

Research on Multi Source Data Driven Financial Volatility Prediction Algorithm and Transfer Learning Framework Integrating BERT Attention and BiLSTM

Fang Li

Southeast University, Nanjing 211102, Jiangsu, China

Keywords: Multi source heterogeneous data fusion, BERT attention mechanism BiLSTM, Transfer learning framework, financial volatility prediction.

Abstract: Financial time series forecasting faces challenges such as market uncertainty, high-frequency data explosion, nonlinear relationships, and multi-source heterogeneous data fusion. Traditional statistical methods such as ARIMA and GARCH have limitations in complex environments, while machine learning/deep learning can capture nonlinear relationships, but face issues of insufficient sensitivity and interpretability in feature engineering. This study integrates multi-source heterogeneous data (historical fluctuations, market sentiment, macroeconomic indicators, etc.), combines BERT attention mechanism to capture text semantic correlations, uses BiLSTM bidirectional modeling capability to extract temporal dependent features, and constructs a transfer learning framework to achieve cross market/cyclical knowledge transfer; Introducing Hybrid Multiscale Decomposition (HMSD) to extract intrinsic features of time series, designing Interval Anomaly Detection and Large Scale Prediction Strategy (LSF), and quantifying prediction uncertainty through interval reliability analysis. Multi-source data fusion and BERT mechanism significantly enhance information richness and model expression ability; The combination of multi-scale decomposition and BiLSTM enhances the ability of nonlinear dynamic modeling; Interval anomaly detection and LSF strategy solve uncertainty quantification problems, improving the robustness and interpretability of results; Transfer learning has validated its adaptability in cross market scenarios such as carbon trading and crude oil prices. Future improvements are needed in concept drift, real-time decomposition techniques, parameter optimization, and social network analysis; Explore dynamic update mechanisms to respond to market changes, introduce real-time decomposition to enhance high-frequency data response capabilities, optimize BERT BiLSTM architecture with automatic parameter tuning, integrate social network analysis to enhance multi market collaborative prediction capabilities, improve the robustness and generalization ability of the model in complex financial environments, and provide stronger methodological support for accurate volatility prediction and risk management.

1. Introduction

The research on multi-source data-driven financial volatility prediction algorithm and transfer learning framework integrating BERT attention and BiLSTM is based on the core position of the financial industry as a modern economic pillar. It explores challenges such as market uncertainty,

high-frequency data explosion, nonlinear relationships, and multi-source heterogeneous data fusion in financial time series prediction. The research background emphasizes that financial time series contain multidimensional data such as prices, trading volumes, and interest rates of assets such as stocks, bonds, and foreign exchange. Their non-stationary, noisy, and chaotic characteristics make traditional statistical methods (such as ARIMA and GARCH) limited in dealing with complex market environments. Although machine learning and deep learning techniques can capture nonlinear relationships, they face problems such as feature engineering sensitivity and insufficient interpretability of black box models. This study integrates multi-source heterogeneous data such as historical data, network search index, social media sentiment, and voice comments, and combines BERT attention mechanism to capture text semantic associations. BiLSTM bidirectional temporal modeling capability is used to extract time-dependent features, and a transfer learning framework is constructed to achieve cross market and cross cycle knowledge transfer, thereby improving the accuracy and robustness of volatility prediction. The research significance lies in the dual dimensions of theoretical innovation and method application: theoretically, integrating multi-source data-driven and mixed multi-scale decomposition techniques, breaking through the limitations of a single data source, and enhancing the adaptability of the model to interval value and functional data; In practice, high-precision prediction tools are provided for financial markets such as carbon trading and crude oil prices to assist investors in risk management, policy formulation, and financial innovation. The current research status at home and abroad shows that existing achievements mostly focus on a single external factor (such as search index or text sentiment) or specific data types (point value, interval value). However, this framework effectively solves the problems of decomposition method selection disputes, outlier interference, and prediction result uncertainty through multi-source information fusion, mixed decomposition integration, interval outlier detection, and sub modal combination prediction. Ultimately, it forms a new paradigm for predicting financial volatility with strong interpretability and high stability, providing key technical support for preventing financial crises and improving economic resilience.

