Professor Wenjie LU, PhD(Corresponding author)
E-mail: lulu@hebust.edu.cn
Business School
Jiangsu Second Normal University
Institute of Science and Technology Development
Hebei University of Science and Technology
Jiazheng LI, Master
E-mail: 69880353@qq.com
School of Information Science and Engineering
Hebei University of Science and Technology
Professor Jingyang WANG, Master
E-mail: jingyangw@hebust.edu.cn
School of Information Science and Engineering
Hebei University of Science and Technology
Student Shaowen WU
E-mail: 2879058513@qq.com
School of Artificial Intelligence
Guilin University of Electronic Technology

A NOVEL MODEL FOR STOCK CLOSING PRICE PREDICTION USING CNN-ATTENTION-GRU-ATTENTION

Abstract. *Predicting stock price to avoid risk is the focus of stock research. A reliable predicting model could offer insights in stock price fluctuations and ultimately could provide the opportunity of gaining significant profits. In this paper, a new composite forecasting model is proposed to forecast the stock closing price of the next trading day. This model consists of three parts. Convolutional Neural Network (CNN) is used to collect: the factors that affect the stock price. attention mechanism (Attention) is used to compute the impact of stock data at different times on stock price. Gate Recurrent Unit (GRU) is used to forecast the stock price. It can make good time series prediction. Through comparison with CNN-Attention-LSTM-Attention, CNN-Attention-GRU, CNN-GRU-Attention and other traditional models. The experimental results indicate that the performance of this model is better to other models, and it has the best performance in evaluation metrics like MAE, RMSE and* R^2*. It is more appropriate for stock price prediction.*

Keywords*: stock price, prediction, Convolutional Neural Network, Attention, Gate Recurrent Unit.*

JEL Classification: O16 G12 C53

1. Introduction

The stock market is regarded as the barometer and weather vane of the current economic and financial activities of a country or a region. Predicting the

DOI: 10.24818/18423264/56.3.22.16

trend of stock price and grasping the changing law of stock market are always the research focus. With the help of accurate prediction technologies, regulators can timely monitor the existence of systemic risks and asset price bubbles in the stock market, which is conducive to the risk management of the stock market (Bukhar et al., 2020). At the same time, precise stock market forecast models can provide reasonable decision-making suggestions for investors. The forecast of stock price changes can also provide a reference for listed companies to make financing decisions, including but not limited to scheme selection, timing determination, cost reduction, uncertainty reduction, risk minimization, et al. Therefore, accurate forecast of the stock price is important to both regulators and investors.

Predicting stock prices has always been a worldwide problem. There are mainly two kinds of stock price prediction methods. One is the traditional econometric methods including regression analysis, time series analysis; the other is machine-learning methods (Oh, 2005; Ren et al., 2019). Early scholars usually used technical analysis methods based on statistics. Over the years, scholars have proposed a variety of traditional prediction methods based on statistics and probability theory, such as the value at risk (VAR) model, Autoregressive Integrated Moving Average Model(ARIMA), error correction model (ECM), and kalman filter model (KFM) (Bhardwaj and Swanson, 2006; Wang, 2015). However, most of these methods require data to satisfy some constraint before they can be used， What is more, the method has also been criticized for its poor accuracy and stability. The stock market is, essentially, a dynamic, non-stationary, and chaotic system with many influencing factors and various uncertainties. The rapid changes in stock price are often dynamic and non-linear. In order to make accurate predictions, it is necessary to ensure that the prediction method can deal with big enough time series data and a large amount of data with certain inductive ability. However, the traditional statistical methods do not have such characteristics. Therefore, the traditional statistical methods are not performing well in predicting stock price.

