Comparison of Prediction Effectiveness in Deep Learning Perspective of China's Data Finance

Meilin Du

Abstract:
The study has developed a financial time series forecasting model, the Transformer-Encoder model, which utilizes the attention mechanism. This model has been applied to predict the closing price of the Shanghai Stock Exchange (SSE) index, a reliable indicator of financial trends. Furthermore, the study has conducted a comparative analysis, evaluating the performance of our model against other deep learning models, machine learning models, and traditional time series data forecasting models across short, medium, and long-term forecasting periods. Our study has yielded the following key findings: Firstly, the Transformer-Encoder model, leveraging the attention mechanism, demonstrates strong performance in predicting closing prices across short, medium, and long-term periods. This indicates the model's viability in handling non-stationary financial data and its potential as a forecasting tool applicable to time series prediction problems within the economic sphere. Secondly, compared to alternative deep learning models, machine learning models, and traditional time series data forecasting models, our proposed model consistently outperforms them.
Keywords: Deep Learning; Data Finance; Transformer-Encoder model

1. Introduction

Time-series studies of financial markets are of great significance among investors and policymakers. The trend of the financial market can reflect the development status of a country's economy. For investors, an accurate grasp of the financial market trend can help them make more informed investment decisions, avoid risks, and maximize benefits. At the same time, policymakers also hope to understand the economy's current status and future development trends through the study of financial market trends to formulate corresponding financial policies and measures to promote economic stability and development. The time series study of the financial market is a fascinating topic because many factors, including macroeconomic indicators, policy changes, market psychology, and various supply and demand factors, influence the financial market trends. Patterns and trends can be discovered by studying the time series of financial markets, providing quantitative analysis and forecasting methods.

In recent years, with the development of technologies such as machine learning and deep learning, researchers have begun to apply these methods to studying time series in the financial market to improve the accuracy and effectiveness of forecasting. Introducing artificial intelligence technologies has brought new opportunities and challenges to time series research in financial markets while also providing investors and policymakers with more tools and methods to avoid risks and maximize benefits. As a powerful machine learning algorithm, deep learning can handle high-dimensional, nonlinear, and discontinuous data. In financial market forecasting, financial data often has complex features and patterns, and traditional statistical models may not be able to capture the nonlinear relationships and implicit patterns well. On the other hand, deep learning models can learn the complex patterns and regularities in the data through the combination and training of multi-layer neural networks to achieve more accurate predictions. Therefore, using deep learning techniques for time series forecasting in financial markets has important theoretical and practical value.

The rest of the paper is organized as follows. Section 2 reviews some important papers after the trade war. Section 3 introduces the data used in this paper. Section 4 shows the results of the discussion. Section 5 concludes.

2. Literature Review

The Transformer-Encoder is an encoder model based on the Transformer architecture. It is a deep learning model designed to process sequential data, and it initially achieved significant success in tasks related to natural language processing [1]. The Transformer-Encoder utilizes a self-attention mechanism to capture contextual relationships within the input sequence without recursive or convolutional operations. This allows for parallel computation, resulting in reduced computational complexity, and facilitates the model's ability to maintain a long memory of context[2].

Yoon and Swales suggested that multivariate analysis techniques may have uncertain effectiveness in stock price prediction, while neural network methods have shown their capability in solving complex prediction problems. They validated the ability of neural networks by comparing them with multiple discriminant analysis methods, which yielded good predictive results[3]. Kaliyaperumal et al. argued that historical data plays an important role in predicting future market directions and has predictive power in financial investment decisions. They designed automated computer programs that utilize data mining and prediction techniques to help investors discover hidden patterns from historical data[4]. Billings et al. compared integrated methods (Random Forest, Kernel Function, and AdaBoost) with single classifier models (Neural Networks, Logistic Regression, Support Vector Machines, and K Nearest Neighbors) to predict the movement of European financial markets. They collected data from 5,767 publicly traded European companies and found that integrated algorithms have the potential to be a new area of research in financial forecasting[5].

Selvin and Vinayakumar et al. applied three deep learning methods, namely RNN, LSTM, and CNN, to predict the share prices of NSE-listed companies. They compared the prediction performance of these methods[6]. Nelson and Pereira et al. investigated the use of LSTM networks in financial markets. They used price history data and technical analysis indicators to predict future trends in stock prices. They found that LSTM achieved an average accuracy of 55.9% in predicting the rise or fall of specific stock prices[7]. Araújo et al. proposed a deep learning model training method based on gradient decay. They evaluated the model's predictive performance using 12 financial time series data from relevant global stock markets. The results showed that the model is competitive and effective in predictive performance[8]. Shin et al. proposed a deep multimodal reinforcement learning method for stock price prediction. They combined long and short-term memory neural networks with convolutional neural networks and used daily data from 256 stocks listed on the KOSPI in South Korea to train and test their model. Their study showed that their model achieved good prediction results[9].

