Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 3
CECS 2023

Research on Stock Price Prediction and Quantitative Stock Picking Strategy Based on Deep Learning

Jiahao Ji*
School of Business, Nanjing Xiaozhuang University, Nanjing, China
*Corresponding author: 15162342136@163.com

Abstract. With the continuous development of the domestic stock market and the continuous improvement of the financial system system, and at the same time, the domestic stock market gradually rises in the financial system, based on the prediction research of the domestic stock market will become more and more important. In order to solve the problems of low precision and poor accuracy of short-term stock price prediction, this paper selects the bi-directional long- and short-term memory network of attention mechanism (WOA-BiLSTM-Attenion) model under the whale optimization algorithm for stock price prediction. The modeling of bi-directional long- and short-term memory network with attention mechanism can reduce the loss of historical information and increase the influence of important information. On this basis, Whale Optimization Algorithm (WOA) is then used for hyperparameter selection to reduce human interference. The experimental results show that compared with BP, LSTM, BiLSTM, BiLSTM-Attention, the WOA-BiLSTM-Attenion model has a better effect on stock closing price prediction, with a value of 13.9446, and the value of 0.9477, which has a higher accuracy, with a view to providing certain reference for the prediction research in other fields.

Keywords: Deep Learning; Whale Optimization Algorithm; Attention Mechanism; Stock Price Prediction.

1. Introduction

In recent years, with the rapid development of social economy, the number of listed companies is increasing, stocks have become one of the hot topics in the financial field. On the one hand, the trend of stock prices will, to a certain extent, determine the direction of many economic behaviors, so the prediction of stock prices is also concerned by more and more financial investors and financial analysts; on the other hand, with the number of investors in the stock market increasing year by year, only by accurately analyzing the trend of future changes in the price of the stock, in order to be one step ahead of others to grasp the market trends, and to obtain more investment income. The stock market conducts high-frequency trading every day, which generates a large amount of stock-related data. But in fact, the huge amount of data is abstract to investors, how to utilize effective methods to extract valuable information from the huge amount of stock data has become a problem that needs to be solved nowadays. In the traditional quantitative investment field, the selection of target stocks and the prediction of stock prices are mostly based on the results of long-term stock market experience, and empirical stock analysis methods are often poor in risk resistance and long-term prediction ability, and are not easy to disseminate and promote, in addition to their analysis speed is often slow. Immediately appeared is based on statistics and finance stock analysis methods, is also the beginning of the stock mathematical modeling, such as autoregressive model, stochastic volatility model and Markov model, etc., this kind of method of predictive analysis is better than the empirical method, in addition, because of the use of mathematical modeling, is very suitable for the use of computers to carry out analysis, but these models are based on the establishment of a small amount of input data, and can not be applied to the current large-scale data scenarios. However, these models are based on a small amount of input data and are not applicable to the current large-scale data scenarios. Therefore, designing an effective stock price prediction model has gradually become a hot research topic in this field.

2. Literature Review

At present, many research scholars at home and abroad have proposed various forecasting methods, such as VAR (Vector Autoregressive Model), ARM (Autoregressive Sliding Average Model), Exponential Smoothing Model based on statistics and probabilistic theory, GM, SVM based on non-statistical principles, as well as ANN Innovative Forecasting Model, Gray Forecasting Method, and Artificial Neural Network Method. In later years, research centers proved that ARIMA algorithm is the most effective for short-term trend prediction of stocks [1].

In 1988 White first used BP neural network to forecast the stock return of IBM Corporation, but did not achieve the desired results [2]. Over time, more and more studies have shown the usefulness of artificial neural networks in stock price prediction. 2008 Senol D et al. proved that artificial neural network models have higher prediction accuracy by comparing the model performance of artificial neural network models and logistic regression models for stock prediction problems [3]. And artificial neural networks are usually combined with other methods to predict the stock market, for example, in 2016 Yaqub M U et al. combined ARMA model and BP neural network multilayer perceptron and other models to conduct prediction research on the Dow Jones index, and constructed a hybrid model that achieved excellent performance [4].