In recent years, more and more scholars have begun to try to use non-linear models to predict stock price. Given that machine learning methods can better deal with non-linear data, machine learning techniques such as logistic regression, decision tree and deep learning are widely used in financial data research. White (1998) applied neural networks to forecast IBM stock. The experimental results showed that neural networks were more suitable for stock price prediction, but it was not accurate enough. Zhang (2003) respectively used neural networks and ARIMA to predict stock. He indicated that neural networks had distinct advantages in non-linear data prediction. Nayak, Misra, and Behera (2017) used multi-layer perceptron (MLP) to forecast the stock index but the result was mediocre. Guo, Han, and Shen (2018) used SVR to forecast stock, but it was seldom applied to stock prediction because the parameters of SVR were difficult to determine. Nabipour, Nayyeri, and Jabani (2020) used long-short term memory (LSTM) to forecast stock price, and the results indicated that LSTM had relatively high accuracy in stock prediction. Sezer and Ozbayoglu (2020) applied convolutional neural network (CNN) to forecast stock price but the accuracy of time series prediction was low because CNN was widely used in feature extraction and image recognition. Gu, Wu, and Pang (2020) combined combination of gate recurrent unit (GRU) and Attention to forecast individual stocks. The results indicated that the GRU and attention was more feasible and effective than using attention or GRU alone.

Focused on the non-linearity, uncertain interaction, randomness, and other features of stock price and on the basis of fully considering the time correlation between the stock price data and the changing trend among the data, a CNN-Attention-GRU-Attention is proposed in this paper to forecast the stock closing price of the next trading day. First, CNN is applied to extract the features of the stock data. Then Attention is applied to compute the influence of different time states on the predicted value. Finally, GRU is applied to calculate the stock price. To prove the effectiveness of CNN-Attention-GRU-Attention proposed in this paper, the daily trading data of 6997 trading days of the Shanghai Composite Index (000001) from January 2, 1992, to August 31, 2020, is used as the experimental data. The data of 6497 trading days from January 2, 1992, to August 10, 2018 is used as the training set. Then, the model is applied to forecast the closing price of the Shanghai Composite Index from August 11, 2018, to August 31, 2020. At the same time, six models: CNN-Attention-LSTM-Attention, CNN-Attention-GRU, CNN-GRU-Attention, CNN-GRU, GRU, and LSTM are selected as comparative experimental models. The results indicate that CNN-Attention-GRU-Attention has higher forecast accuracy and stronger learning ability, and it has important practical value in stock price prediction.

In summary, the contributions of this paper are:

- Through the analysis of the stock data, it is found that the stock data followed time series. The CNN-Attention-GRU-Attention stock price prediction model is proposed.
- Given that the stock data follows time series, it is proposed to use GRU to predict stock price. GRU can avoid the vanishing gradient problem and the exploding gradient problem caused by RNN. And Attention is applied to compute the influence of data at different time states on prediction to improve the forecast accuracy.
- Taking the Shanghai Composite Index data as experimental data, comparative experiments are carried out. By comparing CNN-Attention-LSTM-Attention with six other stock price prediction models, the accuracy and efficiency of CNN-Attention-GRU-Attention is proved. It is more appropriate for stock price prediction.

2. Methodology

2.1 CNN

CNN was proposed by Lecun (1998). It can be effectively applied to feature extraction of data (Livieris et al., 2020). CNN can decrease the number of parameters through the local perception and weight sharing, thus improving the efficiency of model learning (Chung and Shin, 2020). Multiple convolution kernels constitute a convolution layer (Hao and Gao, 2020). The convolution layer extracts data features through convolution operation. However, the feature dimension after the convolution operation is relatively high, so it is necessary to use the pooling layer to settle this issue and decrease the cost of network training. In this paper, the convolutional layer and pooling layer of CNN model are used to extract the input stock trading data feature, abstract information to get feature data,

2.2 GRU

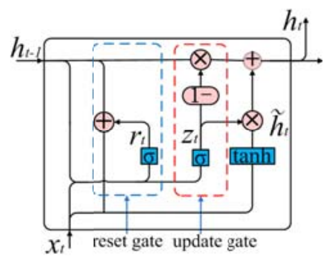

Figure 1. Structure of GRU

GRU was proposed by Cho (2014). It aims to solve the vanishing gradient and the exploding gradient problem in RNN (Ta et al., 2020). GRU is a good variant of RNN. It has a simpler structure and better effect than LSTM. Therefore, it is a very popular network at present. LSTM has the input gate, the forget gate, and the output gate. While GRU only has the update gate and the reset gate (Ergen and Kozat, 2020). The specific structure of GRU is shown in Figure 1. GRU can achieve results similar to LSTM, and it is easier to train than LSTM (Adriano et al., 2019). Therefore, GRU is preferred most of the time.

