ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

Stock price prediction based on deep learning

Qinzheng Hu[1,a,#], Yunchao He[2,b,#], Chengrunyan Yang[3,c,#], Haoyan Zhang[4,d,#], Shengjin Liu[5,e,#]

[1]*International Education School, Suzhou University of Science and Technology, Suzhou, China*
[2]*Xi'an Jiaotong-Liverpool University, Suzhou, China*
[3]*College of Science, Civil Aviation University of China, Tianjin, China*
[4]*Beijing Jingshan School, Beijing, China*
[5]*School of Medical Information Engineering, Shandong First Medical University, Tai'an, China*
[a]*aqing0318@outlook.com*, [b]*hhh1296275969@163.com*, [c]*18129231855@163.com*,
[d]*timhaoyan20070725@gmail.com*, [e]*lsj3456@163.com*
[#]*Co-first author*

Abstract: *Because the LSTM neural network model and GRU neural network model have better advantages in forecasting compared with other neural network models, this paper uses the LSTM model and GRU model to build eight different ensemble neural network models to predict stock prices. The errors of the prediction results will be analyzed by using mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE) as the error measurement standards. A comparison will be made to determine the ensemble model with the smallest error and the highest precision.*

Keywords: *Stock price prediction, LSTM, GRU, ensemble model*

1. Introduction

The stock market is vital to the economy, helping to allocate resources, guiding investment, and reflecting the state of the economy. With the globalization and informationization of the global economy, its volatility and complexity are also increasing, placing higher demands on investors and policy makers. As an economic indicator, the stock market trend directly affects the market value of listed companies and the wealth of investors, and indirectly affects the stability and growth of the economy.

Therefore, it is very important to overcome the difficulties of complex market factors and nonlinearity and uncertainty, so as to build a neural network model that can accurately predict stock prices. This article explores deep learning prediction methods, using LSTM and GRU models to create an integrated and more reliable model for stock price prediction.

2. Related Work

In article [1], Guś and J. Możaryn have compared Hybrid Deep Learning Neural Network Controller (HDLNNC) and the Adaptive Model Predictive Control with Nonlinear Prediction and Linearization along the Predicted Trajectory (AMPC-NPLPT), finding AMPC-NPLPT more accurate.

In article [2], Yani Hou have proposed the Bi-LSTM method, based on the LSTM framework, for bidirectional analysis of temporal data to predict stock prices. Besides, it also uses the Adam algorithm for network parameter optimization and improving predictive accuracy.

In article [3], Chaobin Huang and Ximing Cheng have found LSTM outperforms BP, CNN, RNN, and GRU in predicting stock prices. They also found that it had advantages in dealing with long-term dependencies.

In article [4], Zhen Liu have developed a stock price prediction model by using LSTM and batch processing, which iterates over data fully during training to enhance accuracy.

In article [5], Ruixue Zhang and Yongtao Hao have introduced an encoder-decoder framework blending LSTM, BiLSTM, and an attention mechanism to address long-term dependencies, nonlinearities in time series and mitigating gradient issues. This method has a good effect on stock price prediction.

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

In article [6], Di Zhang analyzed various models like KNN, SVM, CNN, and CNN-LSTM by using CITIC Securities' daily data. After analyzing the error, the CNN-LSTM is found to be the most accurate for stock price predictions.

In article [7], Wenquan Cui and Qingfang Wang introduced a dual encoder-decoder approach with a scoring formula for information extraction and inference. This method integrated multidimensional stock data and market sentiment, improving forecast accuracy and performance.

In article [8],Shimin Huang uses R language software to establish ARIMA model to analyze and forecast the stock price of China Merchants Bank from January 4, 2021 to June 30, 2022, and the result shows that this model is effective in predicting the short-term change of stock price.

In article [9], Jinlei Han and Pingping Xiong developed a method combining sentiment analysis, feature extraction, and LSTM for stock prediction. They improved accuracy by using SnowNLP for sentiment, AdaBoost for features, and LSTM for forecasting, plus GM(1,1) for residual correction.

