Research And Analysis of Neural Networks for Financial Time Series Prediction

Feiyi Feng *

School of Computer Science, City University of Macau, Macau, China

* Corresponding Author Email: D24090150562@cityu.edu.mo

Abstract. Financial time series prediction is crucial for portfolio optimization, risk management, and financial policy-making. However, traditional technical analysis and statistical models, limited by linear assumptions and data distribution, struggle to handle the market's nonlinear, high-volatility traits. Neural networks, with their nonlinear mapping and autonomous learning, offer a key solution. This paper reviews the field's research over the past five years, focusing on four mainstream technologies: Recurrent Neural Network, functional denoising autoencoder, Long Short-Term Memory Neural Network, and hypergraph neural networks. It analyzes each model's principles, features, pros and cons, and verifies their effectiveness using indicators like IC, Sharpe Ratio, and RMSE on datasets from A-shares, U.S. stocks, and cryptocurrencies. Models integrating spatiotemporal features (e.g., hypergraph neural networks) perform exceptionally well. The paper also notes current challenges, such as single data usage and weak cross-market generalization. Future optimization should involve multi-source data fusion and cross-market transfer learning, providing references for financial practice and advancing neural networks' practical application.

Keywords: Neural Networks; Financial Time Series; Long Short-Term Memory; Recurrent Neural Network.

1. Introduction

Financial time series are an intuitive reflection of macroeconomics and market behavior [1]. ACCURATE price forecasting in financial markets is crucial for enabling informed decision-making, managing risks effectively, and supporting strategic planning for investors, businesses, financial institutions, and governments [2]. In recent years, artificial neural networks have been used independently or as an auxiliary tool for predicting time series as a nonlinear method to simulate the nervous system [3]. In traditional forecasting methods, technical analysis relies on empirical indicators, while statistical models are limited by linear assumptions and data distribution constraints [4]. However, the nonlinear correlations, volatile characteristics, and external factor interferences in the market lead to insufficient prediction accuracy of traditional methods. Since the 1990s, neural networks, which possess nonlinear mapping and autonomous learning capabilities, have become the key technology to solve this problem.

Theoretically, sorting out the technological development and model innovation of neural networks in financial forecasting can improve the interdisciplinary theoretical system and provide references for subsequent research. In practice, clarifying the applicable scenarios and performance boundaries of different models can provide financial institutions with accurate forecasting tools and reduce investment risks; at the same time, exploring bottlenecks such as model interpretability can provide a basis for the formulation of regulatory norms.

In recent years, many researchers have utilized neural network to conduct relevant studies in the financial field:

Kady Sako et al. used Recurrent Neural Networks and their variants—Long Short-Term Memory networks and Gated Recurrent Units—to forecast the closing prices of eight stock market indices and the closing prices of six U.S. dollar-pegged currency exchange rates [5]. He obtained forecast results for these two targets in terms of out-of-sample forecasting for univariate currency exchange rates and out-of-sample forecasting for multivariate stock market indices.

Ma Chao et al. constructed a functional-based deep denoising autoencoder neural network to improve the generalization performance of neural networks in time-series forecasting tasks [6]. Meanwhile, by focusing on the forecasting of trends and directions of price sequences, they avoided the impact of original sequences and made up for the inherent shortcomings of traditional neural networks in forecasting the direction of price ups and downs.

Ouyang Hongbing et al. proposed a method of combining wavelet analysis with Long Short-Term Memory neural networks that incorporate time-series dependency features to construct a financial time-series data forecasting model [7]. This is aimed at overcoming the defect of existing models that they cannot reflect the complex characteristics of financial time-series data—such as non-stationarity, non-linearity, and serial correlation—as well as the non-linear interaction relationships between data.

Liu Zhibin proposed a forecasting method based on Transformer neural networks and hypergraph neural networks to construct static and dynamic hypergraphs, thereby conducting ranking forecasting on the price increase and decrease ranges of stocks [8]. Experiments on the A-share market dataset and U.S. stock market dataset verified the forecasting performance and profitability of the proposed model.