The calculation process of GRU are shown in formula (1)-(4):

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{1}$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{2}$$

$$\widetilde{h_t} = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t] + b_{\tilde{h}}) \tag{3}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h_t} \tag{4}$$

Where σ is the sigmoid activation function, W is the weight, b is the bias.

2.3 Attention

Attention was first proposed by Treisman (1980). Attention can focus on key information and ignore irrelevant information. It can solve the information overload problem. Attention can find the features of data at different times in time series. In this paper, the stock price changes with the changes of various factors, that is, the data at different time points in the time series have different effects on the predicted value of the stock. It is generally believed that the stock price characteristic information closer to the predicted value has a greater impact on the predicted value of the stock price. In this paper, Attention was used to calculate the degree of influence of different time points on the predicted stock value, and corresponding weights were assigned to each data item to make the model focus on important data items.

The calculation process of Attention are as follows:

Firstly, the similarity or correlation of features is computed. Secondly, the Softmax function is used to normalize the output value of the previous stage to get the attention score. Finally, the final output value is obtained according to the weight coefficient. The formulas are shown:

$$s_t = tanh(W_h k_t + b_h) \tag{5}$$

$$a_t = softmax(s_t) \tag{6}$$

$$o_t = \sum_t a_t k_t \tag{7}$$

Where W_h is the weight, b_h is the bias.

2.4 The Structure of CNN-Attention-GRU-Attention

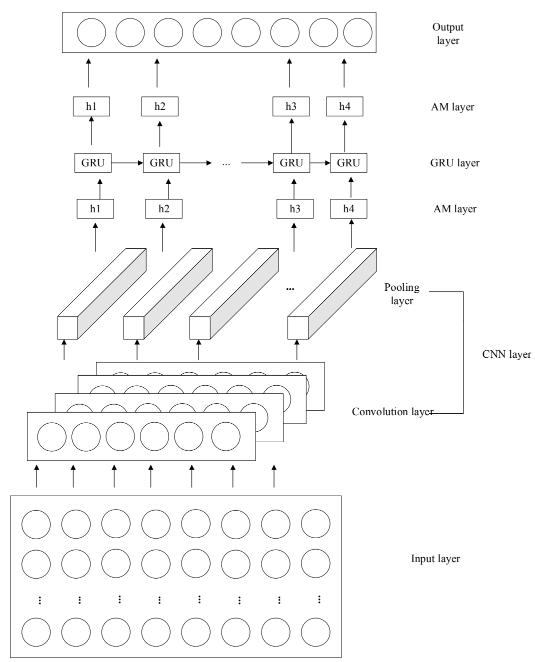

Figure 2. The Structure of CNN-Attention-GRU-Attention Diagram

CNN can extract features from data, so it is used in the feature extraction of data. On the time series. Attention can compute the influence of data at different times on the output data, so it is applied in time series prediction. GRU has a relatively similar performance compared with LSTM. It can learn better from time series data and is often used in time series prediction. Based on CNN, Attention, and GRU, a CNN-Attention-GRU-Attention is proposed to forecast stock price. Its main structure consists of CNN, Attention, and GRU. The model structure is displayed in Figure 2.

2.5 The Training Process of CNN-Attention-GRU-Attention

The training process of CNN-Attention-GRU-Attention is shown in Figure 3.

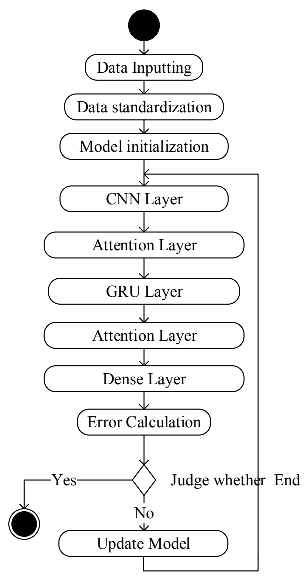

Figure 3. Training Process of CNN-Attention-GRU-Attention Activity Diagram

The steps are as follows:

(1) Data Inputting: Enter the historical characteristic data of the target stock.

