LSTM-XGBoost Application of the Model to the Prediction of Stock Price

Sun Yu[1], Liwei Tian[2(✉)], Yijun Liu[1], and Yuankai Guo[3]

[1] Guangdong University of Technology, Guangzhou 510006, China
[2] Guangdong University of Science and Technology, Dongguan 523083, China
[3] Ankang Vocational Technical College, Ankang 725000, China

Abstract. With the rapid rise of economic globalization and digital economics, the rapid development of the global economy has been promoted. As the most important part of the financial industry, the stock market has an important impact on economic fluctuations. How to improve the changing trend of stock market fluctuations has become a hot topic that many scholars and investors pay most attention to. This paper uses the XGBoost model to train the opening price, closing price, highest price, lowest price, trading volume, change, adjusted closing price, and converted time data information in the processed stock historical data set, and train it The results are saved. Then input each attribute into the LSTM model for prediction, and use the prediction result of each attribute as the test set of prediction after XGBoost training, and continuously adjust the parameters of each model, and finally get the optimal stock fluctuation prediction model, LSTM_XGBoost model. The LSTM_XGBoost model is applied to the five stocks ES = F, YM = F, AAPL, SI = F, and CL = F to predict the rise and fall of five stocks. The model is compared and verified by five evaluation indexes: the root mean square error RMSE, the average absolute error MAE, the coefficient of determination R2, the accuracy rate, and the f1-score. It is found that the LSTM-XGBoost model proposed in this paper has risen and fallen in stocks. There is a certain degree of stability and feasibility in the forecast.

Keywords: LSTM · XGBoost · RNN · Stock price forecast · Time series data

1 Introduction

With the continuous strengthening of economic globalization, financial trade and business exchanges between countries are increasing. The rise and fall of stock price is affected by both internal and external factors. Due to its nonlinearity and irregularity, stock time series prediction is usually considered as one of the most challenging tasks. Therefore, it has important theoretical value and practical significance to study the value of stock time series data and the prediction of its rise and fall trend.

As early as 1879, William Stanley Jevons had applied time series analysis to economic problems. In 1995, Arbabi and Mansur proposed a prediction system and method combining neural network with statistical prediction, which showed good prediction effect, and provided important reference for many scholars in the future research of

X. Sun et al. (Eds.): ICAIS 2021, LNCS 12736, pp. 86–98, 2021.
https://doi.org/10.1007/978-3-030-78609-0_8

neural network combination model. In 2000, Henrique S. Hippert proposed a hybrid prediction system, which combined linear model and multilayer neural network, and showed good performance in time series data prediction. In 2015, Kazuhiro Kohara used prior knowledge and neural network to improve the ability of multiple prediction [1]. The results show that the event knowledge and neural network are effective in stock forecasting. In 2018, Kim, Ha Young proposed a new hybrid long-term short-term memory (LSTM) model to predict stock price volatility. The model combines LSTM model with various generalized autoregressive conditional heteroscedasticity (GARCH) models. Through experimental comparison, it is found that GEW-LSTM hybrid model has better prediction performance [2–4]. Shaolong Sun proposed a hybrid ensemble learning method, which combines AdaBoost algorithm and LSTM network to predict financial time series. The empirical results show that this method is superior to some other single prediction models and ensemble learning methods, and it is a promising method for financial time series data forecasting [5]. Weiling Chen proposed a new RNN boost hybrid model to predict the volatility of China's stock market. The experimental results show that the model is superior to other popular methods and can achieve good prediction performance [6]. In 2019, Sidra Mehtab proposed a hybrid stock price forecasting method using machine learning, deep learning and natural language processing. The results show that the method is effective and practical in stock forecasting [7]. In 2020, the first mock exam of ARI-MA-LS-SVM stock was proposed by Chenglin Xiao. The experimental results show that the hybrid model is better than the single model in the prediction of stock price, and the prediction performance is better [8]. According to the research of the above scholars, it is not difficult to see that the combined model has incomparable advantages in time series prediction.

In this paper, the LSTM-XGBoost combination model is constructed by using the combination forecasting method of LSTM and XGBoost [9, 10].This paper uses the LSTM-XGBoost model to predict the stock price and compares it with the single LSTM network model and RNN network model, so as to verify that the LSTM-XGBoost model proposed in this paper is stable and feasible in the stock price forecast.

2 Related Work

2.1 Long and Short-Term Memory Neural Network (LSTM)

LSTM (Long and Short-Term Memory) is the variant of Recurrent Neural Network (RNN). It is a general term for a series of neural networks that can process time series data [11]. There are input gate, forgetting gate and output gate in LSTM unit. When information enters the LSTM network, it can be selected by request. Only the information that meets the requirements of the algorithm will be left behind, while the information that does not meet the requirements will be forgotten through the forgetting gate. LSTM adds and removes neuron information through gating unit. In order to get the correct information, sigmoid neural network layer and multiplication operation are used to determine the information.

Although in theory, RNN can deal with any long-distance dependence problem, in fact, it is difficult to implement in solving long-range data training due to problems such as gradient disappearance and explosion [12]. In order to solve the problem of

gradient vanishing and gradient explosion in the training of long sequence data, LSTM introduces "gate" mechanism and memory unit, that is to use LSTM unit to replace the hidden layer neural unit of RNN. The difference between LSTM structure and the main data transmission unit of ordinary RNN is shown in Fig. 1.