However, with the development of computer technology, some studies have shown that artificial neural networks may not be suitable for stock prediction, so many studies have begun to start from the deep architecture, so machine learning-based techniques are gradually applied to the problem of financial time series data, and in the study of deep learning algorithms, this study has been a series of research methods have been proposed by experts and scholars.Bao W et al. proposed for the first time to extract the deep features of the stock into stock price prediction, adding wavelet transform to LSTM model and orthogonal decomposition of stock price features at the same time to improve the prediction accuracy, and the results proved that the method is effective [5].Akita R et al. proposed paragraph vector and long and short-term memory model and verified the effectiveness of the method by analyzing the actual data of 50 listed companies in Tokyo Stock Exchange [6].Lin J et al. used symbol aggregation approximation (SAX) to identify similar short-term price movements [7]. And Nair Binoy B et al. improved it by clustering stock price data with self-organizing feature mapping [8]. And in 2022 Gao W et al. proposed to model the stock market nearly based on MFNN and IBHA novel model and analyzed two data sets of Shanghai stock market in 2015.9-2018.8, which finally showed that the calculation results of the new method match the real data better and the prediction accuracy is higher [9].

In summary, scholars at home and abroad have already made many contributions to the problem of stock trend prediction, and most of these studies are based on the analysis of foreign stock markets due to the late start of the domestic stock market. However, with the continuous development of the domestic stock market and the continuous improvement of the financial system system, and at the same time the position of the domestic stock market in the financial system is gradually rising, the research based on the prediction of the domestic stock market will become more and more important. In this paper, we collect the opening price, closing price, high price, low price, previous closing price, increase/decrease amount, and volume of Ping An Bank stock from January 1, 2010 to January 15, 2024 based on Python's tushare financial data interface package, and verify the prediction results based on Attention by comparing them with the prediction results of BP, LSTM, BiLSTM, and BiLSTM-Attention models. By comparing with the prediction results of BP, LSTM, BiLSTM, BiLSTM-Attention and other models, the superiority of optimizing WOA-BiLSTM model based on Attention mechanism is verified. The results show that the WOA-BiLSTM model based on Attention mechanism proposed in this paper has higher prediction accuracy compared with other prediction models.

3. Research Methods

3.1. WOA Whale Optimization Algorithm

The Whale Optimization Algorithm (WOA) is a novel, nature-inspired meta-heuristic optimization algorithm with few optimization parameters, and simple operations, and easy to master. The WOA is an algorithm based on whale hunting behavior and simulates the 'spiral bubble net' strategy of humpback whales (shown in Figure 1). There are two main behavioral patterns in humpback whale hunting: random search and narrow search.

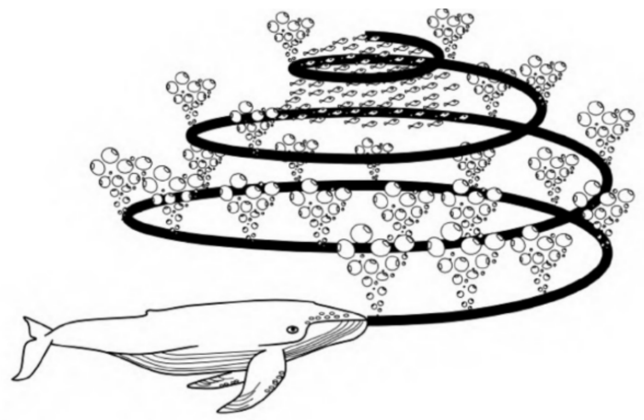

Figure 1. Humpback whale 'spiral bubble net' hunting strategy

The random finding can also be understood as finding the solution to a particular puzzle, which can be converted into the expressions

$$D = |CX_{rand}(t) - X(t)| \tag{1}$$

$$X(t+1) = X_{rand} - AD \tag{2}$$

In equations (1)-(2), theX_{rand} is a vector of randomly selected locations in the current whale population containing feasible solutions; X is a vector of areas where individuals are located; t is the current number of iterations; D is the distance at enclosure and AC are coefficients.