(2) Data standardization: Because different dimension and dimensional unit existence big difference, the input stock data use the Z-score to standardize to ensure that in the training of the model each feature conforms to the same data distribution.

(3) Model initialization: Initialize the weight and bias of each layer of the CNN-attention-Gru-attention model.

(4) CNN Layer calculation and feature extraction : the standardized stock data in Step 2 is calculated through CNN to extract stock data time series features. The output data $output_1$ is obtained.

(5) Attention Layer: the data $output_1$ is inputted into the Attention layer. The attention distribution probability is calculated by comparing the input lag

data of trading days to capture the relationship between data and highlight the influence of the important characteristic data of trading days in the input, and output data $output_2$ is obtained

(6) GRU Layer: the data $output_2$ is inputted into the GRU layer. The output data $output_3$ is obtained by using its hidden layer to calculate the time series.

(7) Attention Layer: the data $output_3$ calculated by the GRU layer is input into the Attention layer again. The influence of data at different times on stock price is calculated again, and output data $output_4$ is got.

(8) Dense Layer: the data $output_4$ calculated by the Attention layer is input into the Dense layer, and the $output_5$ is got.

(9) Error Calculation and evaluation: the error between the $output_5$ calculated by the Dense layer and the real value of the corresponding stock data is calculated.

(10) Judge whether end: the end conditions are to complete the set epoch or the error of predictions is lower than a rated value. If one of the end conditions is met, the model training is ended; otherwise, continue to train the model according to step (11).

(11) Update model: the error is back propagated. The weight and bias of each layer are renewed. Then the whole CNN-Attention-GRU-Attention is updated. Jump to step (4) to continue training the model.

2.6 The Prediction Process of CNN-Attention-GRU-Attention on Stock Price

The precondition of stock price prediction based on CNN-Attention-GRU-Attention is that CNN-Attention-GRU-Attention has been trained. The prediction process of CNN-Attention-GRU-Attention is shown in Figure 4.

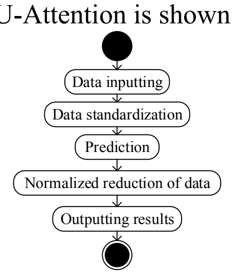

Figure 4. Prediction Process of CNN-Attention-GRU-Attention Activity Diagram

The above steps are as follows:

(1) Data inputting: Enter stock history trading data.

(2) Data standardization. The input stock data is standardized usingZ-score.

(3) Prediction: The data is input into CNN-Attention-GRU-Attention to be predicted.

(4) Normalized reduction of data: Since the data is predicted after standardization, the data predicted by CNN-Attention-GRU-Attention should be restored.
(5) Outputting results: Output the reduced data to finish the stock price prediction process.

3 Experiments

Under the same operating environment, CNN-Attention-GRU-Attention is compared with LSTM, GRU, CNN-GRU, CNN-Attention-GRU, CNN-GRU-Attention, CNN-Attention-LSTM-Attention, and CNN-Attention-GRU-Attention to prove the prediction accuracy. All models use Python and Keras based on Tensor Flow to implement. All the experiments use the same operating environment to carry out. To evaluate the prediction effectiveness of CNN-Attention-GRU-Attention, this paper uses the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error(MAPE), training time, and R^2 as the evaluation index.

3.1 Data description

The stock data selected in this experiment are the daily trading data of the Shanghai Composite Index (000001) for a total of 6997 trading days from January 2, 1992, to August 31, 2020. This data is got through the Wind. Select 8 technical indicators that influence stock price data. That is Opening Price(open) Highest Price (high) Lowest Price(low), Closing Price(closing) , Volume, Turnover, Ups and Downs, Change. Some data are shown in Table 1. The first 6,497 trading days of the stock price data were divided into a training set and the last 500 trading days into a test set

Table 1. Partial Original Stock Price Data

Date	Open	High	Low	Closing	Volume	Turnover	Ups and Downs	Change
1992/1/2	292.75	293.74	293.75	292.76	293.75	92800	1339598	1
1992/1/3	293.75	296.24	296.52	293.75	296.52	143600	1842222	2.77
1992/1/6	296.52	297.68	297.68	296.52	297.68	340000	4236024	1.16
1992/1/7	297.68	298.54	298.77	297.68	298.77	89200	1064008	1.09
1992/1/8	298.77	299.65	299.66	298.77	299.66	65900	930535	0.89

3.2 Model Implementation

In this experiment, the parameter settings of CNN-Attention-GRU-Attention are displayed in Table 2.