This paper aims to systematically integrate the research achievements of neural networks in the field of financial time series forecasting. By sorting out the literature, cutting-edge papers, and practical cases of the past five years, it constructs a framework covering technological development, model innovation, and scenario application. Meanwhile, it sorts out the research context and points out the innovation directions for the academic community; for the financial industry, it provides scientific and practical model selection and optimization solutions.

2. Overview of Mainstream Technologies

2.1. Recurrent Neural Networks and its Variants

2.1.1 Basic principles of RNN

Kady Sako et al. have utilized RNNs and their variants—LSTM networks and GRUs. The core principle of these models centers on "capturing temporal information," and they establish connections between historical data and future predictions through specialized network structures.

The simple RNN takes "recurrent connections" as its core. It uses the hidden state of the previous time step as the input for the current time step, and processes nonlinear features in combination with an activation function. This forms a dependency chain of "historical information - current output," enabling preliminary temporal modeling of sequence data.

LSTM networks, on the other hand, optimize information transmission through a triple gating mechanism consisting of an "input gate," a "forget gate," and an "output gate". By virtue of the continuous transmission of the cell state, LSTM solves the gradient vanishing problem of simple RNNs [9].

GRUs simplify the LSTM structure by retaining only the "reset gate" and the "update gate." They transmit core information through a single memory unit, which reduces parameters while preserving the ability to capture temporal information.

2.1.2 Method Characteristics, Advantages and Disadvantages

It is characterized by an extremely simple structure, consisting only of an input layer, a recurrent layer, and an output layer, with a small number of parameters and intuitive computational logic. Its advantages lie in low computational cost and fast training speed, making it suitable for preliminary temporal modeling or scenarios with limited data volume. However, its disadvantages are significant: the gradient decays or grows exponentially as the number of time steps increases, preventing it from capturing long-term dependencies. When dealing with medium- and long-term fluctuations in financial markets, its prediction accuracy drops sharply. Additionally, it is sensitive to noise and easily disturbed by outliers, so it can only be used as a basic comparison model.

It is characterized by a complex gating mechanism and a large number of parameters, which require fine tuning. Its advantages lie in strong long-sequence dependency capture capability: through the gating mechanism, it selectively retains key historical information and can handle complex cross-time-step correlations, making it theoretically suitable for long-term trend prediction in financial markets. However, its disadvantages are obvious: the large number of parameters leads to high computational complexity and long training time, and high requirements for hardware resources. Moreover, this study only uses short-time-step data of 5 days, so its long-time-sequence advantage is not fully exerted; instead, "overfitting risk" occurs due to parameter redundancy, resulting in suboptimal prediction performance

It is characterized by a balanced structure, simplifying LSTM's gating design with approximately 30% fewer parameters than LSTM while retaining core memory functions. Its advantages are prominent: high computational efficiency, able to balance accuracy and speed in medium-to-short-term time series prediction, and strong robustness to noise. In the face of small fluctuations in financial data, its prediction stability is better than the previous two models. However, its disadvantage is that its long-period capture capability is slightly inferior to LSTM. If the time steps are extended, key long-term dependencies may be lost due to excessive information compression, leading to decreased accuracy.

In summary, the three neural network technologies each have their suitable scenarios: the simple RNN is suitable for lightweight preliminary modeling, LSTM is suitable for long time series requirements, and GRU achieves the optimal balance between efficiency and accuracy in medium-to-short-term predictions.

2.2. Functional Denoising Autoencoder Network

2.2.1 Basic principles of FdAE

Ma Chao et al. have utilized FdAE, which integrates functional analysis with deep learning. Its core lies in converting financial time series from "discrete data points" into "continuous functions" for processing, thereby addressing the issue that traditional deep learning underutilizes the time dimension.