Similarly, Qi et al. constructed a long-term deep-learning model based on a simple recurrent neural network. They compared this model with short-term memory (LSTM), bi-directional long-short-term memory (BiLSTM), and gated recurrent units (GRUs). Their research focused on predicting the euro/sterling currency using 15-minute interval data. They found that their long-term deep-learning model outperformed traditional models in the prediction simulation[10].

Relevant literature shows that traditional financial time series analysis models (e.g., AR model, ARIMA model, etc.) have a certain descriptive ability of the operation mechanism and trend of the financial market in a specific context and play an important role in financial forecasting research. Meanwhile, with the continuous development of artificial intelligence algorithms, scholars have begun introducing machine learning algorithms into finance to construct new algorithmic models to explore the intricate financial system. These new prediction models include machine learning prediction models such as Support Vector Machine (SVM), Random forest, Bayesian network, and deep learning prediction models such as shallow BP Neural Network (Back-Propagation Network) and Convolutional Neural Network (CNN). In recent years, scholars at home and abroad have begun to explore combining basic algorithms into hybrid algorithmic models to predict the direction of financial time series more accurately. For example, the CNN-GRU neural network financial data prediction model constructed based on CNN and GRU, the LSTM-GEM financial data prediction model constructed based on LSTM, and the SDAE-LSTM financial data prediction model have demonstrated excellent prediction accuracy in real datasets, and have made significant contributions to the research field of financial time series prediction. In financial data forecasting tasks, the Transformer-Encoder can be used to learn the temporal patterns present in input sequences. This facilitates financial data forecasting by enabling the model to predict future trends or values. The Transformer-Encoder excels in handling long-term dependencies and capturing complex relationships within sequences, all while exhibiting robust generalization capabilities. As a result, the Transformer-Encoder holds great promise for applications within the financial domain.

3. Data and Method

In this study, I chose the financial data from the Shanghai Stock Exchange as the target for empirical analysis. I selected seven variables as input features: closing price, opening price, high price, low price, change rate, turnover amount, and turnover volume of the SSE index. All the data I obtained from the exchanges include daily data of trading days from 2005 to 2022. The following are our selected variables and their descriptive statistics (see Table 1).

Table 1 Variables and their descriptive statistics

	Average	Standard error	Minimum value	Maximum value
Closing price	2805.41	835.23	1011.50	6092.06
Opening price	2802.42	835.19	1007.90	6057.43
High price	2827.73	844.61	1019.92	124.04
Low price	2776.41	919.10	1112.65	6040.71
Change rate	3085.45	1.76E+13	9.57E+12	6701.26
Turnover amount	1.86E+11	1.69E+11	3.68E+09	1.31E+12
Turnover volume	16386554276	12771974425	676958500	85713280700

To confirm the effectiveness of the constructed Transformer-Encoder model and compare it with different benchmark prediction models, this study carefully evaluated the appropriateness of various evaluation metrics and referred to existing research. The final decision was to use a combination of four evaluation metrics, including Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE), Root Mean Square Error (RMSE), and R-squared (R2), to assess the predictive performance of the Transformer-Encoder model and compare it with other benchmark models. Below are the definitions of each evaluation metric:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|Y_t - \hat{Y}_t\right|}{Y_t}\times 100\%$$

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|Y_t - \hat{Y}_t\right|}{\left(\left|\hat{Y}_t\right| - \left|Y_t\right|\right)/2}\times 100\%$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}\times 100\%$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}{\sum_{t=1}^{n}\left(Y_t - \bar{Y}_t\right)^2}$$

Y_t is the observed value, $\bar{Y}_t$ is the average observed value, and $\hat{Y}_t$ is the predicted value.

To construct the Transformer-Encoder financial data prediction model, this paper adopts the following network structure and parameter settings: encoder_layers is 3, the number of attention heads is 8, the number of hidden state dimensions is 512, and the number of Stacks is six by using the Dropout regularization method. The Dropout rate is set to 0.05; the activation function is selected as real. The learning rate is set to 0.001. to control the number of training iteration steps, the upper limit of the training iteration steps is set to 100 times in this paper.