2. Correlation theory

2.1. Theoretical basis and preprocessing methods of interval and functional time series

Interval time series [1] are composed of interval numbers, which can be defined as the combination of the upper and lower boundaries or midpoints of time and radius, expressed mathematically as $[a, b]$ or (a, b) . Functional time series originated from functional data analysis, a concept proposed by Ramsay in 1982, referring to the observation records when the data generation process is a smooth function process. From the perspective of stochastic processes [2], functional data refers to the trajectories of random process samples defined on bounded intervals, such as the height time curves of each girl in the independently repeated "10 Girl Growth Dataset" [3], or the function series of high-frequency cumulative returns within the S&P 500 days divided by trading days in non independently repeated data [4].

Functional data preprocessing requires converting discrete observations into continuous function form, with core methods including basis function expansion, locally weighted smoothing, and rough penalty [5]. The basis function expansion method approximates latent functions through linear combinations of basis systems, commonly known as Fourier basis and B-spline basis [6]. B-spline bases have the characteristics of locality, continuity, and recursive definition, which are suitable for smoothing non periodic data (such as crude oil prices). Their definition depends on node partitioning and polynomial degree, and coefficient estimation usually uses the least squares method to minimize the sum of squared residuals [7]. This type of preprocessing transforms functional data from a discrete point set to a global functional perspective [8], providing a continuous and smooth

representation basis for subsequent analysis and supporting complex modeling requirements such as financial volatility prediction [9].

2.2. Deep Interest Network Model

Single prediction methods cover three major categories: statistics, machine learning, and deep learning, each with its own characteristics and applicable to different scenarios [10]. In statistical methods, ARIMA models effectively handle stationary time series by combining autoregression, differencing, and moving averages, particularly adept at capturing linear trends and periodicity; The GARCH model focuses on volatility modeling, using conditional heteroscedasticity to dynamically characterize the time-varying variance of financial asset prices, and is suitable for predicting volatility in stocks, exchange rates, and other markets; The Holt Winters exponential smoothing method is an extension of simple exponential smoothing, which updates the three elements of horizontal, trend, and seasonal terms, and performs well in time series with obvious trends or seasonality.

SVR maps to high-dimensional space through kernel functions, combined with an ε -insensitive loss function, to achieve nonlinear regression with limited samples, combining generalization ability and computational efficiency; MLP, as a feedforward neural network, can approximate any complex function through multi-layer hidden layer nonlinear transformations, making it suitable for high-dimensional and nonlinear time series prediction; Extreme Learning Machine (ELM) achieves fast training by randomly initializing hidden layer weights and only training output layer parameters, making it suitable for scenarios with high real-time requirements.

In deep learning methods, Long Short Term Memory (LSTM) networks introduce a triple gating mechanism consisting of input gates, forget gates, and output gates, effectively solving the problem of vanishing gradients in traditional RNNs and demonstrating outstanding performance in long-term dependency modeling; Gated Recurrent Unit (GRU) simplifies the LSTM structure by balancing information retention and updating through reset and update gates, thereby improving computational efficiency; Convolutional neural networks (CNNs) have shown advantages in time series pattern recognition due to their ability to extract local features from convolutional layers, especially for sequence data with spatial local correlations. Each of the three methods has its own focus. Statistical methods have mature theories and strong interpretability, while machine learning and deep learning excel in modeling nonlinear and high-dimensional data, forming a diverse tool library for time series prediction.

3. Research method

3.1. Overview of Data Processing Methods and Model Evaluation Indicators

The data preprocessing stage has established a multi-scale decomposition technology system, integrating mainstream methods such as wavelet analysis, empirical mode decomposition (EMD), variational mode decomposition (VMD), and singular spectrum analysis (SSA). Wavelet analysis captures signal features through the multi-scale decomposition of local basis functions in the time-frequency domain, but requires pre-set decomposition levels and is prone to generating pseudo harmonic components; EMD is adaptively decomposed into intrinsic mode function (IMF) and residual term based on local extremum features, which has the advantage of local adaptability, but there are endpoint effects and mode mixing phenomena; VMD adopts a non recursive optimization strategy to effectively suppress modal aliasing, especially suitable for nonlinear non-stationary time series data; SSA achieves the separation and extraction of trend, period, and noise components through a four step process of trajectory matrix embedding, singular value decomposition, grouping

reconstruction, and diagonal averaging. The quantification of text information needs to follow standardized processes, including data cleaning, word segmentation, stop word filtering, stem merging, feature extraction (such as TF-IDF algorithm), vectorized expression, normalization scaling, and feature matrix construction. Chinese natural language processing can use professional tools such as BosonNLP to complete the entire process operation. The performance evaluation of the model adopts a two-dimensional indicator system, and the point prediction dimensions are quantified using the sum of squared errors (SSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE);

