Stock Price Prediction Based on Spatio-Temporal Coupling with Deep Learning

Heming Lai[1, *, #], Chuyi Yong[2, #] and Qinxin Wang[2, #]

[1]School of Finance Mathematics and Statistics, Guangdong University of Finance, Guangzhou, China

[2]School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai, China

*Corresponding author: Ming944058632@163.com

[#]These authors contributed equally

Abstract. Stock price prediction is a nonlinear dynamic problem, and the stock price is susceptible to its autocorrelation and inertia effect, as well as other stock price fluctuation on the same plate. Traditional Autoregressive Integrated Moving Average Model (ARIMA) only builds a linear prediction model, but the neural network model has strong nonlinear modeling ability. In this paper, we propose a Convolutional neural networks with Long short-term memory (CNN-LSTM) method to predict stock price fluctuations. This is because the selective memory advanced deep learning function of LSTM is used to deeply mine the internal rules of time series information, and the convolution in CNN is used to integrate the original stock data to extract the relationship between features of different variables. Finally, the feasibility of the method and the model's applicability are analyzed by comparing with the results of other prediction models and a conclusion is drawn. The results show that, compared with the prediction model based on time series alone, the model has a significant accuracy advantage. In addition, the hybrid model can better help investors make decisions, expand returns, and avoid risks.

Keywords: Stock Price Prediction; Spatiotemporal Sequence Prediction; Long Short-Term Memory; Convolutional Neural.

1. Introduction

The stock price fluctuation inefficient market is erratic and difficult to predict because it is affected by many factors. An accurate method of stock price forecast can provide a reference for the investors to get higher returns with less risk [1]. Thus, stock price forecast has always been a major topic of discussion in the public and financial industry.

The time series model can infer the possibility, trend, and regularity of future changes by analyzing the past changes of a group of time series. Box et al. introduced the autoregressive integrated moving average (ARIMA) models. With further research, in short-term financial time series forecasting, the ARIMA model is known to be robust and efficient compared to the popular ANNs model. Through real data analysis, the ARIMA model has the potential to help investors make decisions in short-term stock price forecasting [2]. The ARIMA model uses the differencing method to convert non-stationary time series into stationary time series. Then the dependent variable is only regressive to its lag value, the present value, and the lag value of the random error term. The ARIMA model can predict prices solely by analyzing endogenous variables. However, the ARIMA model requires time series data to be stationary or stable with differencing. Besides, it essentially only captures linear relationships. From this foundation, the threshold autoregressive model (TAR) was born. TAR introduces threshold values in the range of time series, divides the time axis into several intervals, and explains the whole system with different AR models in different intervals. TAR model can effectively solve the unstable situation of timing data and achieve great results [3]. Meanwhile, some models such as Autoregressive Conditional Heteroskedasticity (ARCH), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and Threshold Generalized Autoregressive Conditional

Heteroskedasticity (TGARCH) are proposed to explain the general rule of stock price change. But most of these methods are based on single variables and do not consider the influence of other factors.

Because of the shortage of the time series model based on single variables, many machine learning models are approaching to solve the problem with many uncertainties involved and many variables that influence the market value on a particular day. In the literature survey, we found that some machine learning models have reached quite an effective prediction of the stock value. Osman Hegazy et al. proposed a model which uses Particle swarm optimization (PSO) and least square support vector machine (LS-SVM) and compared their performance with the single LS-SVM and ANN-BP algorithm [4]. The PSO algorithm is employed to select the best free parameters combination for LS-SVM to avoid over-fitting and local minima problems and improve prediction accuracy. The proposed model used the dataset from many companies covering all stock sectors, including Information Technology, Financials, Health Care, Energy, Communications, Materials, and Industrials in the S&P 500 stock market. And it was found that the performance of the PSO-LS-SVM algorithm is better than LS-SVM and ANN-BP algorithms. Mr. Rupesh A. Kamble has considered optimizing the stock price trend prediction for the short term using Random Forest with some Technical indicators such as Moving Average Convergence Divergence (MACD), the Relative Strength Index (RSI), the Stochastic Oscillator (KDJ), and Bollinger Band (BB) [5]. Meanwhile, he used J48 Algorithm along with Technical data and Fundamental data (such as its Debt to Equity, Net profit of previous three years, Promoters holding, Dividend yield and PE ratio) to predict the long-term trend. Treating the forecasting problem as a classification problem, Luckyson Khaidem et al. have used a random forest classifier to build the predictive model, and the proposed model has produced effective results [6]. However, the fluctuations of the stock market are dynamic and complex, so most of the time it is hard for the machine learning models to consider more deep information.

