

A Review of the Application of Transformer in Financial Market Risk Prediction

Lewei Yu

Sino-British International College,
University of Shanghai for Science
and Technology, Shanghai, 200031,
China
E-mail address: 15221372824@163.
com

Abstract:

Given that the financial market is influenced by political, economic and social factors, the instability it exhibits leads to an increase in investment risks, and traditional models have certain limitations in predicting financial market risks. Therefore, based on the RNN model and LSTM model, this paper reviews the application of the Transformer model in financial market risk prediction. By utilizing the attention mechanism of the Transformer model to predict complex high-frequency time series, and summarizing the fields where the Transformer model can be applied, a comparative analysis of the usage of each model was conducted, and the limitations of the Transformer model were also proposed. The study found that the Transformer model is more suitable for analyzing high-dynamic data and industries compared to traditional models. This review summarizes the application of the Transformer model in financial market risk prediction and adopts data from Wanfang Database and the National Philosophy and Social Sciences Literature Center. Relevant literature from 2018 to 2025 was collected, and when screening, articles with higher impact factors from journals and international conference papers were given priority. The topics of the literature include the attention mechanism of the Transformer model, risk response measures in the financial market, the prediction accuracy of the model, the field of natural language processing (NLP), and the spread trading market.

Keywords: Transformer model, RNN model, LSTM model, financial time series analysis, financial market risk prediction, machine learning model

1. Introduction:

1.1 Research background

The risks in the financial market exhibit characteristics of complexity and globalization. The financial market is influenced by political factors, inflation and interest rates, changes in the financial market struc-

ture, as well as ESG and extreme climate risks, thereby increasing the risks associated with the returns on the invested funds. The prediction of financial market risks can prevent systemic risks, optimize the financing structure of the financial market, establish strategies for financial risk hedging, and enhance the response to tail risks. In recent years, high-frequency time series prediction using transformers has become the main tool for predicting financial risks.

The significance of this study lies in that, through a comprehensive review of the application of the transformer model in financial market risk prediction, it can provide investors with suggestions and methods for financial market risk management, reduce the instability in the financial market, and point out the future direction for academic research. At the same time, this study also helps to promote the optimization of the training algorithm and model structure of the transformer model, enhance the real-time performance and the ability to handle complex time series of the model, and improve the prediction accuracy of the model.

1.2 Review of the Research

In recent years, scholars have begun to pay attention to the application of the Transformer model in the high-frequency risk prediction of finance. Ji Naigeng [6], Wang Yuhua [13] (2024), and others have proposed the attention mechanism of the Transformer model. They described the processing method of the Transformer model for long sequence problems and the issue of capturing sequence dependencies. The core advantage of the Transformer model is its self-attention mechanism, which breaks through the short-term memory limitations of traditional RNN/LSTM and avoids the problem of gradient disappearance. It allows them to measure the importance of different time steps in the sequence and learn complex patterns and relationships across long periods. The Transformer model is more flexible than traditional models and can handle the non-linearity in data, making it very suitable for the rapidly changing financial market. Due to the uncertainty and volatility of the financial market, measures for responding to financial market risks have emerged. Ding Bing [3] (2021), Ding Xin [4] (2022), and Bai Lu [1] et al. proposed financial market risk response measures such as optimizing the financing structure and efficiency of the financial market, expanding financial tools, innovating financial products, establishing financial risk transfer strategies, strengthening financial market risk management, and improving the Probit model of the Firefly Algorithm to guide financial risk management. The Probit model enhances the modeling ability of high-dimensional nonlinear relationships, has robustness for extreme event prediction, can monitor dynamic risks in real time, such as the high-frequency trading risk control system can complete

the risk revaluation of the entire market within 5 seconds. It is beneficial for investors to make decisions in the rapidly changing financial market. Wen Xinxian [12] (2022) and Joel Paul [7] (2024) et al. proposed research on the prediction accuracy, prediction ability, and establishment of early warning models of the Transformer model. The improvement of prediction accuracy and prediction ability mainly depends on the enhancement of data quality. The financial market is prone to be affected by noise. Using CleanLab can automatically identify wrongly labeled samples and achieve the function of noise cleaning.

Financial data has certain heterogeneity, multi-source nature and imbalance. When establishing an early warning model, heterogeneous graph neural networks can be used to fuse multi-source data to enhance the early warning for extreme events. Ji Naigeng [6] (2024) and Kai Hanl [8] (2021) et al. proposed the application of the Transformer model in the field of natural language processing (NLP). Using the self-attention mechanism to capture the context of words in sentences and paragraphs, in these basic translation, classification, annotation and other tasks, the accuracy has increased by 30%-50%. At the same time, it also unifies the modeling of text, images, sounds and videos, making cross-modal fusion a reality. Kima [10] (2019) proposed using deep learning (DL) methods to predict the behavior of retail investors in the spread trading market. The deep learning method can deeply mine the cross-asset relationships and dynamically capture the behavior patterns, thereby improving the prediction accuracy of hedge fund behaviors. The prediction accuracy in noisy environments has steadily increased by 35%.