The optimization effect of the model is quantified by comparing error indicators (such as SSE improvement rate). The greater the increase in indicator values, the more prominent the improvement of the model, which can intuitively reflect the improvement effect of prediction accuracy and stability.

3.2. Carbon price prediction method driven by multi-source fusion and mixed decomposition

Accurate carbon trading price prediction plays a crucial role in reducing carbon dioxide emissions and addressing global warming, but the carbon trading market faces dual challenges of external factors diversity and uncertainty. Traditional prediction methods are often inefficient and unstable due to disputes over a single information source or decomposition method selection. This chapter proposes a carbon trading price prediction framework based on Multi Source Information Fusion (MSIF) and Hybrid Multiscale Decomposition (HMSD) to systematically address the aforementioned issues. This framework consists of four steps: firstly, integrating historical carbon price data, influencing factors (such as energy prices), and unstructured data (such as search indices and social media sentiment) through multi-source information fusion (MSIF), constructing a dynamic prediction system, and using principal component analysis (PCA) and sentiment analysis to achieve data predictability; Secondly, using the Hybrid Multiscale Decomposition (HMSD) technique, combined with four signal processing methods including wavelet transform, Empirical Mode Decomposition, Singular Spectrum Analysis, and Variational Mode Decomposition, the intrinsic features of multi-source information are comprehensively extracted to avoid the uncertainty caused by the selection of a single decomposition method, decomposition frequency, and sequence; Next, a combined prediction strategy is implemented using statistical methods (such as Holt, ARIMA) and artificial intelligence methods (such as SVR, BPNN, LSTM) to address the linear and nonlinear characteristics of the sub modes, in order to stabilize the prediction results; Finally, the effectiveness of the framework is verified through multiple types of comparative experiments. Experiments have shown that Multi Source Information Fusion (MSIF) can significantly improve prediction performance compared to a single information source, and Hybrid Multiscale Decomposition (HMSD) outperforms single decomposition methods in terms of information extraction capability. The innovation of this method includes: for the first time, the fusion of multiple sources of information is applied to carbon trading price prediction, which compensates for the limitations of single source information; Establish efficient data processing technology HMSD to reduce data complexity and improve prediction stability; The combination prediction model combining statistics and intelligent strategies effectively handles linear and nonlinear features, reduces the uncertainty of results caused by model selection, and provides an advanced and stable prediction tool for the carbon trading market.

3.3. Carbon Price Prediction Model and Verification Based on Multi source Fusion and Hybrid Decomposition

The carbon price prediction model constructed in this study integrates historical prices, energy

prices (coal, oil, gas, electricity prices), and unstructured data (search index, social media sentiment) through multi-source information fusion (MSIF). Data feature analysis shows that carbon prices exhibit significant non-stationary (ADF test $P=0.797$) and nonlinear (BDS test z statistic increases to 91.2752) characteristics, with a standard deviation of 4.0, skewness of 1.39, and kurtosis of 1.5 indicating severe fluctuations and abnormal distribution. Coal and crude oil prices have a significant negative impact, and unstructured data forms a high contribution emotional sequence after PCA dimensionality reduction (emotional score of about 0.55). Experimental verification shows that HMSD's decomposition ability is superior to any single technology, and the combined forecasting (CFM) strategy is significantly superior to the single model - such as HMSD-CFM-H's MAPE (0.0070), RMSE (0.3415) and other indicators are optimal, which is 85.85% higher than the average error of the single model, and 95.65% higher than the existing model. Although the model has a high computational cost (average running time of 1713.15 seconds), it effectively improves prediction accuracy and stability by reducing uncertainty through multi-source information, improving feature extraction efficiency through HMSD, and combining complex sub modal features for prediction processing. The contribution of this chapter lies in pioneering a collaborative framework for multi-source information fusion and hybrid decomposition, addressing the limitations of single source information, and proposing HMSD technology to avoid disputes in the selection of decomposition methods; Limitations include the need to explore the applicability of interval value data, optimize the sensitivity of abnormal data, and further improve computational efficiency. Future research will expand to interval value prediction, anomaly detection, and cross domain application verification.