Based on a large number of variables in the stock market and the negligible signal-to-noise ratio, deep learning has gradually attracted attention in predicting stock prices. As an effective method of time series modeling, the long short-term memory (LSTM) model has been proved to be effective and widely used [7, 8]. Compared with the recurrent neural network (RNN) model, neurons in LSTM models can transfer data to the previous layer or the same layer. Non-single flowing data enables the neural network to have long and short-term memory simultaneously, thus solving the vanishing gradient issue when RNN processes long data sequences. David M. Q. Nelson et al. found that LSTM neural networks can learn from high-dimensional data onto any dimension reduction techniques (such as feature selection) [7]. Although other deep learning models may have good effects in different directions, LSTM models still retain great potential in long-term prediction [9].

In this paper, a series prediction mechanism based on time and space fusion is proposed to solve the simultaneous influence of time series and multiple variables on stock price volatility. We use LSTM to consider time series prediction, introduce spatial multidimensional variables, and use CNN neural network to extract spatial feature vector information. Finally, the comprehensive influence of different influencing factors on stock price is considered. Combined with the neural network, connection and feedback mechanism, the extracted feature vectors are transferred to the time series prediction neural network to modify and supplement the internal information of the time series. Compared with traditional stock price forecasting methods. Our model has higher accuracy and solves the prediction bias caused by the joint action of multiple factors in actual shares.

The rest of the text is arranged as follows. The second part introduces the prediction method of space-time series, the third part provides the principle of using the model, the fourth part replaces the data, uses the model for experimental verification, and the fifth part draws the conclusion and summary. In order to solve the impact of time series and multi-variables on stock price volatility simultaneously, a Spatio-temporal series prediction mechanism is proposed. That is, the time series prediction is considered, the spatial multidimensional variables are introduced on this basis, and the information of spatial feature vectors is extracted by using a neural network. Finally, the comprehensive influence of different influence factors on stock prices at the same time is considered.

Combined with the connection and feedback mechanism of neural network, the extracted feature vectors are transmitted to the time series prediction neural network. In addition, the internal information of time series is corrected and supplemented, which improves the accuracy of the whole model and solves the problem of prediction deviation caused by the combined action of various factors in actual stocks.

The rest of the text is arranged as follows. The second section introduces the prediction method of the Spatio-temporal series, the third section provides the principle of using the model, the fourth section replaces the data and uses the model for experimental verification, and the fifth section draws conclusions and summarizes.

2. Spatiotemporal sequence prediction model

The past time series of stock prices and the changes of other variables, including the day's high and low prices and prices of other stocks, will have different impacts on stock prices. Neither the model that only considers time series factors (as shown in Figure 1) nor the model that only considers other variables (as shown in Figure 2) can comprehensively predict stock prices. Therefore, it is necessary to consider the effects of time series and other variables in stock price prediction. Our model is composed of a variable influence module and a time series influence module, which are considered together to predict stock price.