In article [10], Xiaojie Li and Huan Xia used sklearn's linear regression to predict stock prices, demonstrating the benefits of quantitative trading. While it is accurate, complex factors can reduce its precision. So, it requires further research for reliability.

In article [11], Meng Li presentd an LSTM-Wavelet model. It processed each layer of wavelet coefficients and effectively captured data nonlinearity to minimize the impact of noise on predictions.

3. The methodology employed in this article

3.1. The network models in this article—LSTM model and GRU model

3.1.1. The LSTM (Long Short-Term Memory) model

Long Short-Term Memory (LSTM) is a special type of recurrent neural network (RNN) designed to handle long-term dependencies. It consists of four interconnected layers forming a chain-like structure. The core of LSTM lies in its cell state, functioning like a conveyor belt ensuring consistent information flow through linear operations. Its unique gates, such as Forget, Input, and Output Gates, allow selective passage of information, enabling effective information management.

3.1.2. The Gated Recurrent Unit (GRU) model

The Gated Recurrent Unit (GRU) is a neural network similar to LSTM, designed for handling long-term dependencies in sequential data with fewer parameters, reducing computational costs. GRU utilizes gating mechanisms, including reset and update gates, to control information flow between time steps, addressing the gradient vanishing problem in traditional RNNs.

3.2. The approach/methodology of this paper

This study explores a novel recurrent neural network architecture that integrates two popular models, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). This article examines eight different combinations of these neural networks, including single-layer LSTM model (LSTM), single-layer GRU model (GRU), double-layer LSTM model (D-LSTM), double-layer GRU model (D-GRU), single-layer LSTM with double-layer GRU model (LSTM-D-GRU), double-layer LSTM with single-layer GRU model (D-LSTM-GRU), double-layer LSTM with double-layer GRU model (D-LSTM-D-GRU), and single-layer LSTM with single-layer GRU model (LSTM-GRU). The eight ensemble models are evaluated on the test set, and their predictions are compared to select the optimal model.

Here are the diagrams of various ensemble neural network models (Figure 1-8):

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

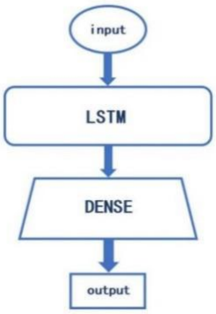

Figure 1: The schematic diagram of a single-layer LSTM model

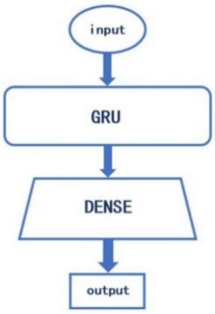

Figure 2: The schematic diagram of a single-layer GRU model

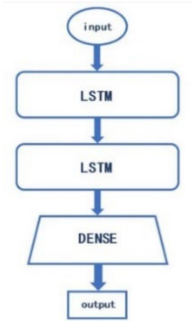

Figure 3: The schematic diagram of a double-layer LSTM model (D-LSTM)

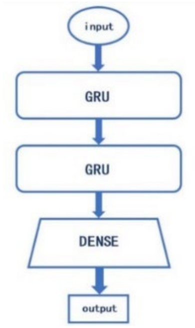

Figure 4: The schematic diagram of a double-layer GRU model (D-GRU

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

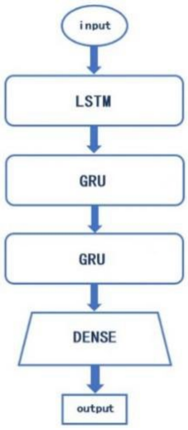

Figure 5: The schematic diagram of a single-layer LSTM with double-layer GRU model (LSTM-D-GRU)

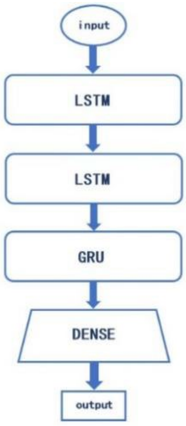

Figure 6: The schematic diagram of a double-layer LSTM with single-layer GRU model (D-LSTM-GRU)