Table 2. The CNN-Attention-GRU-Attention Parameter settings

Parameters		Value
Convolution layer	Filters	64
	Kernel-size	1
	Padding	Valid
Pooling layer	Pool-size	1
	Padding	Valid

GRU layer	Number of hidden units	64
	Activation function	tanh
Training parameters	epoch	80
	loss function	MAE
	optimizer	Adam
	batch size	64
	time step	5
	learning rate	0.001

In this experiment, all models are trained with the same parameters. Each model is trained 100 times and the best training model is retained.

The data input by the model is a two-dimensional vector *(None, 5, 8)*. The vector first enters the convolutional layer to extract features and the vector *(None, 5, 64)* is also obtained. This vector is then entered into the pooling layer and the vector *(None, 5, 64)* is also obtained. This vector enters the Attention layer to capture important features, and an output vector *(None, 5, 64)* is obtained. The output vector is then entered into the GRU layer for calculation, and an output vector *(None, 5, 64)* is obtained. This output vector enters into the Attention layer to capture important features and also obtains an output vector *(None, 5, 64)*. Finally, the output value is obtained through the output layer.

3.3 Results

All models are trained 100 times. The best training model is retained. The test set is forecasted. The comparison between the predicted value and the real value is shown in Figures s 5-11.

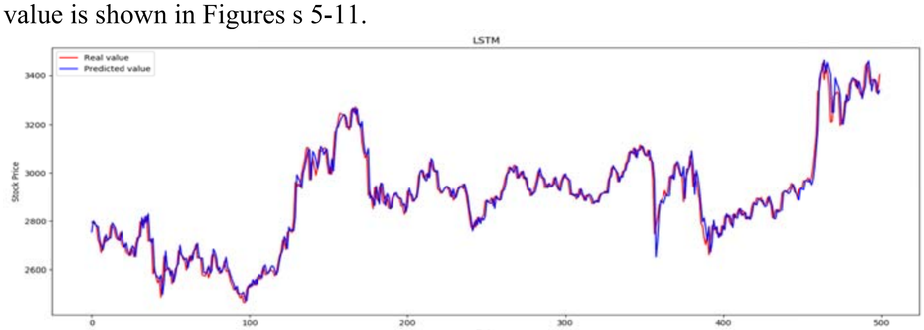

Figure 5. Real Value and LSTM Predicted Value