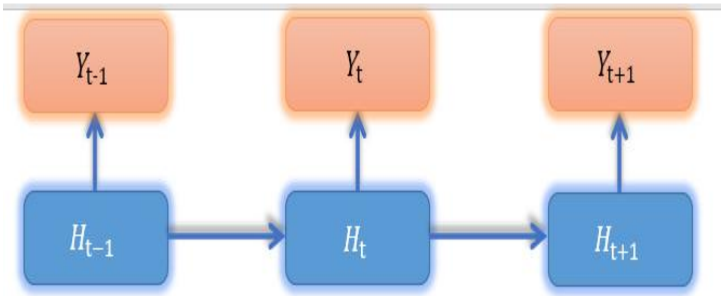

Fig. 1 Structure diagram of stock price prediction model with only time series factors

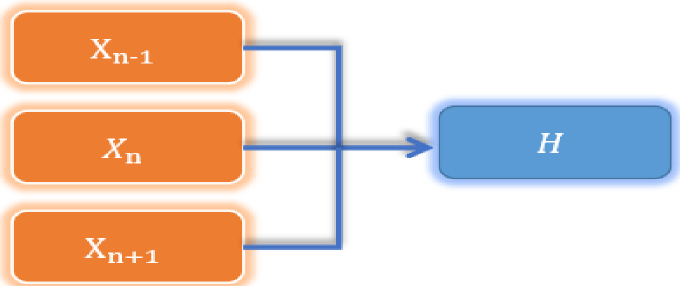

Fig. 2 Structure diagram of stock price prediction model with only variable factors

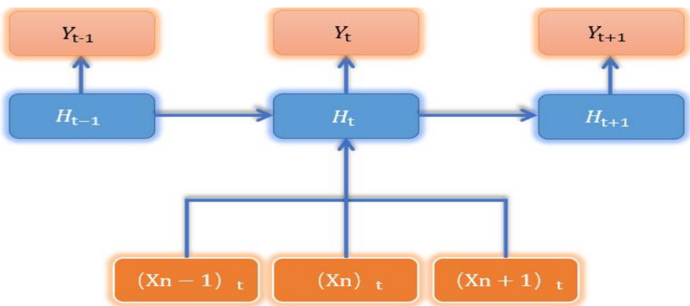

Fig. 3 Structure diagram of stock price prediction model considering time series and other variables

3. Spatiotemporal sequence prediction model

3.1 Recurrent neural network

A recurrent neural network (RNN) is a specific neural network. Unlike ordinary neural networks, RNN models can combine the past information to learn the present information when using sequential data. This is thanks to the feedback loops within the RNN that allow information to pass from one step of the network to the next. [9]

Initially, (time step t) after input X_t, the result is calculated by RNN as follows:

$$H_t = f(X_t) \tag{1}$$

Since RNN considers the association between different moments, that is, in the next time step (time step t+1), RNN will receive two inputs, X_{t+1} and H_t, and then:

$$H_{t+1} = f(X_{t+1}, H_t) \tag{2}$$

As a result, information on the past has been preserved for a relatively long time. Intuitively, the simple internal information transfer structure of RNN is shown in Figure 4.

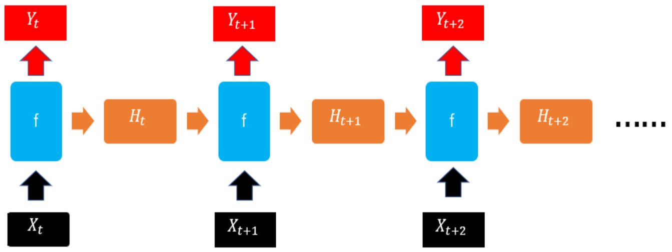

Fig. 4 Internal information transfer structure of RNN

Although RNN can solve the problem of neural network time-series information, it does not pay too much attention to the coupling relationship between variables. The stock price is not only related to time change series but also affected by the interaction of different stock price changes.