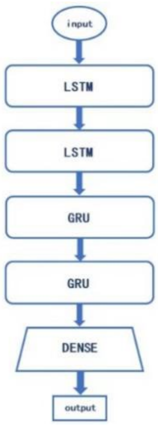

Figure 7: The schematic diagram of a double-layer LSTM with double-layer GRU model (D-LSTM-D-GRU)

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

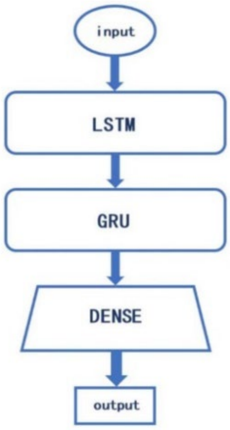

Figure 8: The schematic diagram of a single-layer LSTM with single-layer GRU model (LSTM-GRU)

4. Experiment

4.1. Experimental Data

4.1.1. Data processing

The dataset used in this study is related to Inner Mongolia First Machine (formerly known as Northern Entrepreneurship), which contains 14 columns of data set variables. Since some of the data are not useful, this paper will conduct data screening to remove the useless part and retain the useful part.

4.1.2. The processed data

After data processing, there are 12 data remaining, including: closing price (close), opening price (open), highest price (high), lowest price (low), latest trading price (last), change amount (change amount), change percentage (change percent), turnover rate (turnover rate), total trade quantity (total trade quantity), turnover (turnover), total volume (total volume), circulation market value (circulation market value). The dataset contains a total of 3701 entries from May 18, 2004, to June 3, 2020 (excluding 204 days of suspension in between). The data format is shown in Figure 9.

Close	High	Low	Open	Last	Change_Amount	Change_Percent	Turnover_Rate	Total_Trade_Quantity	Turnover	Total_Volume	Circulation_Market_Value
10.02	10.1	9.93	10.06	10.03	-0.01	-0.0997	0.6001	10138863	1E+08	16930110806	16930110806
10.03	10.24	10.01	10.13	10.1	-0.07	-0.6931	0.5922	10005585	1E+08	16947007125	16947007125
10.1	10.12	9.92	9.92	9.86	0.24	2.4341	0.599	10120644	1E+08	17065281352	17065281352
9.86	10.02	9.82	9.93	9.97	-0.11	-1.1033	0.4319	7297438	7.2E+07	16659769716	16659769716
9.97	10.07	9.85	9.92	9.92	0.05	0.504	0.3456	5839237	5.8E+07	16845629216	16845629216
9.92	10.14	9.83	9.9	9.93	-0.01	-0.1007	0.4273	7219491	7.2E+07	16761147625	16761147625
9.93	9.97	9.69	9.69	9.69	0.24	2.4768	0.4907	8290857	8.2E+07	16778043943	16778043943
9.69	10.02	9.6	9.97	9.94	-0.25	-2.5151	0.4191	7080591	6.9E+07	16372532307	16372532307
9.94	10.15	9.87	10.07	10.05	-0.11	-1.0945	0.5826	9843796	9.8E+07	16794940261	16794940261

Figure 9: Example of Data Format

4.2. Experimental Content and Results

In this study, various neural network models were trained using a training dataset, and then these models were tested for performance on a testing dataset. The predicted results of each model on the testing dataset are as follows (Figure 10-25):