1.3 Purpose of Research

This review aims to summarize the application of Transformer models in financial market risk prediction in recent years, explore the advantages of Transformer models compared to traditional prediction models, as well as the challenges faced by Transformer models, and propose possible future research directions.

The purpose of financial market risk prediction is to reduce the increase in investment risks caused by the instability and increasing volatility of the market. Traditional RNN models have difficulty in maintaining long-term information and suffer from problems such as gradient disappearance and gradient explosion, making the models unable to rely on them for a long time. The LSTM model has certain lag in predicting highly fluctuating situations, and thus has insufficient "memory update" speed for extreme events. Both of these models also lack parallel processing capabilities. The VaR model usually assumes a normal distribution, so it will seriously underestimate financial data with "fat-tailed" characteristics, which will mislead investment decisions and lead to the failure of hedging strategies.

In order to comprehensively summarize the application of the deep learning Transformer model in financial market risk prediction, this review adopted data from Wanfang Database and the National Philosophy and Social Sciences Literature Center, collecting relevant literature from 2018 to 2025. The keywords include “Financial time series analysis”, “Transformer model”, “RNN model”, “LSTM model”, “Financial market risk prediction”, “Machine learning model”, etc. During the screening process, journal articles with higher impact factors and international conference papers were given priority.

In the literature analysis, this review categorizes all the literature according to the research questions (such as the self-attention mechanism, multi-head attention mechanism, financial market risk response measures, how to improve the prediction accuracy of the model, etc.) and application fields (such as the natural language processing (NLP) field, spread trading market, etc.). At the same time, the experimental settings, data sets, selection of evaluation indicators, and model performance of different studies are evaluated to provide a comprehensive comparative analysis.

The innovation of this review lies in comparing various models for financial market risk prediction, including traditional models such as RNN, LSTM, CNN, and GNN. It reveals their advantages and limitations in different situations, as well as the scenarios in which they are applicable. Therefore, the Transformer model emerged. This review also studies the effectiveness of financial market risk prediction under the combination of various traditional models, such as the GNN-Transformer model, CNN-Transformer model, and iTransformer model. By combining a series of traditional models, new models with broader applicability are established, which is the innovation point of this review.

2. Research Method

The content of this review study is to use the Transformer model to predict financial market risks. In this field, the main future research directions are to use multivariate features to study single-variable or multivariate multi-step prediction problems, solve the problem of inference delay in the Transformer model, improve the prediction efficiency, adapt to the nanosecond-level high-frequency trading risk control in the financial market, and improve the performance of the model in handling complex time series and nonlinear relationships. The theoretical framework adopted in this review mainly includes the theoretical framework of time series modeling, such as the Transformer model processing long-term dependencies based on the self-attention mechanism, combining position encoding to capture the temporal features of market data. The multi-head attention mechanism analyzes patterns at

different time scales in parallel, enhancing the ability to extract nonlinear features. And the weighted evaluation methods used in risk prediction and management, which are operations for weighting the factors affecting financial risks, thereby calculating the risk level. The technical means involved in this review include the comprehensive ability of CNN-Transformer to model short-term and long-term dependencies in time series. GNN-Transformer combines graph neural networks to model the correlation of assets. Machine learning (ML) methods improve performance by processing high-dimensional and nonlinear feature interactions in an unmodelled way. Deep learning (DL) methods are applied to numerical time series data, with process automation, eliminating the need for domain expertise.

3. Research Outcome

3.1 The attention mechanism of the Transformer model

The core of the Transformer model lies in the self-attention mechanism and the multi-head attention mechanism. The self-attention mechanism captures the temporal characteristics of market data and handles long sequence dependencies by calculating the correlations between various positions in the input sequence. The self-attention mechanism can directly calculate the correlation weights between any two points in the sequence without the need for layer-by-layer information transmission, thus avoiding the risk of gradient disappearance. The multi-head attention mechanism can concurrently focus on different parts of the input sequence, more comprehensively capturing the complex features and long-distance dependencies of the data, and enhancing the ability of nonlinear feature extraction.

3.2 The application of Transformer models in the financial field

The Transformer model has demonstrated certain advantages in both the futures market and the treasury bond market. Leveraging its ability to capture jumps in risk, the Transformer identifies the impact of extreme events on prices. For instance, in the futures market, the optimization of the crude oil hedging strategy by the Transformer reduced the volatility prediction error of the crude oil futures market in 2023 by 34% compared to the GARCH model, and the hedging cost decreased by 18%. In the treasury bond market, the interest rate risk model driven by the Transformer, during the period of aggressive interest rate hikes by the Federal Reserve in 2022, reduced the prediction error of the 10-year treasury bond yield by 23% compared to LSTM, and warned of the yield inversion

risk one month in advance.