Breaking free from the limitations of traditional numerical networks, the input layer treats time series as continuous real-valued function mapping relationships. It replaces summation operators with linear integral operators and combines time-varying functions to capture the temporal continuity characteristics of data. No discretization process is required, thus preserving the temporal correlation information of the original sequence.

The hidden layer extracts high-level features through function composition operations and introduces sparsity constraints to reduce redundant features. Meanwhile, it incorporates a denoising mechanism: by randomly disrupting part of the input information, the network is forced to learn the inherent laws of data rather than noise, enhancing the robustness of features.

The decoding layer retains linear summation operators. It expands inputs and weights through basis functions, converting continuous functions into computable parameters. Based on Parseval's theorem, an error function is defined, and weights and thresholds are optimized using algorithms such as gradient descent. Finally, through stack integration and supervised fine-tuning, a complete prediction model is formed, outputting prediction results for the PI index.

2.2.2 Method Characteristics

Based on functional analysis, it directly processes continuous time series, avoiding information loss caused by discretization and better aligning with the inherent temporal continuity of financial data.

By combining input disruption for denoising and sparsity constraints, it provides dual guarantees for the network to learn core laws, reducing the impact of high noise in financial data.

2.2.3 Advantages and Disadvantages

The Functional Denoising Autoencoder Network features strong generalization ability. It breaks through the overfitting bottleneck in traditional networks. Experiments show that the R^2 of the test set only decreases by 0.11 compared with the training set, which is far better than the decrease of more than 0.8 in the NARX network. It can still stably predict even when facing changes in the distribution of financial data. Additionally, combined with the PI index that focuses on trend and turning point information, it effectively avoids noise in original prices, solves the problem of "accurate values but poor direction" in traditional networks, and is more suitable for actual trading needs.

Meanwhile, this technology can perform automatic feature extraction without the need for manual design of numerous technical indicators. It automatically mines deep temporal features through the network, reducing biases caused by human intervention.

The Functional Denoising Autoencoder Network requires designing the network structure based on functional analysis theory, involving basis function selection, integral operator calculation, etc. The difficulty of model construction and debugging is far greater than that of traditional neural networks. The processing of continuous functions, stack training, and fine-tuning processes also require substantial computing power support, with high hardware requirements, resulting in low cost-effectiveness in small-scale data scenarios. Moreover, the processes of functional transformation and deep feature extraction are quite abstract, making it difficult to intuitively explain the logic behind prediction results, which does not meet some requirements for model interpretability in the financial field.

2.3. Long Short-Term Memory Neural Network

2.3.1 Basic principles of LSTM

Ouyang Hongbing et al. constructed a combined technical model of "wavelet analysis preprocessing + LSTM neural network prediction" based on the needs of financial time series forecasting, while clarifying the basic architecture of deep neural networks and the core settings of comparative models.

LSTM is an improved version of recurrent neural networks, specifically designed for processing time series data. Its core innovation is the introduction of a "memory cell" structure, which realizes the selective processing of information through three control units: input gate, forget gate, and output gate.

The forget gate determines which information to discard from the cell state, outputting values between 0 and 1 through a sigmoid function. The input gate is responsible for updating the cell state.This gating mechanism enables LSTM to effectively capture long-term dependencies in time series, solving the gradient vanishing or exploding problem that occurs in traditional RNNs when processing long sequences. It is particularly suitable for financial time series, which contain long-term memory features.

2.3.2 Advantages and Disadvantages

It possesses the ability to capture long-range dependencies. Through its gating mechanism and the circulation of cell states, it can retain and utilize important long-term information in time series, which is crucial for financial data forecasting. Via multi-layer nonlinear transformations, it can effectively characterize the complex nonlinear relationships in financial markets, outperforming linear models.

When combined with wavelet analysis for preprocessing, it can better filter high-frequency noise, exhibit strong anti-interference capability and enable it to focus on learning core trends. Meanwhile, this model does not require manual feature design; it can automatically extract effective features from raw data, reducing reliance on professional knowledge in the financial field.