Narrowing down the search can be seen as an extrapolation procedure after approaching a valid solution, and the position will be updated when the most valid current solution to the problem is determined, expressed as follows.

$$D = |X(t) - CX_{best}(t)| \tag{3}$$

$$X(t+1) = X_{best} - AD \tag{4}$$

In equations (3)-(4), theX_{best} is the best solution produced by each iteration. By assigning equal probabilities to the two search methods to simulate the real behavior pattern of humpback whales, the search is judged to be over when the maximum number of iterations is reached.

3.2. Long Short-Term Memory Network

LSTM consists of three control gates, i.e., input gate, forgetting gate, and output gate, and is able to solve the gradient vanishing and gradient explosion problems of RNNs during long sequence training. The forget gate, which determines how much the information in the memory unit is forgotten.

$$\mathrm{f_t} = \sigma(\mathrm{W_f} \bullet [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_f}) \tag{5}$$

By the sigmoid layer of the input door, the value of the update is determined. After that, the tanh layer will generate a new candidate value vector, and the combination can be used to generate a update status value.

$$\mathrm{i_t} = \sigma(\mathrm{W_i} \bullet [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_i}) \tag{6}$$

$$\tilde{C}_t = \tanh(W_C \bullet [h_{t-1}, x_t] + b_C) \tag{7}$$

Multiply the old state by the discarding information defined by the forget gate, forget the information that decides to discard, and add the new candidate value to get the updated cell.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{8}$$

Finally, the output part is determined by the sigmoid layer of the output gate, and the neuron state is passed through the layer and multiplied by the output of the sigmoid threshold to get the result we expect.

$$o_t = \sigma(Wo \bullet [h_{t-1}, x_t] + bo) \tag{9}$$

$$h_t = ot * \tanh(C_t) \tag{10}$$

3.3. Bidirectional Long and Short-term Memory Neural Network

Although LSTM makes up for the defects of RNN, LSTM can only utilize the information of the past data, but ignores the future information of the data, but the future information of the data also plays a very important role in the construction of the model. So there is a bidirectional long and short-term memory neural network (BiLSTM), BiLSTM is an extension of the LSTM model, which consists of two LSTM neural networks of forward sequence and backward sequence, and these two LSTM neural networks are connected to the input layer and the output layer, and the results are shown in Figure 2. The application of BiLSTM can further explore the intrinsic connection between the current stock price data and the price data of the past and future moments, and further improve the prediction accuracy of the model.

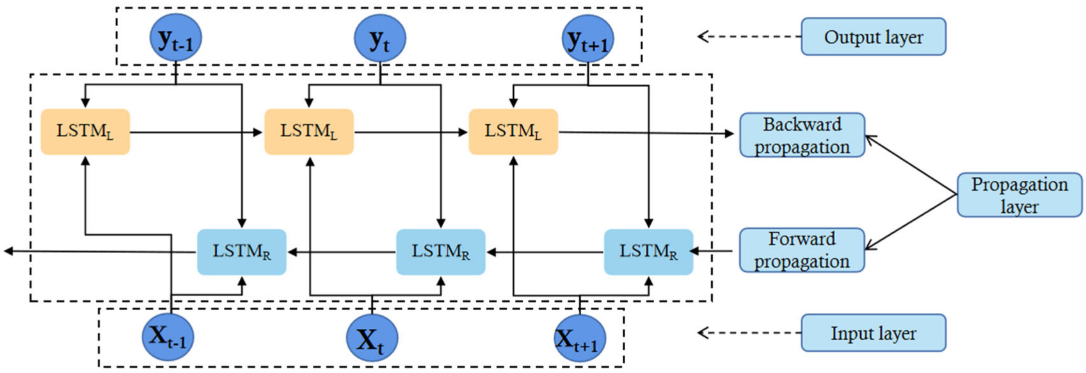

Figure 2. Model of BiLSTM

3.4. Attention Mechanism

Attention mechanism is a technique commonly used in computer science and machine learning, in traditional neural networks, the output of each neuron depends only on the output of all the neurons in the previous layer, by introducing the attention mechanism, the output of each neuron can also be weighted according to the different parts of the input data, i.e., different weights are given to different parts. In this paper, different weights are given to the implicit states of BiLSTM by mapping the weighting and learning parameter matrices, the structure of which is shown in Figure 3.