3.2 Long Short-Term Memory

As we all know, stock prices change over time. In the short term, stock prices may be affected by occasional events and fluctuate, thus showing a certain temporal correlation. Even though stock prices often fluctuate, they still show a clear trend in the long run. Therefore, in the direction of the time axis, the LSTM system is more consistent with the change law of stock prices. LSTM system is mainly composed of LSTM cells, which consists of (I) input gate, (II) cell state, (iii) forget gate, (iv) update gate (v) output gate along with (vi) a sigmoid layer, (vii) tanh layer, and (viii) pointwise multiplication and the structure is shown in Figure 5:

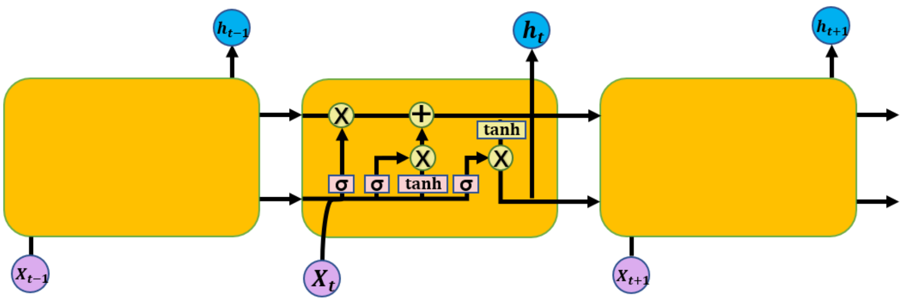

Fig. 5 The repeating module in an LSTM contains four interacting layers

The horizontal line across the top of the LSTM is a cell state that links each repeated cell structure. Moreover, the cell state moves along the chain, like a conveyor belt, and has small linear interactions. Inside the cell structure, the input gate and output gate are similar to RNN, which are the results of input vector x_t and output generated by LSTM. Both the sigmoid layer and the tanh layer represent a "neuron", that is, a $W_t X + b$ operation, differing in the activation function used. The sigmoid layer uses the sigmoid function, which prints numbers between 0 and 1. A value of 0 means "let nothing pass"; A value of 1 means "let everything pass." The tanh layer uses the hyperbolic tangent function, which outputs the numbers between -1 and 1. Its primary function is to normalize the content.

In addition to the hidden layer, another characteristic of LSTM cells is the gate structure. The gates that update the unit status are: (I) forget gate, (ii) update gate. Forget Gate contains a sigmoid layer, which runs as follows: $[h_{t-1}, x_t]$ (information after $h_{t-1}$and x_t are connected) outputs a result ft through the sigmoid layer, and then f_t is multiplied by the elements of cell state C_{t-1}. If a value of f_t is 0, the corresponding information of C_{t-1} will be deleted. If the value is 1, all are reserved. When the value is between 0 and 1, part of the information is reserved. Another gate structure, called the update gate, is divided into two parts. One of the parts is in accordance with the structure of the forget gate, and the tanh layer normalizes the other part to form a new cell state $\tilde{C}_t$ according to the fusion of the two parts. Add the updated information of the last two gates to get the new unit status C_t.

Finally, the updated Ct is screened through an output gate consisting of a tanh layer and a sigmoid layer to obtain the information to be output, and the output is. The mathematical expression of LSTM is the expression of each gate, which can be summarized as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{6}$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = O_t \cdot \tanh(C_t) \tag{8}$$

3.3 Variables Influence Model

The stock price is also affected by other variables. In addition to the daily closing price, there are also the highest price, lowest price, and other changes in the stock price, which will make people make different investment decisions, thus affecting the stock price. Therefore, in our model, we use CNN to capture variable relations. Convolutional neural network (CNN) is a deep feedforward neural network containing convolutional computation and deep structure, which can realize function mapping from input to output. It is widely used in image classification, target recognition, natural language processing, and other fields.

A typical CNN network consists of six parts: input layer, convolution layer, pooling layer, flatten layer, full connection layer and output layer. The function of the convolution layer is to extract the input characteristic information. The activation function layer makes the characteristics of sample output nonlinear. The function of the pooling layer is to process further the feature mapping results obtained by a convolution operation. The fully connected layer is a nonlinear combination of features extracted from the previous layers. The network itself has the characteristics of "local link" and "weight sharing", which simplifies the complexity of the network link, improves the ability to extract abstract features of the model, and to some extent, alleviates the problems of slow training speed and easy over-fitting of the fully connected network. The process of convolution operation is expressed as

$$S(i,j) = (IK)(i,j) = \sum_m \sum_n I(m,n)K(i-m, j-n) \tag{9}$$

Where S is the output characteristic sequence. i, j are the input data; Here, we take the values of various variables as input data. K is the convolution kernel.