4.2.1. Scheme 1: Prediction Results of Single-layer LSTM Model

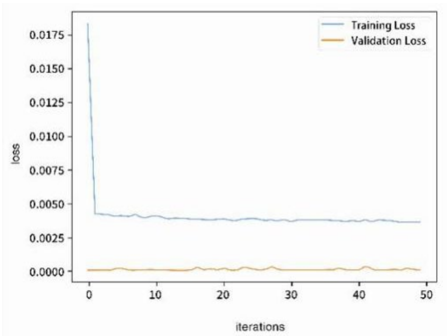

Figure 10: Training Loss and Validation Figure Loss Plot

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

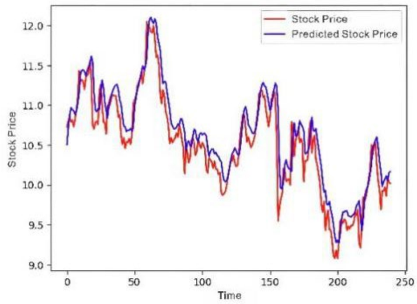

Figure 11: Prediction Plot of Predicted Values and True Values

4.2.2. Scheme 2: Prediction Results of Single-layer GRU Model

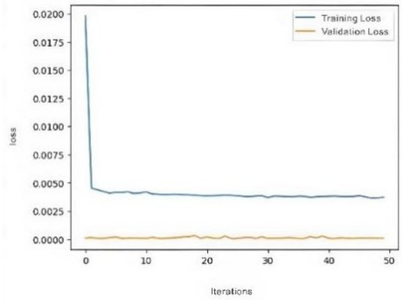

Figure 12: Training Loss and Validation Loss Plot

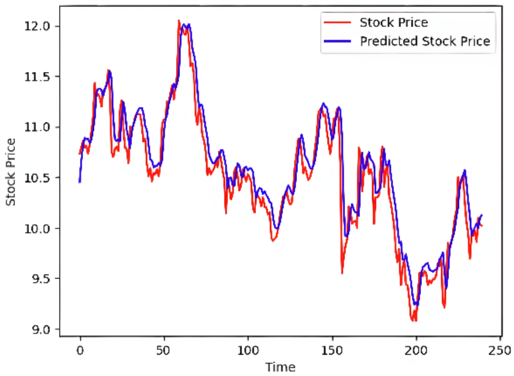

Figure 13: Prediction Plot of Predicted Values and True Values

4.2.3. Scheme 3: Prediction Results of Double-layer LSTM Model (D-LSTM)

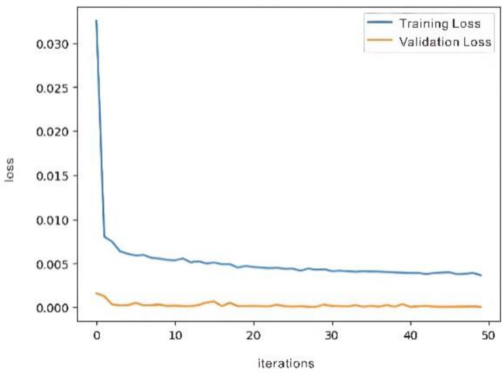

Figure 14: Training Loss and Validation Loss Plot

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

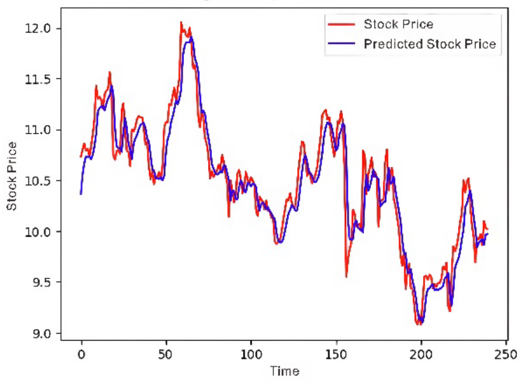

Figure 15: Prediction Plot of Predicted Values and True Values

4.2.4. Scheme 4: Prediction Results of Double-layer GRU Model (D-GRU)

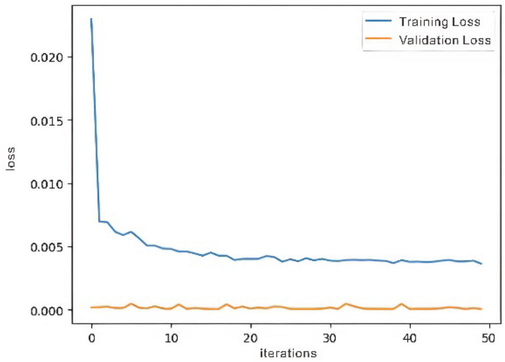