4. Results and discussion

4.1. Dual stage BEMD Error Correction Carbon Price Interval Prediction Model

This paragraph constructs an interval value carbon price combination prediction model based on ICEEMDAN and I-ksigma to address challenges such as complex data features, outlier interference, and uncertainty in prediction method selection. The model first constructs the ICEEMDAN decomposition framework by combining interval variables (Minmax, CRM, L+2R, U-2R) to decompose the interval value sequence into high-frequency, mid frequency, low-frequency, and trend sub modes. The sub modes are reconstructed using permutation entropy to improve stability; Secondly, the I-ksigma criterion is introduced to identify outliers in sub modalities, combined with KNN interpolation smoothing to reduce the interference of outliers on prediction; Finally, five types of models including IMLP, MSVR, LSTM, GRU, and CNN are used for combined prediction of sub modalities, and the final result is obtained by summing them up. Experiments have shown that different combinations of interval variables significantly affect the decomposition and prediction results - the CRM combination factor has small modal deviation, effective outlier identification, and the best prediction performance, while the Minmax combination performs the worst due to large low-frequency deviation and multiple outliers. The advantages of the model are reflected in three aspects: ICEEMDAN overcomes the modal mixing and endpoint effects of traditional BEMD, and enhances feature extraction capabilities; I-ksigma achieves effective detection and smoothing of interval outliers; Multi model combination avoids the uncertainty of single method selection and enhances robustness. Research contributions include proposing interval adaptive noise decomposition method, constructing interval anomaly detection process, and verifying the effectiveness of multi model combination in interval prediction, providing new technical path and theoretical support for interval value carbon price prediction.

4.2. Model experiment

This chapter validates the effectiveness of the interval value carbon trading price prediction framework based on ICEEMDAN and I-ksigma through five experimental systems, while retaining key error matrix data to support the conclusions. Experiment 1 compares benchmark models such as IMLP, MSVR, LSTM, GRU, CNN, etc. The results are shown in Table 1.

Table 1. Comparison of Error Indicators for Interval Carbon Price Prediction

model name	IMAPE	IRMSE	IARV	IU
IMLP	0.0695	2.3434	1.3120	1.9716
MSVR	0.0705	2.2969	1.2944	1.9584
LSTM	0.0672	2.4023	1.2968	1.9602
GRU	0.0740	2.4772	1.3723	2.0165
CNN	0.0681	2.2988	1.1846	1.8735
Minmax-proposed model	0.0357	1.3443	0.5007	0.7260
CRM-proposed model	0.0231	0.9372	0.2670	0.5302
L+2R-proposed model	0.0293	1.1256	0.3541	0.6105
U-2R-proposed model	0.0374	1.2848	0.4563	0.6931

In the benchmark model, CNN performed the best in the four indicators of IMAPE, IRMSE, IARV, and IU (IMAPE=0.0681, IRMSE=2.2988, IARV=1.1846, IU=1.8735). However, the proposed Minmax proposed, CRM proposed, L+2R proposed, and U-2R proposed models were significantly better than the baseline models in all indicators (such as the CRM proposed model with IMAPE=0.0231, IRMSE=0.9372, IARV=0.2670, IU=0.5302), confirming that a single prediction model is difficult to meet the requirements. The combined framework achieved high-precision prediction through the synergistic effect of ICEEMDAN decomposition, I-ksigma anomaly detection, and sub modal combination prediction. Experiment 2 focused on the comparison of interval decomposition techniques and found that BEMD suffers from modal mixing (mid frequency and low frequency in the interval), endpoint effects, and high-frequency outliers. The proposed ICEEMDAN effectively overcomes these problems, and although BEMD-I-ksigma-CFM is superior to BEMD-CFM without anomaly detection, it still lags behind the proposed model, indicating that ICEEMDAN is more effective in feature extraction and outlier detection. Experiment 3 verifies the necessity of I-ksigma. The ICEEMDAN-CFM series models without anomaly detection (such as CRM-ICEEMDAN-CFM with IMAPE=0.0282) have weaker predictive performance than the corresponding proposed models, confirming that I-ksigma improves accuracy by reducing the impact of outliers. However, it should be noted that the combination of interval variables (such as CRM) has a significant impact on the prediction results. Experiment 4 compared the combination prediction with a single prediction model, and found that the proposed combination framework (such as CRM proposed model IMAPE=0.0231) was significantly better than the single sub modal prediction models based on IMLP, MSVR, LSTM, GRU, and CNN, proving that combination prediction can overcome method selection uncertainty and enhance stability. Compared with existing carbon price prediction models such as CEEMDAN/VMD-LSTM's IMAPE=0.0613 and interval prediction models such as BEMD-LSTM's IMAPE=0.0679 in Experiment 5, the proposed model has the advantage in the minimum error metric (CRM proposed model IMAPE=0.0231, Figure 1), attributed to ICEEMDAN's information extraction ability, I-ksigma's outlier robustness, and interval variable combination strategy.