Compared with traditional models, the LSTM model requires more computing resources and longer training time, and parameter tuning is more difficult. Additionally, as a black-box model, LSTM is hard to explain the formation logic of prediction results, which limits its application in the financial field where risk controllability is emphasized. Furthermore, it can only exert its advantages when there is a large amount of high-quality time series data, and performs poorly in scenarios with scarce data.

2.4. Hypergraph Neural Network

2.4.1 Basic principles of HGNN

As an emerging artificial intelligence algorithm, Graph Neural Network has attracted a lot of attention in the financial field in recent years [10]. The HGNN is an extension of Graph Neural Networks. Its core lies in modeling the collective relationships between nodes through "hyperedges" (edges that can connect multiple nodes), making it suitable for multi-asset linkage scenarios in financial markets. Its principle revolves around the definition of hypergraph structure and spatial-domain hypergraph convolution.

Based on hypergraphs, Liu Zhibin modeled asset correlations through a two-stage message passing process of "node-hyperedge-node".

2.4.2 Method Characteristics

This experiment breaks through the limitation of traditional GNNs that can only handle pairwise relationships. Hyperedges can connect multiple nodes, enabling direct modeling of collective multi-asset relationships in financial markets, such as "inter-industry linkage" and "sector synergy". Meanwhile, static hypergraphs and dynamic hypergraphs can be constructed, catering to both long-term stable correlations and short-term dynamic correlations of assets.

By leveraging hyperedge-vertex attention to generate continuous-value incidence matrices, it replaces the traditional 0/1 discrete matrices. This allows the representation of dynamic membership degrees between assets and hyperedges, aligning with the complexity of relationships in financial markets.

2.4.3 Advantages and Disadvantages

The I-Transformer improves the quality of temporal features, while hypergraphs capture multi-dimensional correlations—these together enhance stock prediction accuracy, with indicators such as IC and Rank IC all outperforming traditional models. Among them, the asset portfolio management model also has advantages including a high Sharpe ratio and low maximum drawdown; on the cryptocurrency dataset, its SR reaches 1.0322 and MDD is only 0.0762. Additionally, the model features strong generalization: it is effective across A-share, U.S. stock, and cryptocurrency datasets, making it adaptable to different financial markets.

However, the model still has shortcomings such as high complexity, sensitivity to hyperparameters, and strong data dependence. It has a large number of parameters, requiring long training and testing times, and also has high hardware requirements. Moreover, hyperparameters such as the 1D convolution kernel size and k-nearest neighbor parameters need to be optimized via grid search; improper parameter settings can easily affect performance. At the same time, the model's application also requires a large amount of high-quality financial data. When dealing with small datasets or data with high noise, the effectiveness of hypergraph construction and feature learning tends to decline.

3. Experimental Results

3.1. Data set

3.1.1 Experiments Related to RNN and Its Variants

Daily data of 8 stock market indices, including NYSE Composite and NASDAQ Composite from the Americas, Euronext 100 and Deutsche Boerse DAX index from Europe, SSE Composite index and Nikkei 225 from Asia, as well as Global X MSCI Nigeria ETF and FTSE/JSE Africa index from Africa. The data spans from January 2, 2008 to May 28, 2021, with a small number of missing values in some indices. Each dataset contains 6 features: High, Low, Close, Volume, Adjust Close, and Open, with the research focusing on closing price prediction.

3.1.2 Experiments Related to Functional Deep Neural Networks

Financial futures data: The front-month continuous contract of CSI 300 stock index futures (IFL8), with two-time spans: one from January 30, 2013 to August 31, 2015 (this period was selected because subsequent stock index futures restriction measures may have caused abnormal price characteristics), and the other from January 5, 2015 to March 10, 2017.

3.1.3 Experiments Related to LSTM Neural Networks

Experiment selected the daily closing price data of the Dow Jones Industrial Average (DJIA), with a time span from January 4, 2012 to June 31, 2017, and the data was sourced from Wind Information. Fourteen indicators were selected as feature input vectors from dimensions such as trading volume and trading price.