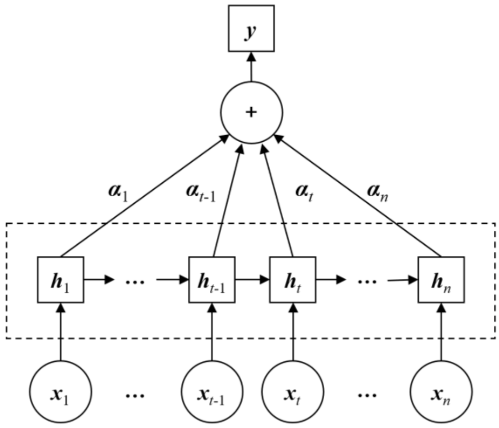

Figure 3. Structure of the Attention mechanism

4. Research Analysis

4.1. Data Set Selection and Data Preprocessing

In this paper, an actual stock price dataset is used to experiment with the model, and modeling methods such as BP, LSTM, BiLSTM, BiLSTM-Attention are selected to compare the results with the method proposed in this paper. In this paper, the opening price, closing price, high price, low price, previous closing price, increase/decrease amount, and volume of Ping An Bank stock from January 1, 2010 to January 15, 2024 are collected based on Python's tushare financial data interface package. The multivariate distribution of this stock dataset is shown in Figure 4.

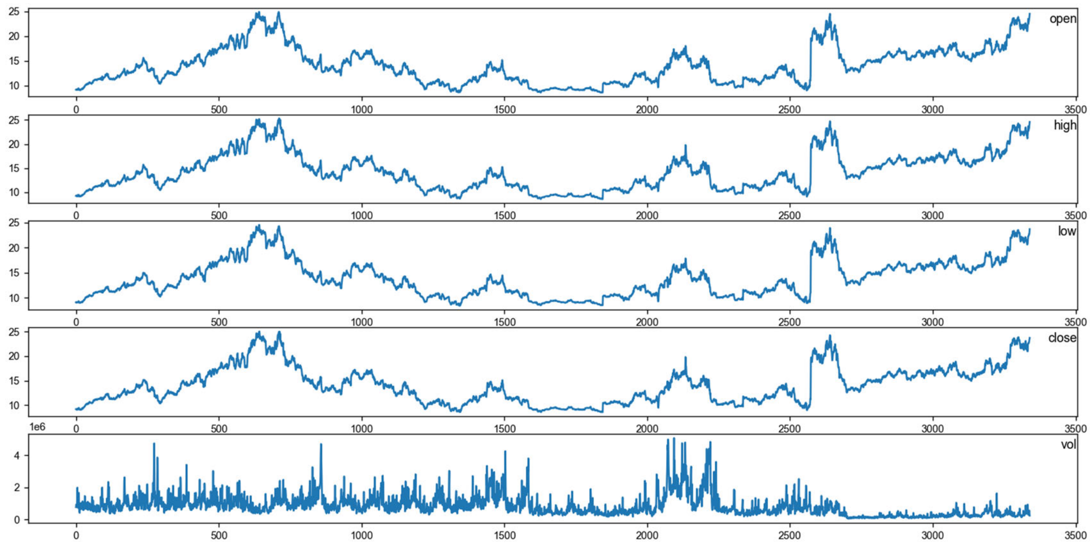

Figure 4. Multivariate distribution of the data set

A total of 2540 data from January 1, 2010 to September 24, 2020 were selected as the training set, and a total of 800 data from September 25, 2020 at 22:00:00 to January 15, 2024 were selected as the test set, and the results of the division between the training and test sets are shown in Figure 5.