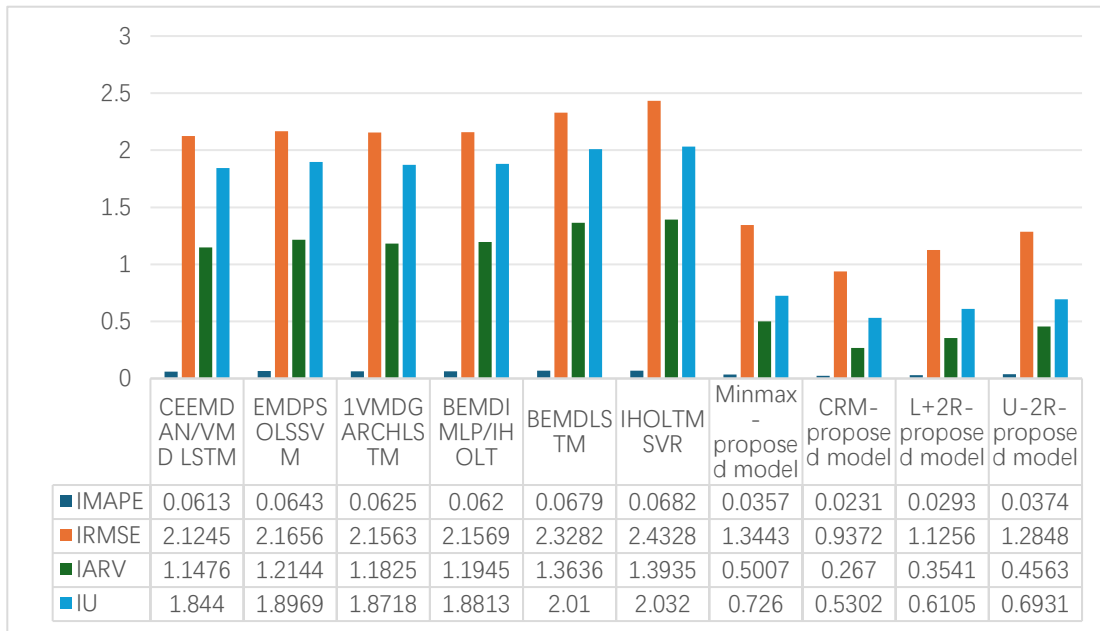


Figure 1. Comparison of Error Indicators for Interval Carbon Price Prediction

Statistical analysis was conducted using Wilcoxon signed rank test, with the CRM proposed model as the benchmark. The p-values of all compared models were less than 0.05, confirming that the prediction error was significantly smaller; However, there were no significant differences between the Minmax based, L+2R based, and U-2R based models, indicating that the combination of interval variables has a significant impact on the prediction results. In summary, the framework proposed in this chapter achieves significant improvement in interval carbon price prediction performance through ICEEMDAN decomposition, I-ksigma anomaly detection, and sub modal combination prediction, combined with different interval variable combination strategies (such as the best CRM effect), and provides new data processing techniques, which have theoretical innovation and application value.