3.1.4 Experiments Related to Hypergraph Neural Networks

A-share datasets: Including CSI300 and CSI100, with a time span from August 1, 2013 to August 1, 2020. The Alpha360 features from the open-source quantitative investment platform Qlib are adopted, covering 6 indicators. Daily data for each stock is constructed into 360-dimensional features.

U.S. stock datasets: NASDAQ (1,026 stocks) and NYSE (1,737 stocks) datasets, with a time span from January 2, 2013 to August 12, 2017, including information on stock industry affiliation.

3.2. Evaluation Metrics

3.2.1 Rank Performance Metrics

Information Coefficient (IC): It measures the linear correlation between the model's predicted values and the actual values. The calculation formula is as follows: $IC\ (y', \hat{y'}) = corr(y', \hat{y'})$, where corr refers to the correlation coefficient calculation function. The value range of IC is between -1 and 1, and a larger value indicates stronger predictive ability of the model.

Rank IC: First, sort the daily factor values and target values in descending order, then calculate the correlation coefficient between the two sorted values. The calculation formula is as follows: $RankIC\ (y', \hat{y'}) = corr(rank_{y'}, rank_{\hat{y'}})$ .Similarly, a larger value indicates a better prediction effect.

Precision@N: Evaluates the precision of the top N predicted results of a model.

3.2.2 Profitability Indicators

Average Daily Investment Return (AVR): Measures the profitability of an investment project or asset. The calculation formula is as follows:

$AVR = \frac{1}{L}(\frac{1}{|F^{t-1}|}\sum_{i\in F^{t-1}}\frac{p_i^t - p_i^{t-1}}{p_i^{t-1}})$, where L represents the number of trading days, Ft−1 denotes the set of stocks in the portfolio on date t−1, and pit is the price of stock i on date t.

Sharpe Ratio (SR): Evaluates the risk-adjusted performance of an investment portfolio. The calculation formula is as follows: $\mathrm{SR} = \frac{E(R_a - R_f)}{\mathrm{std}(R_a - R_f)}$, where Ra represents the asset return rate, and Rf denotes the risk-free rate.

Maximum Drawdown (MDD): Measures the maximum magnitude of decline from the peak of the cumulative portfolio value.

The calculation formula is as follows: $MDD = max_{\tau\in(0,T)}\left(max_{t\in(0,\tau)}\frac{q_t^c - q_\tau^c}{q_t^c}\right)$. where qtc represents the return value at time t. A smaller MDD indicates that the investment portfolio has stronger risk resistance.

3.2.3 Prediction Error Indicators

Theil Inequality Coefficient (T): Measures the goodness of fit between predicted values and actual values. The calculation formula is as follows:

$T = \frac{\sqrt{1/T\sum_{t=1}^{T}(Y_t^{pre}-Y_t^{true})^2}}{(\sqrt{1/T\sum_{t=1}^{T}(Y_t^{pre})^2}+\sqrt{1/T\sum_{t=1}^{T}(Y_t^{true})^2})}$, Its value ranges between [0, 1]; T = 0 means the predicted values are perfectly fitted to the actual values.

Mean Absolute Error (MAE): Calculates the average of the absolute errors between predicted values and actual values. The calculation formulas are as follows: $\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}\frac{|Y_t^{pre}-Y_t^{tre}}{Y_t^{pre}}$ or $\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}\frac{|Y_t^{pre}-Y_t^{tre}}{Y_t^{pre}}$.

Mean Squared Error (MSE): Calculates the average of the squared errors between predicted values and actual values. The calculation formula is as follows: $\mathrm{MSE} = \frac{1}{2N}\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2$, A smaller MSE indicates a smaller prediction error.