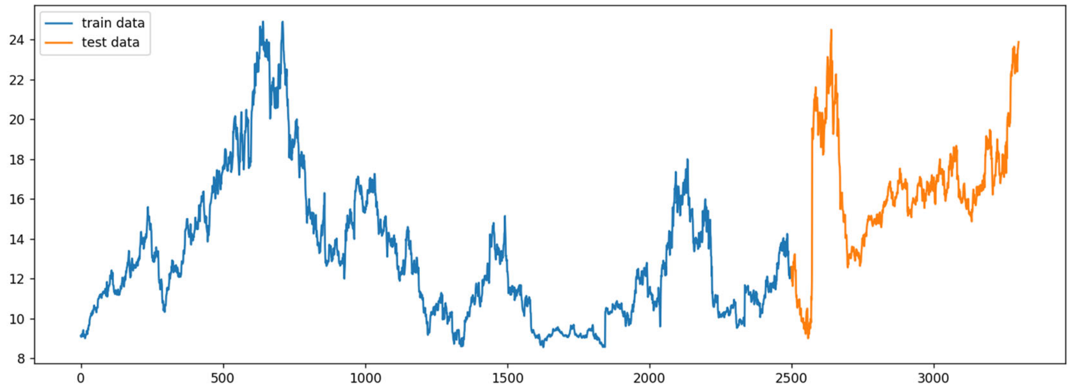

Figure 5. Division of the training set and dataset

The stock price data were normalized to between [0,1] using the normalization method of Max-Min Normalization (Min-Max Normalization) with the following formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{11}$$

4.2. Model Evaluation Indicators

In order to compare the prediction effect of different models, this paper adopts the evaluation indexes that are often used in stock price prediction, which are Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2), which are calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{13}$$

where: N is the number of samples. $\hat{y}_i$ is the predicted value of the ith sample.y_i is the actual value of the ith sample. $\bar{y}$ is the mean of the N sample data.

4.3. Model Construction and Experimentation

The preprocessed data divided into features and labels are input into the attention layer separately to get the weighted feature vector. At the same time, the hyperparameters of BiLSTM model are optimized by WOA, including learning rate, batch size, maximum iteration number, number of nodes in two-layer BiLSTM, number of nodes in fully-connected layer, and then the weighted training set data are inputted into the BiLSTM-Attention model, so as to construct the WOA-BiLSTM-Attention model, and the results of optimized hyperparameters are as follows: the batch size is 134, the number of units in two hidden layers is 12 and 8, the maximum number of units in one hidden layer is 12 and 8 respectively, and the maximum number of iterations is 20, and the learning rate is 0.00972358434.

When comparing with other models in the experiment, in which for BP, the maximum number of iterations is 10, there is a hidden layer containing 300 neurons; for LSTM, a two-layer structure with the number of hidden units of 10 each, the batch size is 100, the maximum number of iterations is taken as 50, and the learning rate is 0.01; and for BiLSTM, whose number of units in the two layers of the hidden layer is 20 and 10, the batch size is 100, the maximum number of iterations is taken as 20, and the learning rate is 0.00972358434. The maximum number of iterations is 20 and the learning rate is 0.001; for BiLSTM-Attention, the number of units in both hidden layers is 10, the batch size is 100, the maximum number of iterations is 20 and the learning rate is 0.001. Finally, the test set data is fed into the trained model and the results are back-normalized and output to get the stock price prediction results.

4.4. Results of the Experiment

Therefore, to address the unstable prediction results of the BiLSTM-Attention model, the WOA algorithm is chosen to optimize the network hyperparameters to improve the prediction stability and accuracy of the model. The WOA-BiLSTM- Attention model is used for training and prediction, and the results are shown in Figure 6. From the figure, it can be seen that the model fits the up and down fluctuations of stock prices and the overall trend more accurately, and it can obviously determine the highest value of stock prices in the short term. Figure 7 shows the loss profile of the whale algorithm for optimizing the deep network model as the loss varies with the number of iterations.