3.4 A hybrid model for stock price prediction based on CNN-LSTM

In order to extract more comprehensive information on stock price changes, we use CNN-LSTM hybrid neural network structure to predict stock price. The CNN model's convolutional and pooling layers are first used for feature extraction of all known variables. Then the extracted feature data will be input into the LSTM model through the activation function layer for time series prediction. Finally, the model's predicted values are obtained by accessing the full connection layer. The operation of convolution and pooling reduces the complexity of input data and can deal with outlier data more effectively, avoiding the phenomenon of over-fitting. By adding the LSTM model, the forgetting gate and input gate of LSTM can be used to effectively screen and update the time series data, and the multi-factor influence of stock market can be taken into more comprehensive consideration. [10]

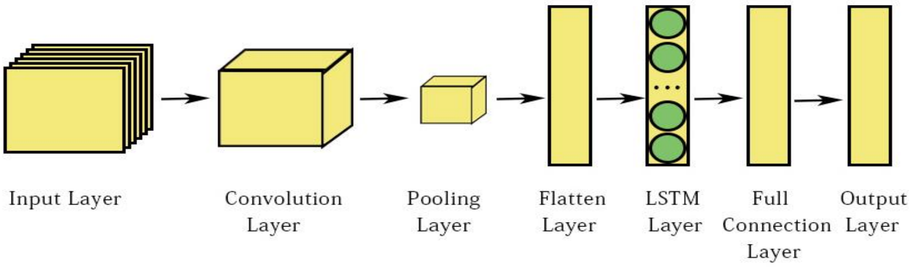

Fig. 6 CNN-LSTM model structure

4. Experiments

In this part, we select the actual stock data of GRANDLAND GROUP, ERA and XIUQIANG to verify the effectiveness of the proposed method. The three stocks chosen in this paper are all the leading stocks of photovoltaic building integration. The concept of photovoltaic building conforms to the new policy of carbon peak and carbon synthesis. The industry is currently in the sunrise industry and has great development prospects. Guangtian Group, as a forecast target stock, has a deep industrial foundation in the greater Bay Area infrastructure, green decoration, and other directions. The company's gross profit rate is always positive, that is, it can ensure long-term and stable profits. The other two stocks are both in the construction and photovoltaic sectors and have maintained rapid appreciation in recent years. So these three underlying stocks have predictive value.

In our experimental data, the stock data of GRANDLAND GROUP, ERA and XiuQiang (hereinafter referred to as company A, B, and C, respectively) all contain four variables. What we want to predict is the closing price of Company A, with the training data of 2012, and output the closing price of the next 100 days according to the training result. In the proposed method, we take the opening, highest and lowest prices of A, B, and C companies as inputs, and jointly consider the influence of time series and other variables to predict the closing price of A Company. Figure 1 shows the prediction effect of the proposed model, and the prediction error is 0.028.

For comparison, we also use ARIMA [1], SVM [3], LSTM [7], and RNN [8] models. ARIMA model is A univariate model established only considering the time series transformation of the closing price of company A. Figure 8 shows the prediction results of the ARIMA model, and the prediction error is 0.058. The SVM model considers the relationship between the opening price, the highest price, the lowest price, and the closing price of company A, without considering the influence of time series on the closing price. Figure 9 shows the prediction results of the SVM model, and the prediction error is 0.041. LSTM and RNN models comprehensively consider the influence of company A's opening price, high price, low price, and closing price time series. However, they do not consider the mutual influence of other companies. Figure 10 and Figure 11 respectively show the prediction effects of the LSTM model and RNN model, and the prediction errors are 0.045 and 0.043.