Root Mean Squared Error (RMSE): The square root of the Mean Squared Error (MSE). The calculation formula is as follows: $\mathrm{RMSE} = \sqrt{\frac{1}{l}\sum_{i=N-l+1}^{N}(Y_i - \hat{Y}_i)^2}$, It can better reflect the degree of deviation between predicted values and actual values.

Coefficient of Determination (R^2): Measures the model's ability to explain the data, with a value range between 0 and 1. The closer R^2 is to 1, the better the model's fitting effect. The calculation formula is as follows: $R^2 = 1 - \frac{\sum_{i=1}^{N}(Y_i-\hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i-\overline{Y})^2}$

Directional Accuracy Rate (SR): Measures the accuracy of a model in predicting the direction of price movements. It is a crucial indicator in quantitative investment for evaluating a model’s ability to predict trend directions.

3.3. Analysis of Experimental Results

In the prediction of futures PI indicators using Ma Chao's FdAE model, the model exhibits strong generalization ability. The R^2 values of the training set and test set are 0.6809 and 0.5780 respectively, with a small difference; in contrast, the R^2 of the NARX network on the test set is only 0.18, indicating severe overfitting.

In the research by Sako et al., the GRU model demonstrates the overall optimal performance. It performs prominently in single-variable exchange rate prediction and multi-variable stock index prediction. With fewer parameters and higher efficiency under short time steps, its performance is superior to that of the LSTM model.

In the prediction of stock price change ranking, the TSDHG model proposed by Liu Zhibin achieves outstanding performance. On the A-share CSI 300, CSI 100, U.S. stock NASDAQ, and NYSE datasets, its ranking indicators such as IC, Rank IC, and Precision@N all outperform those of baseline models including MLP, LSTM, and GATs. For example, on the CSI 300 dataset, the IC of TSDHG reaches 0.0743, the Rank IC is 0.0725, and the Precision@30 is 56.68%. In the investment backtesting, the daily average return of the Top30 portfolio is 0.0316, with a Sharpe Ratio of 1.0660, showing excellent risk-return performance. Ablation experiments indicate that modules such as the improved Transformer and static-dynamic hypergraph can all enhance the model's performance.

In terms of cryptocurrency portfolio management, the APMHG model has significant advantages. Its AVR reaches 0.0882, Sharpe Ratio is 1.0322, and Maximum Drawdown is 0.0762, which are far superior to those of traditional strategies such as UBAH and UCRP, as well as deep learning models such as PPN and GPM. The performance of the ablation model with the market hypergraph module removed drops sharply, proving that this module is crucial for capturing the implicit correlations between assets.

The LSTM model combined with wavelet analysis proposed by Ouyang Hongbing et al. achieves excellent performance in predicting the Dow Jones Industrial Average. For the model based on wavelet-reconstructed data, the Theil coefficient of the test set is 0.0005 and the Mean Absolute Error is 0.0008. Its accuracy is higher than that of models such as MLP, SVM, and GARCH, and it can accurately capture the long-term and short-term dynamic trends of the index.

4. Suggestions

4.1. Current Challenges

Most existing models rely heavily on traditional financial data such as prices and trading volumes, while underutilizing unstructured data like news texts and social media sentiment. Additionally, noise and outliers in the data tend to interfere with model learning; for some assets, explicit correlation relationships are lacking, making it difficult to mine implicit relationships.

Most models perform well in specific markets but experience performance degradation when migrated across markets. When facing extreme market volatility, models are prone to failure and struggle to capture the impact of unexpected factors on time series.

Deep learning models (e.g., hypergraph neural networks, LSTM) have the characteristic of being "black boxes"—their prediction logic is difficult to interpret, which is not conducive to building trust in financial decision-making. Moreover, some models have complex parameters and high computational costs, resulting in a high threshold for practical application deployment.

4.2. Recommendations for Future Outlook

Introduce natural language processing technology to extract text features from news, research reports, and public opinions, and combine them with industry financial reports and macroeconomic data to construct multi-modal inputs. Additionally, develop adaptive denoising algorithms to reduce the interference of data noise and improve feature quality.