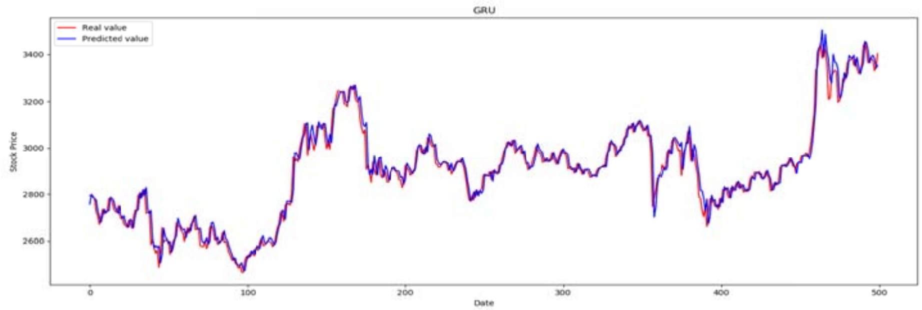

Figure 6. Real Value and GRU Predicted Value

Figure 7. Real Value and CNN-GRU Predicted Value

Figure 8. Real Value and CNN-GRU-Attention Predicted Value

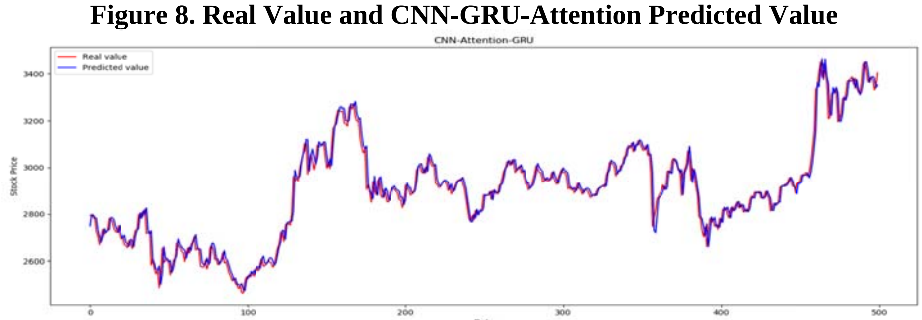

Figure 9. Real Value and CNN-Attention-GRU Predicted Value

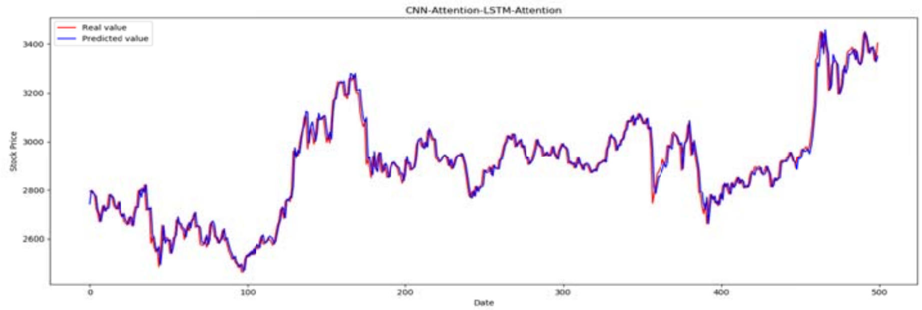

Figure 10. Real Value and CNN-Attention-LSTM-Attention Predicted Value

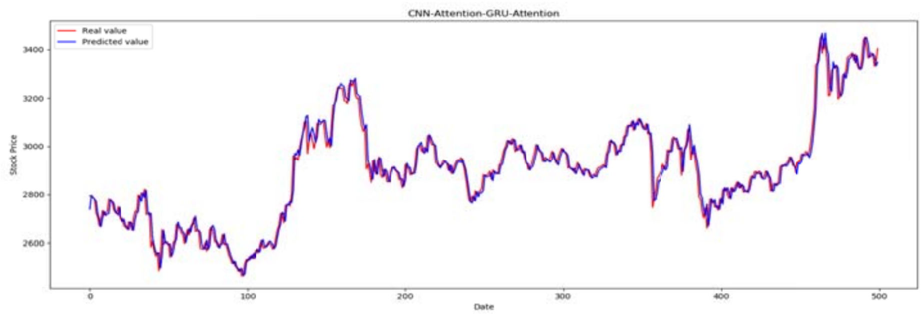

Figure 11. Real Value and CNN-Attention-GRU-Attention Predicted Value

Figures 5-11 shows that among the seven prediction models, the fitting degree of the real value and model-predicted value are ranked from high to low as CNN-Attention-GRU-Attention, CNN-Attention-LSTM-Attention, CNN-Attention-GRU, CNN-GRU-Attention, CNN-GRU, GRU, LSTM. CNN-Attention-GRU-Attention has the highest fitting degree and the two lines almost completely coincide. It shows that the model has relatively high accuracy in predicting stock price.

According to the prediction value of each model, the MAE, RMSE, R^2, MAPE can be calculated. The results are shown in Table 3.

Table 3.Coding modes for non-numerical

Method	MAE	RMSE	R^2	MAPE	Training time
LSTM	27.551	39.795	0.9644	0.9498	44.38
GRU	27.457	39.875	0.9642	0.9642	39.47
CNN-GRU	27.340	38.984	0.9658	0.9406	48.23
CNN-GRU-Attention	27.021	39.523	0.9649	0.9299	64.13
CNN-Attention-GRU	26.648	38.593	0.9665	0.9191	63.29
CNN-Attention-LSTM-Attention	26.567	38.449	0.9668	0.9153	83.15
CNN- Attention-GRU-Attention	26.266	38.266	0.9671	0.9050	80.13