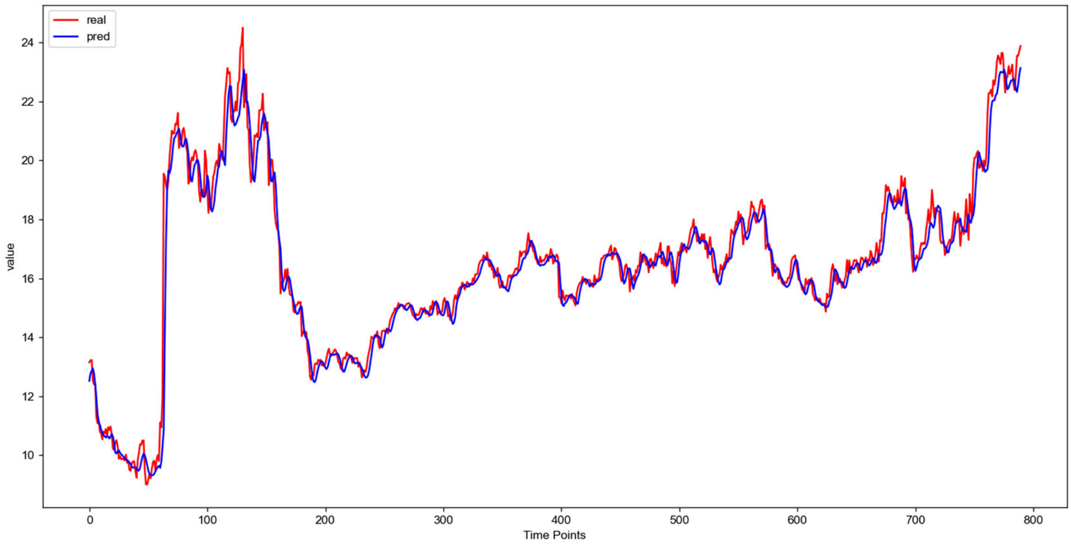

Figure 6. Results of comparison of real and predicted values

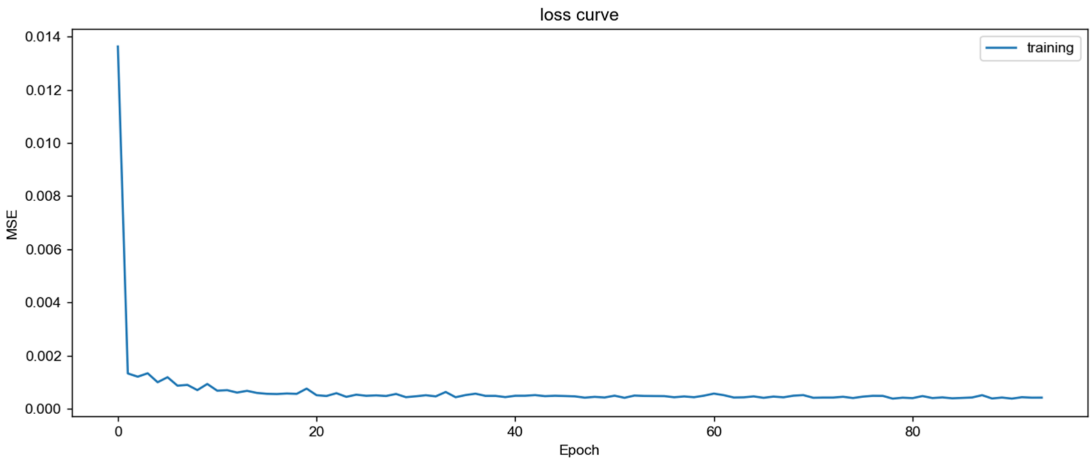

Figure 7. Loss curve diagram

The methodology of this paper has been improved in and indicators, which indicates that the stock price prediction model proposed in this paper has improved the overall prediction accuracy and model performance in the prediction process.

Table 1. Comparison of forecasting accuracy in different models

Models	$RMSE$	R^2
BP	74.4799	0.9119
LSTM	55.3879	0.9155
BiLSTM	36.3151	0.9188
BiLSTM-Attention	21.5360	0.9412
WOA-BiLSTM-Attention	13.9446	0.9477

As can be seen from Figure 8 and Table 1, the WOA-BiLSTM-Attention prediction model has a higher prediction accuracy and better fitting effect than the other four models. Especially in the valley value of stock price change, this paper's method predicts better, can more accurately capture the law of change of stock price, and at the same time can better predict the trend of change to improve the prediction accuracy.