Design cross-market transfer learning frameworks to leverage the common knowledge of different markets and improve model adaptability; integrate extreme risk simulation training to enhance the model’s ability to respond to sudden market fluctuations; explore lightweight model structures to reduce computational costs and promote practical application.

Combine technologies such as attention mechanisms and causal inference to visualize the key decision-making basis of the model; deeply integrate prediction results with quantitative trading strategies, develop end-to-end decision-making systems, and provide more practical risk prevention and control measures as well as investment recommendations for investors and regulatory authorities.

5. Conclusion

This study focuses on financial time series prediction, covering assets such as stocks, cryptocurrencies, futures, and exchange rates. It proposes models including hypergraph neural networks, LSTM, and functional deep neural networks, and integrates technologies like wavelet analysis and reinforcement learning to optimize prediction performance from two dimensions: temporal feature extraction and asset correlation modeling. The effectiveness of the models is verified using multi-market datasets.

Experimental results show that hypergraph neural networks can effectively capture the static and dynamic correlations of assets, achieving outstanding performance in stock price change prediction (with an IC of 0.0743 in the A-share market) and cryptocurrency portfolio management (with an AVR of 0.0882); LSTM combined with wavelet analysis improves the accuracy of index prediction; FdAE alleviates the overfitting problem in futures prediction; and GRU has an efficiency advantage in short-time-step prediction.

Overall, deep learning models integrating spatiotemporal features outperform traditional models in both risk-return performance and accuracy in financial time series prediction. Future efforts should focus on the following aspects: fusing multi-source unstructured data to enhance the cross-market generalization ability of models; exploring lightweight and interpretable designs to lower the threshold for practical application; and combining causal inference with quantitative strategies to promote the transformation of models into practical decision-making systems.

References

[1] Zhang H and Wang W G. Neural Network Integration Model for Financial Time Series Prediction. Journal of Beijing University of Posts and Telecommunications, 2025, 48(1):127-132.

[2] Wu J, Zhang W, Liu Y, Wang Y, Zhang Y and Yin B, Multi-attributeSD: An Adaptive Self-Knowledge Distillation Method for Financial Time Series Prediction, in IEEE Transactions on Computational Social Systems, 2025.

[3] Wang J, Wang J, Fang W, Niu HL, Financial Time Series Prediction Using Elman Recurrent Random Neural Networks, Computational Neuroscience, 10.1155/2016/4742515, 2016, (1-14), 2016.

[4] Ma Y L, Myeong C C, Weng Y L, GNN-MAM: A graph neural network based multiple attention mechanism for regional financial risk prediction,Alexandria Engineering Journal,Volume 127, 2025, Pages 1004-1014, ISSN 1110-0168

[5] Kady S, Berthine N M and Paulo C R, Neural Networks for Financial Time Series Forecasting, Entropy, 2022.

[6] Ma C, Hou T C, Xu J H, Zhang Z H, Lan B, Application of functional deep neural network in financial time series prediction, Journal of Xuzhou Institute of Technology (Natural Sciences Edition), 2017.

[7] Yang H B, Huang K, Yan H J. Prediction of Financial Time Series Based on LSTM Neural Network. Chinese Journal of Management Science, 2020, 28(4): 27-35.

[8] Liu Z B, Research on financial time series prediction and its application based on hypergraph neural network. Shanxi University of Finance & Economics. 2024.

[9] Zeng Z H, Wang J H, Hu L Y, et al. Study on the accuracy of LSTM model under optimization algorithm in financial prediction. Computer Programming Techniques and Maintenance, 2023, (11):60-64+75. DOI: 10.16184/j.cnki.comprg. 2023. 11.019.

[10] Sun Y M, Financial Transaction Network Risk Prediction Model Based on Graph Neural Network, Procedia Computer Science, Volume 261, 2025, Pages 763-771, ISSN 1877-0509