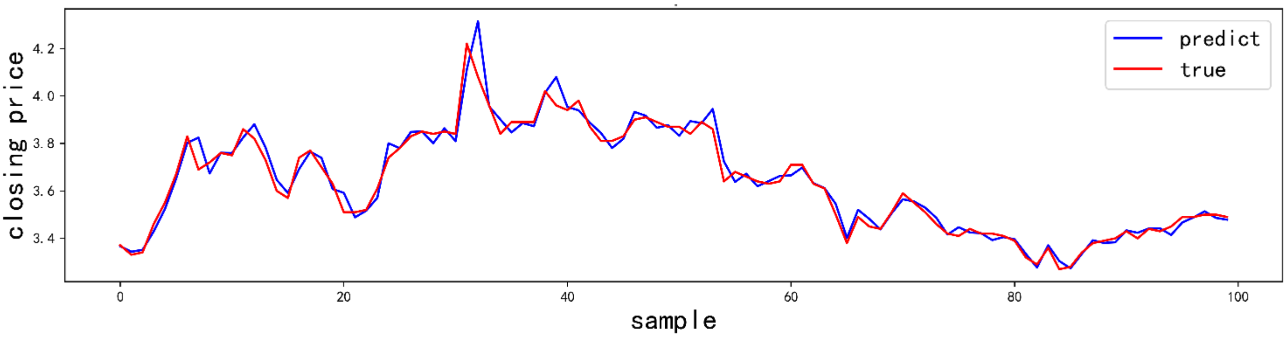

Fig. 7 The prediction results of the proposed model (CNN-LSTM)

