in the Automotive Industry. International Journal of Social Sciences and Economic Management (2025), Vol. 6, Issue 2: 162-170.
[29] Huijie Pan. Design of Data-Driven Social Network Platforms and Optimization of Big Data Analysis. Machine Learning Theory and Practice (2025), Vol. 5, Issue 1: 133-140.
[30] Jingzhi Yin. Research on Financial Time Series Prediction Model Based on Multi Attention Mechanism and Emotional Feature Fusion. Socio-Economic Statistics Research (2025), Vol. 6, Issue 2: 161-169