3.3 The limitations of the Transformer model

However, the application of Transformer models in financial market risk prediction still faces both technical and non-technical challenges. Technically, the issue of inference delay in Transformer models needs further optimization to meet the millisecond-level decision-making requirements of the financial market. The performance of Transformer models in complex time series needs to be further improved by introducing more advanced training algorithms and regularization techniques to optimize the model structure and parameters. Non-technically, the cross-border data flow is restricted by regulations such as GDPR and CCPA, which affects model training. The cognitive gap between the technical team and traders leads to difficulties in the actual implementation of the model, and resolving the conflicts between the teams is the key to the model's successful deployment.

4. Conclusion

In conclusion, the Transformer model represents a significant advancement in the field of financial market risk prediction. It can capture and handle long-term dependencies between data through its attention mechanism, and effectively overcomes the limitations of traditional RNN models and LSTM models that cannot handle the complex nonlinear relationships in financial data. By utilizing the self-attention mechanism and multi-head attention mechanism of the Transformer model, it can also analyze and predict highly dynamic financial markets, such as stock prices and currency exchange rate predictions, which are highly volatile and uncertain data influenced by economic, political, and social factors. Therefore, the Transformer model can help investors in the financial field to do a thorough job of risk prediction, help avoid investment decisions' mistakes and capital losses, and provide new possibilities in investment decision-making, financial management, and risk management. With the development of artificial intelligence machine learning models and deep learning fields, the Transformer model will have greater development potential and space. Of course, the technical problem of inference delay currently existing in the Transformer model also needs to be further optimized and explored to optimize its parameter structure. The Transformer model also needs to overcome the cognitive gap caused by human factors and whether this technology can pass the relevant laws of the industry also needs further research.

5. Reference

- [1] Bai Lu, Wen Wen. (2024) Performance Analysis of Probit Model Based on Improved Firefly Algorithm in Digital Financial Risk Prediction. *Journal of Pingdingshan University*, 2024, 39(2).
- [2] Caosen, X. , Jingyuan, Li., Bing, F., and Baoli, L., (2023) A Financial Time-Series Prediction Model Based on Multiplex Attention and Linear Transformer Structure. *Appl. Sci.* 2023, 13, 5175.
- [3] Ding Bing. (2021) Financial Market Risk Prediction and Countermeasures. DOI:10.16301/j.cnki.cn12-1204/f.2021.05.056
- [4] Ding Xin. (2022) Financial Market Risk Prediction and Countermeasures. DOI:10.19995/j.cnki.CN10-1617/F7.2022.18.070
- [5] Fan Yangyang, Guo KaiTai, Li Yiyang. (2025) Prediction and Analysis of Soybean Oil Futures Prices Based on LSTM-Transformer Model.
<https://doi.org/10.12677/aam.2025.144152>
- [6] Ji Naigeng, Zhou Xinbo. (2024) An institution reserve fund prediction method based on LSTM, Transformer and LightGBM. <https://doi.org/10.12677/csa.2024.142025>
- [7] Joel, P., (2024) Financial Time Series Analysis with Transformer Models.
<https://www.researchgate.net/publication/387524930>
- [8] Kai, H., An, X., Enhua, W., Jianyuan, G., Chunjing, X., Yunhe, W., (2021) Transformer in transformer. In: 35th Conference on Neural Information Processing Systems
- [9] Katarina, V., Tomas, K., Lucia, S., and Peter, A., (2018) Financial Risk Measurement and Prediction Modelling for Sustainable Development of Business Entities Using Regression Analysis. *Sustainability* 2018, 10, 2144.
- [10] Kima , A., Yangc , Y., Lessmanna , S., Mab , T., Sungh, M.-C., Johnsonb, J.E.V. (2019) Can Deep Learning Predict Risky Retail Investors? A Case Study in Financial Risk Behavior Forecasting.
- [11] Qian Jun, (2024) Financial Risk Prediction of Internet Finance Emotions. DOI: 10.12248/j.issn.1007-676X.2024.016.034
- [12] Wen Xinxian. (2022) High-frequency Financial Time Series Prediction by Integrating Empirical Mode Decomposition and Linear Transformer.
Modern Electronic Technology, 2022, 45(23).
- [13] Wang Yuhuan, Liang Zhiyong. (2024) Financial Time Series Prediction Based on iTransformer Model. *National Philosophy and Social Sciences Literature Center*, 2024, 15.
- [14] Zhen, Z., Rachneet, K., Suchetha, S., Saba, R., Tucker, B., Manuela, V., (2023) Financial Time Series Forecasting using CNN and Transformer. *Association for the Advancement of Artificial Intelligence* (www.aaai.org).