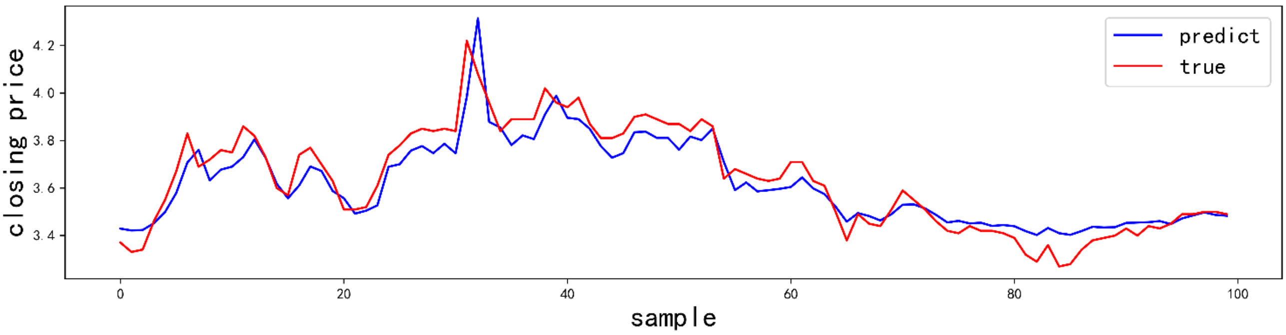

Fig. 8 The prediction results of ARIMA model

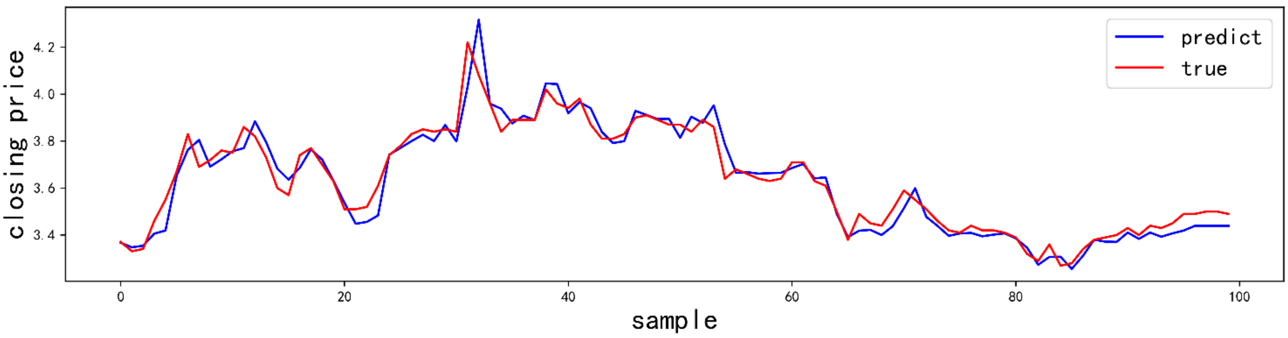

Fig. 9 The prediction results of SVM model

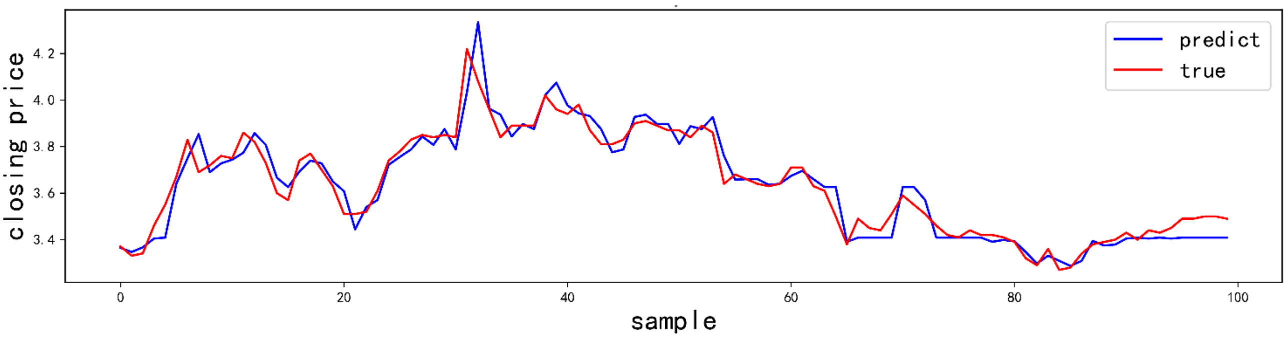

Fig. 10 The prediction results of LSTM model

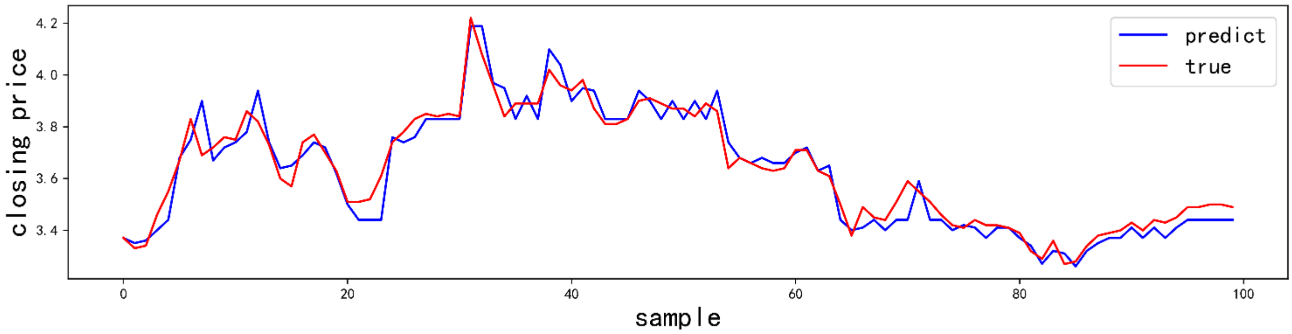

Fig. 11 The prediction results of RNN model

The above results show that the error value of the CNN-LSTM algorithm proposed is the smallest, only 0.028. It can also be noted that the prediction curve of CNN-LSTM algorithm is closest to the real curve and has a good fitting effect in some fluctuation cases. In the long and short-term memory networks, the longer the distance from the prediction time, the less the influence on the prediction target. Therefore, through experience and debugging in the experiment, we finally select the parameter that has the most significant influence on the prediction target. When the training duration increased, there was no significant effect on the final prediction results. However, due to insufficient training data, there will be a large error when the training time is short. Therefore, when the duration of training changes, the prediction results generated by the parameters we selected can also remain robust.