3.4 Discussion

As shown in table 3, every evaluation index of CNN-Attention-GRU-Attention is the best. Its MAE, RMSE, MAPE is the smallest, and its R^2 is the biggest, indicating its highest fitting degree. With the increase of model complexity, the training time will be increased, but the increased time belongs to the normal range. The performance rank of all models from high to low is CNN-Attention-GRU-Attention, CNN-Attention-LSTM-Attention, CNN-Attention-GRU, CNN-GRU-Attention, CNN-GRU, GRU, LSTM. Compared with LSTM, the MAE, RMSE, MAPE and R^2 of GRU are respectively 27.457, 39.875, 0.9642% and 0.9642, while those of LSTM are respectively 27.551, 39.795, 0.9498% and 0.9644. The results show that LSTM and GRU have almost the same prediction accuracy, but the training time of GRU is obviously reduced by 5s compared with LSTM. Compared with GRU, in CNN-GRU formed by adding CNN in front of GRU, MAE reduces from 27.457 to 27.340. RMSE reduces from 39.875 to 38.984. MAPE reduces from 0.9642% to 0.9406% and R^2 adds from 0.9642 to 0.9658, indicating that the introduction of CNN can improve the forecasting accuracy. CNN can extract the features that affect the stock price. Compared with CNN-GRU, in CNN-Attention-GRU, MAE reduces by 2.5% from 27.340 to 26.648. RMSE reduces by 1% from 38.984 to 38.593. MAPE reduces from 0.9406% to 0.9191%. R^2 adds by 0.0007 from 0.9658 to 0.9665. The results indicate that attention can improve the forecasting accuracy. Attention captures the effect of stock data at different times on stock prices. Compared with CNN-Attention-GRU, every performance of CNN-Attention-GRU-Attention is improved. It is MAE decreases by 1.4% from 26.648 to 26.266. RMSE decreases by 0.8% from 38.593 to 38.266. MAPE decreases by 0.0141. R^2 increases by 1.5%. The results show that the introduction of double Attention is more effective. Compared with CNN-Attention-LSTM-Attention, the MSE of CNN-Attention-GRU-Attention reduces by 0.301 from 26.567 to 26.266; RMSE reduces by 0.183 from 38.449 to 38.266; MAPE reduces from 0.9153 to 0.9050; R^2 increases to 0.9671; the training time reduces from 83.15s to 80.13s. The results show that when using the composite model, GRU has a higher performance than LSTM. Therefore, CNN-Attention-GRU-Attention can make full use of the respective features of CNN, Attention, and GRU. CNN is applied to extract the features that affect the stock price. Attention is applied to compute the effect of stock data at different times on stock prices. GRU is applied for time series forecasting. The accuracy of stock price prediction is improved, which provides guidance for regulators and investors.

4 Conclusions and future research

In this article, a composite model (CNN-Attention-GRU-Attention) composed of CNN, Attention, and GRU is proposed to forecast the closing price of Shanghai Stock Composite index on the second trading day. In the model, to improve the forecasting effect, two aspects of characteristic screening and model structure are improved. CNN is applied to extract the features of the stock data.

Attention is applied to compute the influence of different time states on the predicted value. GRU is the main model to calculate the stock closing price. This paper uses the Shanghai Composite Index data to testify the accuracy of the prediction model and compares CNN-Attention-GRU-Attention with six models. Through comparison, it is concluded that CNN-Attention-GRU-Attention has the highest accuracy, and all its performance evaluation indicators are the best. The more complex compound model structure has higher accuracy than the single or simple compound model in stock price prediction. The proposal of CNN-Attention-GRU-Attention has broad application prospect and has great significance for regulators and investors to understand the stock market.

The future research will mainly have two aspects:

- Every parameter of CNN-Attention-GRU-Attention is to be adjusted to improve its prediction accuracy.
- To see if it is feasible to combine other neural networks and try to use the bidirectional GRU model to replace the GRU to improve the prediction accuracy of stock price.