Figure 16: Training Loss and Validation Loss Plot

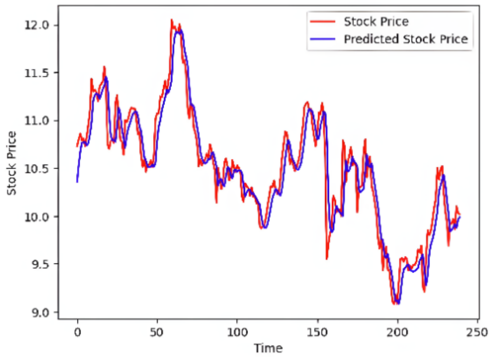

Figure 17: Prediction Plot of Predicted Values and True Values

4.2.5. Scheme 5: Prediction Results of Single-layer LSTM - Double-layer GRU Model (LSTM-D-GRU)

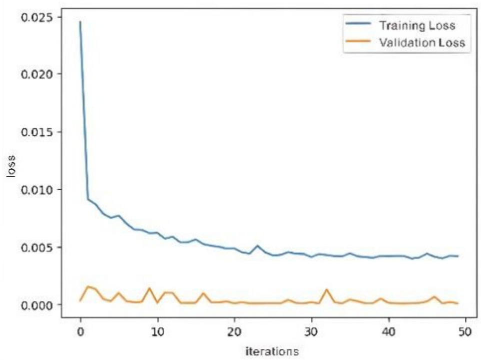

Figure 18: Training Loss and Validation Loss Plot

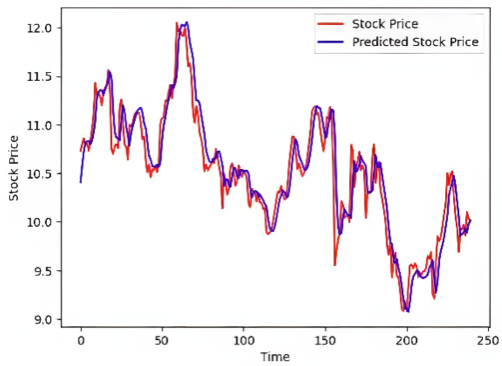

Figure 19: Prediction Plot of Predicted Values and True Values

4.2.6. Scheme 6: Prediction Results of Double-layer LSTM - Single-layer GRU Model (D-LSTM-GRU)

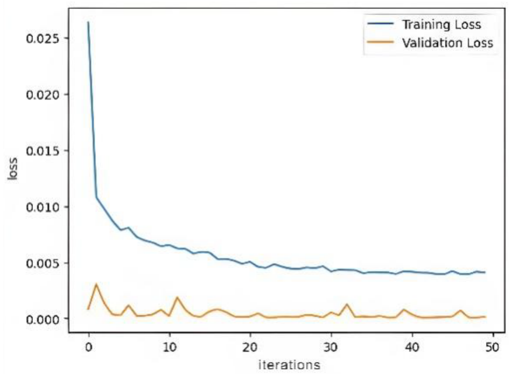

Figure 20: Training Loss and Validation Loss Plot

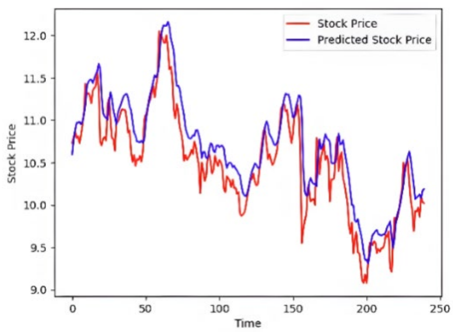

Figure 21: Prediction Plot of Predicted Values and True Values

4.2.7. Scheme 7: Prediction Results of Double-layer LSTM - Double-layer GRU Model (D-LSTM-D-GRU)

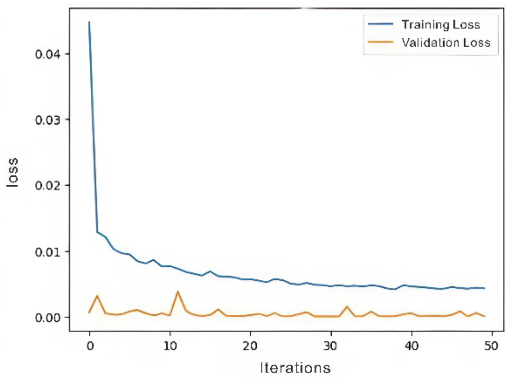

Figure 22: Training Loss and Validation Loss Plot

ISSN 2616-5775 Vol. 7, Issue 4: 80-89, DOI: 10.25236/AJCIS.2024.070411