4.3. Effect analysis

This chapter proposes a large-scale interval value carbon trading price prediction framework (IMSIF-MEMD-COF-LSF) based on interval multi-source information fusion (IMSIF) and connectivity outlier factor (COF), which achieves high-precision prediction and reliability analysis through a four step process. The first step is to construct an interval multi-source information fusion system, integrating historical carbon prices, supply and demand influencing factors (such as international carbon emission futures prices, crude oil futures prices, etc.), and unstructured data (Baidu search index, social media sentiment), and using principal component analysis to reduce the dimensionality of keyword data; The second step is to perform data preprocessing, using multiple empirical mode decomposition (MEMD) to decompose the upper and lower bounds of the fused data and extract stable sub modes. Combined with connectivity based outlier factors (COF) to detect multivariate outliers, outlier points are corrected through mean interpolation to reduce data complexity; The third step is to implement large-scale combination prediction, integrating statistical models (ARIMAX, MLR), machine learning (SVR, MLP), and deep learning (LSTM, CNN) to process sub modal linear and nonlinear features, and optimizing the results through combination strategies such as simple averaging, optimal weights, and nonlinear ensemble; The fourth step is to establish an interval reliability prediction method, which calculates the prediction interval based on

85%, 90%, 95%, and 99% confidence levels. Interval reliability analysis shows that the model achieves high coverage probability, narrow interval width, and low coverage error at different confidence levels, providing probabilistic decision support for carbon market risk management. Although the model still has limitations (such as interval time series only involving endpoint information, multi-source information fusion needs to solve the problem of synchronizing data with different frequencies and dimensions), by turning to functional time series prediction and optimizing frequency synchronization mechanisms, the applicability and accuracy of prediction can be further improved.

5. Conclusion

This article focuses on a systematic study of the multi-source data-driven financial volatility prediction algorithm and transfer learning framework that integrates BERT attention and BiLSTM. The core conclusions are as follows: through the fusion of multi-source heterogeneous data (covering historical volatility, market sentiment, macroeconomic indicators, etc.), combined with BERT attention mechanism to capture complex correlations between data, the information richness and model expression ability of financial volatility prediction are effectively improved; Introducing the Hybrid Multiscale Decomposition (HMSD) technique to extract intrinsic features of time series, combined with the BiLSTM network to handle sequence dependencies, enhances the model's ability to model nonlinear dynamics; The designed interval outlier detection method and large-scale prediction strategy (LSF), combined with interval reliability analysis, solve the uncertainty quantification problem in volatility prediction, making the prediction results more robust and interpretable; Based on the transfer learning framework, the model achieved knowledge transfer between different financial scenarios such as carbon trading and crude oil prices, verifying the adaptability of cross market prediction. However, there is still room for improvement in research: in the future, attention should be paid to the problem of concept drift and dynamic update mechanisms should be explored to cope with changes in market conditions; The introduction of real-time decomposition technology can enhance the model's responsiveness to high-frequency financial data; Parameter optimization needs to be combined with automatic parameter tuning technology to explore the optimal configuration of BERT BiLSTM architecture; Social network analysis can further integrate the interactive information between markets (such as policy linkage and information dissemination), enhancing the collaborative predictive ability of multi market volatility. These directions will promote the robustness and generalization ability of models in complex financial environments, providing stronger methodological support for accurate prediction and risk management of financial volatility.

References

- [1] Huang, J. (2025). *Research on Resource Prediction and Load Balancing Strategies Based on Big Data in Cloud Computing Platform*. *Artificial Intelligence and Digital Technology*, 2(1), 49-55.
- [2] Xu, H. (2025). *Supply Chain Digital Transformation and Standardized Processes Enhance Operational Efficiency*. *Journal of Computer, Signal, and System Research*, 2(5), 101-107.
- [3] Liu, Y. (2025). *The Importance of Cross-Departmental Collaboration Driven by Technology in the Compliance of Financial Institutions*. *Economics and Management Innovation*, 2(5), 15-21
- [4] Zhou, Y. (2025). *Improvement of Advertising Data Processing Efficiency Through Anomaly Detection and Recovery Mechanism*. *Journal of Media, Journalism & Communication Studies*, 1(1), 80-86.