To demonstrate the advantages of the Transformer-Encoder model in financial data prediction, this paper carries out a comparative analysis with other methods, including deep learning methods such as LSTM, GRU, and BP, as well as SVR and the traditional time-series method ARIMA.In constructing the model input set, the LSTM, GRU, and BP neural network processing is the same as the Transformer-Encoder method in this paper. In contrast, the ARIMA model only selects the closing price as the input data for univariate prediction of future closing prices.

Specifically, the LSTM and GRU neural networks are set with one hidden layer and 16 neurons. The BP neural network adopts a network structure containing three fully connected layers; the hidden layer contains 64 neurons, and the Relu activation function is applied. The SVR model uses the default parameters in Scikit-learn, in which the gamma value is set to 10, and the error term penalty factor C is set to 1. The ARIMA model was selected as a parametric ARIMA(1,1,1) for data prediction.

4. Discussion

4.1 Analysis of short-term forecasts

In the short-term training, the optimal prediction window is seven by performing a sensitivity analysis with the training window of the Transformer-Encoder financial data prediction model from period 1 to period 10. To compare the effectiveness of the Transformer-Encoder financial data prediction model proposed in this paper with six other comparative modeling approaches for short-term financial data prediction, I conducted predictions on the test set for the closing price of the next time step and compared the evaluation metrics. The results are presented in Table 2. The models used include Transformer-Encoder, BP, SVR, LSTM, GRU, and ARIMA, with a training window of 7 periods for each model. The values marked with * in the table indicate the best results in terms of the evaluation metrics(see Table 2).

Table 2 Comparison of short-term forecast accuracy

	Transformer-Encoder	BP	SVR	LSTM	GRU	ARIMA
MAPE	0.0312	0.0449	0.0101*	0.0393	0.0493	0.0503
SMAPE	0.0307*	0.0491	0.0874	0.0351	0.0485	0.0312
RMSE	0.0154*	0.0261	0.0514	0.0191	0.0194	0.0414
R^2	0.9328	0.8621	0.5423	0.9012	0.8952	0.9515*

According to the comparison results, when using the Transformer-Encoder model for short-term forecasting of financial data, both evaluation metrics, SMAPE and RMSE, show superiority over other forecasting models. This indicates that the Transformer-Encoder model is superior in prediction accuracy compared to other methods. Despite being slightly lower than the ARIMA model in terms of the coefficient of determination, the Transformer-Encoder model still exhibits a higher predictive ability than other deep learning and machine learning models. This result indicates that the Transformer-Encoder model has a strong adaptive ability and high accuracy in short-term financial data prediction, and it can be applied in the short-term prediction of financial data.

4.2 Analysis of medium-term forecasts

Comparisons were made with training windows of periods 30, 60, 90, 120, and 150, and it was found that the best predictions were produced when the training window was period 90. Therefore, in this paper, I chose the condition with a training window of 90 minutes for medium-term prediction. Also, I compared the mid-term prediction results under models of Transformer-Encoder, BP, SVR, LSTM, GRU, and ARIMA. The values marked with * in the table represent the best-performing results in the corresponding evaluation metrics (see Table 3).

Table 3 Comparison of medium-term forecast accuracy

	Transformer-Encoder	BP	SVR	LSTM	GRU	ARIMA
MAPE	0.0787*	0.0905	0.1847	0.0881	0.1174	0.0920
SMAPE	0.0743*	0.0934	0.1690	0.0831	0.1260	0.0776
RMSE	0.0395*	9.3489	0.0900	0.0428	0.0657	0.0983
R^2	0.6624	0.4349	-0.9288	0.5779	-0.0248	0.7166*

Based on the comparison of evaluation indexes, I can conclude that the Transformer-Encoder financial data forecasting model outperforms other comparative forecasting models in terms of MAPE, SMAPE, and RMSE, which indicates that the model outperforms other methods in terms of medium-term forecasting accuracy. Although the coefficient of determination of the Transformer-Encoder model is slightly lower than that of the ARIMA model, it is still higher than that of other deep learning and machine learning models, which indicates that the model has a high ability to fit and has good forecasting performance. In summary, the Transformer-Encoder financial data prediction model proposed in this paper performs better than the other six benchmark models selected in comparison and is suitable for medium-term prediction of financial data.

Analysis of long-term forecasts

By comparing the results under the training window of 120, 180, 240, 300, and 360 periods, the Transformer-Encoder financial data prediction model gives the best prediction in 240 periods. In this study, models of Transformer-Encoder, BP, SVR, LSTM, GRU, and ARIMA were used for comparison to predict the closing price after 240 trading days. Values marked with * indicate that optimal results were achieved in the evaluation metrics (see Table 4).