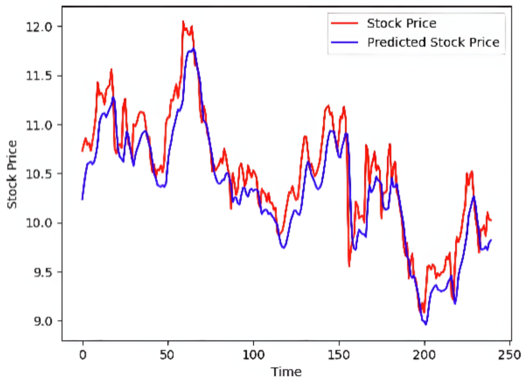

Figure 23: Prediction Plot of Predicted Values and True Values

4.2.8. Scheme 8: Prediction Results of GRU-LSTM Model

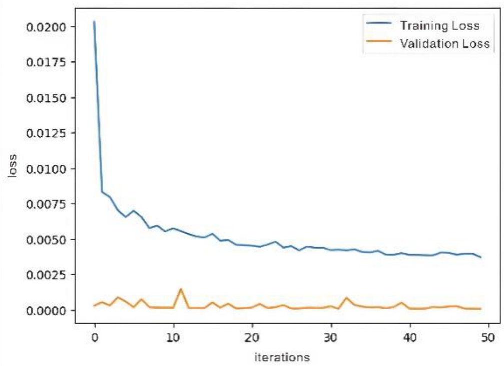

Figure 24: Training Loss and Validation Loss Plot

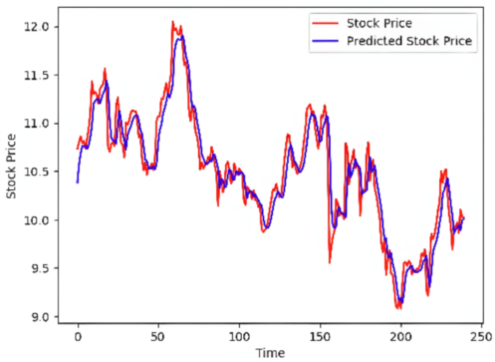

Figure 25: Prediction Plot of Predicted Values and True Values

4.3. Error analysis

Below is the error analysis for the eight different schemes (Table 1):

Table 1: Summary of Error Analysis

model structure	MSE	RMSE	MAE
LSTM	0.061469	0.247930	0.187994
GRU	0.048553	0.220347	0.158238
D-LSTM	0.045855	0.214139	0.158179
D-GRU	0.046794	0.216320	0.157026
LSTM-D-GRU	0.049578	0.222662	0.160625
D-LSTM-GRU	0.085074	0.291674	0.229323
D-LSTM-D-GRU	0.078444	0.280078	0.226964
LSTM-GRU	0.044019	0.209806	0.152030

This article compares the eight model combinations primarily using RMSE as the main metric, revealing that the LSTM-GRU neural network model has the smallest prediction error and highest accuracy in forecasting stock price changes.