REFERENCES

[1] **Adriano, G., Sugimoto, D.N., Tasinaffo, P.M., et al. (2019),** ***A New Approach to River Flow Forecasting: LSTM and GRU Multivariate Models.*** *IEEE Latin America Transactions*, 17 (12), 1978-1986;
[2] **Bhardwaj, G., Swanson, N.R. (2006),** ***An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series.*** *Journal of Econometrics*, 131 (1-2), 539-578;
[3] **Chung, H., Shin, K. (2020),** ***Genetic Algorithm-Optimized Multi-Channel Convolutional Neural Networks for Stock Market Prediction.*** *Neural Computing & Applications,* 32, 7897–7914;
[4] **Cho, K., Van, B., Gulcehre, C., et al. (2014),** ***Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation.*** *Computer Science*, 1724-1734;
[5] **Ergen, T., Kozat, S.S. (2020),** ***Unsupervised Anomaly Detection with LSTM Neural networks.*** *IEEE Transactions on Neural networks and Learning Systems*, 31 (8), 3127-3141;
[6] **Guo, Y., Han, S. & Shen, C., et al. (2018), A*n Adaptive SVR for High-Frequency Stock Price Forecasting.*** *IEEE Access*, 6, 11397-11404.
[7] **Gu, L., Wu, Y., Pang, J. (2020).** ***An Attention-based GRU Model for Stock Predicting.*** *Systems Engineering*, 30 (1), 134-140,
[8] **Hao, Y., Gao, Q. (2020),** ***Predicting the Trend of Stock Market Index Using the Hybrid Neural networks Based on Multiple Time Scale Feature Learning.*** *Applied Sciences*, 10 (11), 3961-3974;

[9] **Lecun, Y., Botou, L., Bengio, Y., Haffner, P. (1998),** ***Gradient-based Learning Applied to Document Recognition.*** *Proceedings of the IEEE*, 86 (11), 2278-2324;
[10] **Livieris, I., Pintelas, E., Pintelas, A. (2020),** ***A CNN–LSTM Model for Gold Price Time-series Forecasting.*** *Neural Computing & Applications*, 32, 17351–17360;
[11] **Nayak, S., Misra, B., Behera, H. (2017),** ***Artificial Chemical Reaction Optimization of Neural Networks for Efficient Prediction of Stock Market Indices.*** *Ain Shams Engineering Journal*, 8 (3), 371-390;
[12] **Nabipour, M., Nayyeri, P., Jabani, H. et al. (2020),** ***Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms via Continuous and Binary Data; A Comparative Analysis.*** *IEEE Access*, 8, 150199-150212;
[13] **Oh, L.H.S. (2005),** ***Technological Forecasting at the Korean Stock Market: A Dynamic Competition Analysis Using Lotka–Volterra Model.*** *Technological Forecasting and Social Change,* 72 (8), 1044-1057;
[14] **Ren, R., Wu, D.D. & Liu, T. (2019),** ***Forecasting Stock Market Movement Direction Using Sentiment Analysis and Support Vector Machine.*** *IEEE Systems Journal*, 13 (1), 760-770;
[15] **Sezer, O. B., Ozbayoglu, A.M. (2020),** ***Financial Trading Model with Stock Bar Chart Image Time Series with Deep Convolutional Neural Networks.*** *Intelligent Automation & Soft Computing*, 26 (2), 323–334;
[16] **Ta, V., Liu, C., Tadesse, D. (2020),** ***Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading.*** *Applied Sciences*, 10 (2), 437-456;
[17] **Treisman, A., Gelade, G. (1980),** ***A Feature-integration Theory of Attention.*** *Cognitive psychology*, 12 (1), 97-146;
[18] **Wang, L. (2015),** ***Dynamical Models of Stock Prices Based on Technical Trading Rules—Part III: Application to Hong Kong Stocks.*** *IEEE Transactions on Fuzzy Systems*, 23 (5), 1680-1697;
[19] **White, H. (1998),** ***Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns.*** *Earth Surface Processes & Landforms*, 8 (5), 409-422;
[20] **Yu, P., Yan, X. (2020),** ***Stock Price Prediction Based on Deep Neural Networks.*** *Neural Computing & Applications*, 32, 1609–1628;
[21] **Zhang, G. (2003),** ***Time Series Forecasting Using a Hybrid ARIMA and Neural Networks Model.*** *Neurocomputing*, 50, 159-175.