- [5] Yang D, Liu X. Collaborative Algorithm for User Trust and Data Security Based on Blockchain and Machine Learning[J]. *Procedia Computer Science*, 2025, 262: 757-765.
- [6] Zhang, Xuanrui. "Automobile Finance Credit Fraud Risk Early Warning System based on Louvain Algorithm and XGBoost Model." In *2025 3rd International Conference on Data Science and Information System (ICDSIS)*, pp. 1-7. IEEE, 2025.
- [7] Xu, H. (2025). Optimization of Packaging Procurement and Supplier Strategy in Global Supply Chain. *European Journal of Business, Economics & Management*, 1(3), 111-117.
- [8] Jing X. Real-Time Risk Assessment and Market Response Mechanism Driven by Financial Technology[J]. *Economics and Management Innovation*, 2025, 2(3): 14-20.
- [9] Liu X. The Role of Generative AI in the Evolution of Digital Advertising Products[J]. *Journal of Media, Journalism & Communication Studies*, 2025, 1(1): 48-55.
- [10] Han, Wenxi. "The Practice and Strategy of Capital Structure Optimization under the Background of the Financial Crisis." *European Journal of Business, Economics & Management* 1, no. 2 (2025): 8-14.
- [11] Chang, Chen-Wei. "AI-Driven Privacy Audit Automation and Data Provenance Tracking in Large-Scale Systems." (2025).
- [12] Zhang K. Research on the Application of Homomorphic Encryption-Based Machine Learning Privacy Protection Technology in Precision Marketing[C]//2025 3rd International Conference on Data Science and Network Security (ICDSNS). IEEE, 2025: 1-6.
- [13] Truong T. The Research on the Application of Blockchain Technology in the Security of Digital Healthcare Data [J]. *International Journal of Health and Pharmaceutical Medicine*, 2025, 5(1): 32-42.
- [14] Gao Y. Research on Risk Identification in Legal Due Diligence and Response Strategies in Cross border Mergers and Acquisitions Transactions [J]. *Socio-Economic Statistics Research*, 2025, 6(2): 71-78.
- [15] Li W. Building a Credit Risk Data Management and Analysis System for Financial Markets Based on Blockchain Data Storage and Encryption Technology[C]//2025 3rd International Conference on Data Science and Network Security (ICDSNS). IEEE, 2025: 1-7.
- [16] Zhou Y. Cost Control and Stability Improvement in Enterprise Level Infrastructure Optimization [J]. *European Journal of Business, Economics & Management*, 2025, 1(4): 70-76.
- [17] Li, W. (2025). Research on Optimization of M&A Financial Due Diligence Process Based on Data Analysis. *Journal of Computer, Signal, and System Research*, 2(5), 115-121.
- [18] Hu, Q. (2025). Implementation and Management of a Cross-Border Tax System Oriented Towards Global Tax Administration Informatization. *Economics and Management Innovation*, 2(4), 94-101.
- [19] Wang, C. (2025). Exploration of Optimization Paths Based on Data Modeling in Financial Investment Decision-Making. *European Journal of Business, Economics & Management*, 1(3), 17-23.
- [20] Cai, Y. (2025). Research on Positioning Technology of Smart Home Devices Based on Internet of Things. *European Journal of AI, Computing & Informatics*, 1(2), 80-86.
- [21] Wei, X. (2025). Practical Application of Data Analysis Technology in Startup Company Investment Evaluation. *Economics and Management Innovation*, 2(4), 33-38.
- [22] Huang, J. (2025). Promoting Cross-field E-Commerce Development by Combining Educational Background and Technology. *Economics and Management Innovation*, 2(4), 26-32.
- [23] Jing, X. (2025). Research on the Application of Machine Learning in the Pricing of Cash Deposit Products. *European Journal of Business, Economics & Management*, 1(2), 150-157.
- [24] Yang D, Liu X. Collaborative Algorithm for User Trust and Data Security Based on Blockchain and Machine Learning[J]. *Procedia Computer Science*, 2025, 262: 757-765.

- [25] Tang X, Wu X, Bao W. *Intelligent Prediction-Inventory-Scheduling Closed-Loop Nearshore Supply Chain Decision System*[J]. *Advances in Management and Intelligent Technologies*, 2025, 1(4).
- [26] Wu X, Bao W. *Research on the Design of a Blockchain Logistics Information Platform Based on Reputation Proof Consensus Algorithm*[J]. *Procedia Computer Science*, 2025, 262: 973-981.