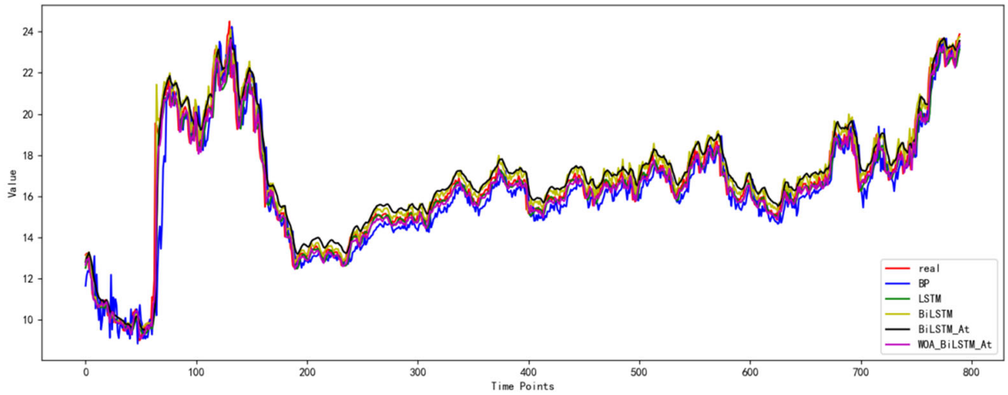

Figure 8. Plot of predicted results and true values for each model

5. Discussion and Conclusion

In order to solve the complex stock prediction problem, this paper proposes to use the whale algorithm to optimize the stock price prediction method of BiLSTM-Attention, and establishes the WOA-BiLSTM-Attention model. From the simulation results and process analysis, it can be seen that:

(1) The whale algorithm can well optimize the hyperparameter selection problem of BiLSTM-Attention model. It improves the global optimization-seeking ability of BiLSTM, reduces human interference, reduces the randomness of the quality of the prediction results and the prediction cost caused by the perceived setting parameters, and improves the accuracy of the model;

(2) The experimental results show that compared with BP, LSTM, BiLSTM, BiLSTM-Attention and other network models, the proposed model in this paper has higher prediction accuracy on the dataset, the interval coverage are higher, and the generalization ability of the model is stronger.

References

[1] Ariyo A A, Adewumi A O, Ayo C K. Stock price prediction using the ARIMA model[C]//2014 UKSim-AMSS 16th international conference on computer modelling and simulation. IEEE, 2014: 106-112.

[2] White H. Economic prediction using neural networks: The case of IBM daily stock returns[C]//ICNN. 1988, 2: 451-458.

[3] Senol D, Ozturan M. Stock price direction prediction using artificial neural network approach: The case of Turkey[J].Journal of Artificial Intelligence, 2009,1(2):70-77.

[4] Yaqub M U, Al-Ahmadi M S. Application of combined ARMA-neural network models to predict stock prices[C]//Proceedings of the The 3rd Multidisciplinary International Social Networks Conference on SocialInformatics 2016, Data Science 2016. 2016: 1-5.

[5] Bao W, Yue J, Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory[J]. PloS one, 2017, 12(7): e0180944.

[6] Akita R, Yoshihara A, Matsubara T, et al. Deep learning for stock prediction using numerical and textual information[C]//2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS). IEEE, 2016: 1-6.

[7] Lin J, Keogh E, Wei L, et al. Experiencing SAX: a novel symbolic representation of time series[J]. Data Mining and knowledge discovery, 2007, 15: 107-144.

[8] Nair Binoy B, Mohandas V P, Sakthivel N R. A genetic algorithm optimized decision tree-SVM based stock market trend prediction system[J]. International journal on computer science and engineering, 2010, 2(9): 2981-2988.

[9] Gao W. Modeling stock market using new hybrid intelligent method based on MFNN and IBHA[J]. Soft Computing, 2022, 26(15): 7317-7337.