Compared to the original models, the LSTM-GRU ensemble model has shown significant improvements in accuracy and stability. It is believed to be of great assistance to financial investors.

5. Conclusion

In this study, we used long short-term memory network (LSTM) and gated cycle unit (GRU) models to construct eight different configurations of prediction model combinations to perform experimental prediction. By comparing the prediction results with the actual data, the prediction errors are analyzed in detail and the related performance indexes are calculated. Further, this study compares the performance of eight models on key performance indicators such as mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE), and the results show that LSTM-GRU integrated neural network model has the best performance in stock price prediction. The experimental results show that the LSTM-GRU integrated model not only has significant advantages in improving the accuracy of stock price prediction, but also do well in improving the prediction performance. This finding provides an effective supplement to the existing prediction methods and provides an important reference for the follow-up research.

This article recognizes that the factors affecting stock price are extremely complex, including but not limited to 12 variables such as circulating “Open”, “Last”,“Turnover”, “Total Volume”, “Circulation” and so on. However, the current study does not take into account market sentiment, external political factors and other possible influencing factors. Therefore, future research should consider introducing a wider range of data sources, such as news reports, international trade data, etc., in order to use unsupervised learning methods to analyze and extract these data, so as to optimize the model to adapt to the new market environment and maximize the potential of LSTM-GRU integrated model in stock price prediction.

References

[1] Guś and J. Możaryn, "Comparison of artificial neural network adaptive control techniques for a nonlinear system with delay," 2023 27th International Conference on Methods and Models in Automation and Robotics (MMAR), Międzyzdroje, Poland, 2023, pp. 111-116, doi: 10.1109/MMAR58394. 2023. 10242473.

[2] Yani Hou. Stock Price Prediction Based on Bi-LSTM Deep Learning [J]. Statistics and Application, 2021, 10(3): 538-546.

[3] Chaobin Huang, Ximing Cheng. Research on stock price prediction based on LSTM neural network [J]. Journal of Beijing Information Science & Technology University, 2021, 36(1): 79-83.

[4] Zhen Liu, Huimin Wang, Siyv Hua, Yv Chen. Research on stock price prediction based on deep learning [J]. Science and Technology Innovation Herald, 2018, 13(247): 247-248.

[5] Ruixue Zhang, Yongtao Hao. Research on stock price prediction based on deep learning [J]. Computer Knowledge and Technology, 2023, 19(33):8-10. DOI:10.14004/j.cnki.ckt.2023.1739.

[6] Di Zhang. Research on Data Mining and Prediction Model Based on KNN and Neural Network Algorithm [J]. Journal of Taiyuan Normal University(Natural Science Edition),2023,22(02):29-34.

[7] Wenquan Cui, Qingfang Wang. Stock price prediction based on dual encoder using online social network information [J].Journal Of University Of Science And Technology Of China, 2020, 50(08):1093-1101.

[8] Shimin Huang. Stock price analysis and prediction based on ARIMA model -- taking China Merchants Bank as an example [J]. Management & Technology of SME, 2022, (11):184-187.

[9] Jinlei Han, Pingping Xiong, Jihong Sun. Stock price time series forecasting based on LSTM and grey model [J].Journal of Nanjing University of Information Science & Technology,2023,15(06):631-642.DOI:10.13878/j.cnki.jnuist.20221008002.

[10] Xiaojie Li, Huan Xia. Research on stock price regression prediction based on machine learning algorithm [J]. Science & Technology Information, 2023, 21(14):227-231. DOI:10.16661/j.cnki.1672-3791. 2211-5042-9348.

[11] Meng Li, Zhangjie Huang, Jianhui Xv. Stock price prediction based on LSTM-Wavelet model based on deep learning and Wavelet analysis [J]. J Chongqing Technol & Business Univ (Nat Sci Ed), 2023, 40(02):99-105. DOI:10.16055/j.issn.1672-058X.2023.0002.015.