Table 4 Comparison of long-term forecast accuracy

	Transformer-Encoder	BP	SVR	LSTM	GRU	ARIMA
MAPE	0.1365*	0.7126	0.2146	0.1567	0.8360	3.7255
SMAPE	0.1324	0.1151*	0.2279	0.143	0.1351	0.3757
RMSE	0.0937*	0.3324	0.1076	0.9821	0.3979	0.4216
R^2	0.1439*	-39.8078	-2.3355	-1.1852	-44.5781	-0.2931

Based on the evaluation metrics, it can be concluded that the Transformer-Encoder financial data forecasting model outperforms other comparative forecasting models in both evaluation metrics, MAPE and RMSE. This indicates that the model is more accurate than other methods in long-term forecasting. In addition, the coefficient of determination of the Transformer-Encoder financial data prediction model is also higher than that of other deep learning models and machine learning models, which indicates that the model has a higher degree of fit relative to other methods. However, it should be noted that the Transformer-Encoder financial data forecasting model has a slightly different model fit in long-term forecasting relative to short-term and medium-term forecasting.

5. Conclusion

The study has developed a financial time series forecasting model, the Transformer-Encoder model, which utilizes the attention mechanism. This model has been applied to predict the closing price of the Shanghai Stock Exchange (SSE) index, a reliable indicator of financial trends. Furthermore, the study has conducted a comparative analysis, evaluating the performance of our model against other deep learning models, machine learning models, and traditional time series data forecasting models across short, medium, and long-term forecasting periods.

Our study has yielded the following key findings:

Firstly, the Transformer-Encoder model, leveraging the attention mechanism, strongly predicts closing prices across short, medium, and long-term periods. This indicates the model's viability in handling non-stationary financial data and its potential as a forecasting tool applicable to time series prediction problems within the economic sphere.

Secondly, compared to alternative deep learning models, machine learning models, and traditional time series data forecasting models, our proposed model consistently outperforms them.

Nonetheless, it is worth noting that our study has certain limitations. In our window sensitivity analysis for medium and long-term forecasts, the study has only selected a few representative windows, with some intervals between them. Thus, there is a possibility that the study may have missed the best training windows, potentially impacting the accuracy of our predictions. To address this, future research endeavors can explore a denser selection of windows, which will help improve the model's prediction accuracy. By conducting a more expansive and continuous evaluation of different windows, I will better understand their influence on medium and long-term predictions.

Our study successfully develops and applies a Transformer-Encoder financial time series forecasting model with the attention mechanism. The model showcases robust performance in predicting the SSE index closing prices across various forecasting periods. Additionally, it outperforms other models in the comparative analysis. Nevertheless, future research should enhance our window selection methodology to improve the model's predictive accuracy and reliability.

References

[1] Kalyan K S, Rajasekharan A, Sangeetha S. Ammus: A survey of transformer-based pre-trained models in natural language processing[J]. ar**v preprint ar**v:2108.05542, 2021.

[2] Khan S, Naseer M, Hayat M, et al. Transformers in vision: A survey[J]. ACM computing surveys (CSUR), 2022, 54(10s): 1-41.

[3] Yoon Y, Swales G. Predicting stock price performance: A neural network approach[C]//Proceedings of the twenty-fourth annual Hawaii International conference on system Sciences. IEEE, 1991, 4: 156-162.

[4] Kannan K S, Sekar P S, Sathik M M, et al. Financial stock market forecast using data mining techniques[C]//Proceedings of the International Multiconference of Engineers and Computer Scientists. 2010, 1(4).

[5] Ballings M, Van den Poel D, Hespeels N, et al. Evaluating multiple classifiers for stock price direction prediction[J]. Expert Systems with Applications, 2015, 42(20): 7046-7056.

[6] Selvin S, Vinayakumar R, Gopalakrishnan E A, et al. Stock price prediction using LSTM, RNN, and CNN-sliding window model[C]//2017 international conference on advances in computing, communications, and informatics (cocci). IEEE,

2017: 1643-1647.
[7] Nelson D M Q, Pereira A C M, De Oliveira R A. Stock market's price movement prediction with LSTM neural networks[C]//2017 International joint conference on neural networks (IJCNN). Ieee, 2017: 1419-1426.
[8] Araújo R A, Nedjah N, Oliveira A L I, et al. A deep increasing–decreasing-linear neural network for financial time series prediction[J]. Neurocomputing, 2019, 347: 59-81.
[9] Shin H G, Ra I, Choi Y H. A deep multimodal reinforcement learning system combined with CNN and LSTM for stock trading[C]//2019 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2019: 7-11.
[10] Qi L, Khushi M, Poon J. Event-driven lstm for forex price prediction[C]//2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). IEEE, 2020: 1-6.