These results reflect the accuracy of CNN-LSTM model in forecasting because it can comprehensively consider the influence of the time series of closing price itself and the influence of opening price, high price, low price, and other stocks on the closing price. However, other models are obviously not comprehensive enough: ARIMA only inputs the time series of the opening price, which cannot consider the influence of other variables; The SVM model only takes the opening price, the highest price, and the lowest price as input, so it lacks the time series and the influence of other stocks on the closing price. LSTM and RNN models consider the influence of opening price, high price, low price, and the time series of closing price, but do not consider the influence of other stock price fluctuations on the target forecast stock.

5. Conclusion

In view of the complexity of stock price fluctuation, this paper proposes a hybrid neural network model based on CNN and LSTM to predict stock price. The traditional ARIMA model can only describe the linear model well, while the machine learning model can only capture the information of other variables in the horizontal direction. For the nonlinear stock prediction problem, the hybrid neural network model proposed by us can capture the information of other variables while collecting the information of time series to obtain better prediction results. In the model of this paper, the CNN model uses a convolutional neural network to extract information about four price fluctuations of stocks, including closing price, opening price, highest price and lowest price, as well as other stocks in the same sector, and integrate and analyze the correlation between variables. At the same time, LSTM uses long and short-term memory networks to extract the change information of the same stock and the same price in different periods to analyze the influence of historical information on the existing information. After the fusion of the two models, the model can capture the multi-temporal and multi-faceted information in the market, which improves the prediction accuracy of the whole model and makes the results closer to its own internal logic. The results show that the prediction accuracy of the hybrid model is about 0.025 higher than that of ARIMA, about 0.01 higher than that of single LSTM and RNN, and about 0.005 higher than that of SVM. The prediction result of the hybrid model is more accurate and reliable.

The model in this paper not only considers the change of stock price in the time series direction, but also considers the coupling effect of other stocks on the predicted targets. Moreover, the model in this paper proves that the volatility of stock price is consistent under the coupling effect. Therefore, in the actual investment process, we can refer to the training results of the proposed method to make better investment decisions. In terms of financial market supervision, if the stock price deviates greatly from the model's stock price, it indicates an inconsistency between the predicted target and the coupling target, that is, abnormal volatility. The possibility of artificial operation of the financial market may need to be taken into account, and intervention and review by the financial market supervision department are required.

References

[1] Ferson, Wayne E. "Stock price predictability." The New Palgrave Dictionary of Economics (2nd ed.). Palgrave Macmillian Ltd 10 (2008): 978-1.

[2] Ariyo, Adebiyi A., Adewumi O. Adewumi, and Charles K. Ayo. "Stock price prediction using the ARIMA model." 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. IEEE, 2014.

[3] Domian, Dale L., and David A. Louton. "A threshold autoregressive analysis of stock returns and real economic activity." International Review of Economics & Finance 6.2 (1997): 167-179.

[4] Hegazy, Osman, Omar S. Soliman, and Mustafa Abdul Salam. "A machine learning model for stock market prediction." arXiv preprint arXiv:1402.7351 (2014).

[5] Kamble, Rupesh A. "Short and long term stock trend prediction using decision tree." 2017 International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2017.

[6] Khaidem, Luckyson, Snehanshu Saha, and Sudeepa Roy Dey. "Predicting the direction of stock market prices using random forest." arXiv preprint arXiv: 1605.00003 (2016).

[7] Khare, Kaustubh, et al. "Short term stock price prediction using deep learning." 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT). IEEE, 2017.

[8] Nelson, David MQ, Adriano CM Pereira, and Renato A. De Oliveira. "Stock market's price movement prediction with LSTM neural networks." 2017 International joint conference on neural networks (IJCNN). IEEE, 2017.

[9] Selvin, Sreelekshmy, et al. "Stock price prediction using LSTM, RNN and CNN-sliding window model." 2017 international conference on advances in computing, communications and informatics (icacci). IEEE, 2017.

[10] Wu, Jimmy Ming-Tai, et al. "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators." Multimedia Systems (2021): 1-20.