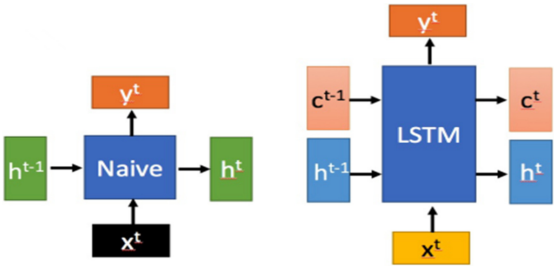


Fig. 1. Comparison of input and output between LSTM and RNN

It can be seen from Fig. 1 that there is only one transfer state ht in RNN unit, while there are two transmission states ht and ct in LSTM.Among them, ct is the unit state value of the current hidden layer, and ct-1 is the unit state value of the previous hidden layer. ht is the output of the hidden layer, and ht-1 is the output of the previous hidden layer.

2.2 XGBoost Model

XGBoost (eXtreme Gradient Boosting) is a machine learning system based on Gradient Boosting. It is an optimized distributed gradient enhancement library, which aims to achieve high efficiency, flexibility and portability. XGBoost provides parallel tree promotion, which can quickly and accurately solve many data science problems. Tree set model is used to treat trees as a set of classification and regression trees for basic learners, so it has good effect in retrospective prediction. XGBoost is an improved algorithm of GBDT, and the value of optimization objective function is the core of XGBoost. The specific process of objective function optimization is as follows.

The prediction model of XGBoost can be expressed as formula (1)

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{1}$$

K: Total number of trees, f_k: The k-th tree, $\hat{y}_i$: the prediction result of sample x_i.

The objective function is represented by formula (2)

$$obj(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}) + \sum_{k=1}^{K} \Omega(f_k) \tag{2}$$

$l(y_i, \hat{y})$: Training error of sample x_i, $\Omega(f_k)$: The regular term of the k-th tree.

$$\mathrm{obj}^{(t)} = \sum_{i=1}^{n} l((y_i, \hat{y}^{(t-1)}) + f(x_i)) + \Omega(f_t) + C \tag{3}$$

Taylor expansion is used to approximate the original objective function

$$obj^{(t)} \approx \sum_{i=1}^{n} l[(y_i, \hat{y}^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) + C \tag{4}$$

In model training, the objective function can be expressed by formula (5)

$$obj^{(t)} = \sum_{j=1}^{t} [(\sum_{i \in I_j} g_i) w_j + \frac{1}{2} (\sum_{i \in I_j} h_i + \lambda) w_j^2] + \gamma T \tag{5}$$

Definition formula (6),

$$G_j = \sum_{i \in I_j} g_i, H_j = \sum_{i \in I_j} h_i \tag{6}$$

By introducing formula (6) into formula (5), Formula (7) is obtained,

$$\begin{aligned} obj^{(t)} &= \sum_{j=1}^{t} [G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2] + \gamma T \\ &= -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \end{aligned} \tag{7}$$

According to the optimal results of formula (6) and formula (7), the optimal decision tree node is found, and then the calculation and prediction are carried out.

3 Model Construction

LSTM-XGBoost is constructed by combining LSTM and XGBoost. The specific flow chart is shown in Fig. 2. In order to prevent over fitting phenomenon in the training process of LSTM model, dropout layer is added and the forgetting rate is set to 0.2.

The construction process of LSTM-XGBoost model is as follows:

(1) Obtain the historical data of stock index and process the missing value;
(2) The "date" attribute in the dataset is decomposed into "year", "month" and "weekday", which are referred to as time component;
(3) Using the LSTM method of keras package, the LSTM model is implemented. The attributes of "open", "high", "low", "close", "volume" and "adj close" in the stock history data set are trained respectively, and the LSTM stock price prediction model is constructed, in which units = 128, batch in LSTM model_ size = 64;
(4) The LSTM stock price forecasting model is used to predict the attributes of "open", "high", "low", "close", "volume" and "adj close";
(5) The prediction results are recombined with the "time component" to construct the "text" test set.
(6) Using XGBRegressor method in sklearn package, XGBoost algorithm is implemented. The attributes of "Open", "High", "Low", "Close", "Volume", "Adj Close" in stock historical data set are trained to build XGBoost stock price prediction model;

(7) XGBoost stock price forecasting model is used to predict the reconstructed data set "text" after LSTM prediction. In order to compare with other models, XGBoost model uses default parameters;
(8) Compare the difference between the real value and the predicted value, and judge the performance of LSTM-XGBoost model in the stock price forecast.

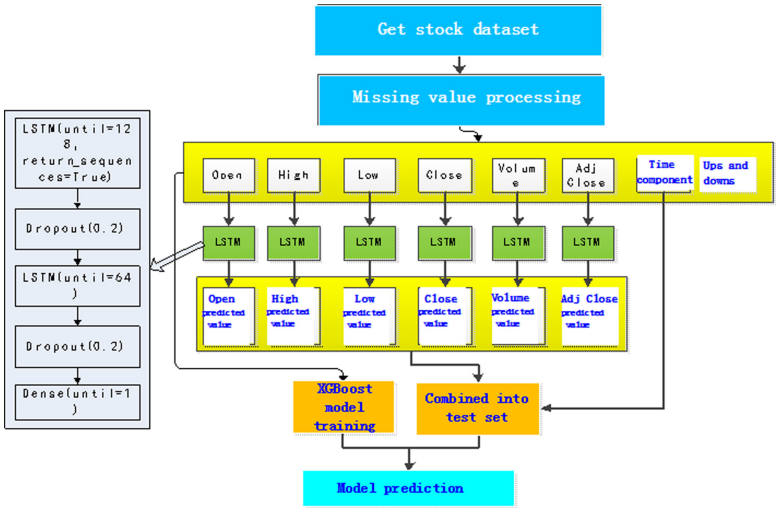
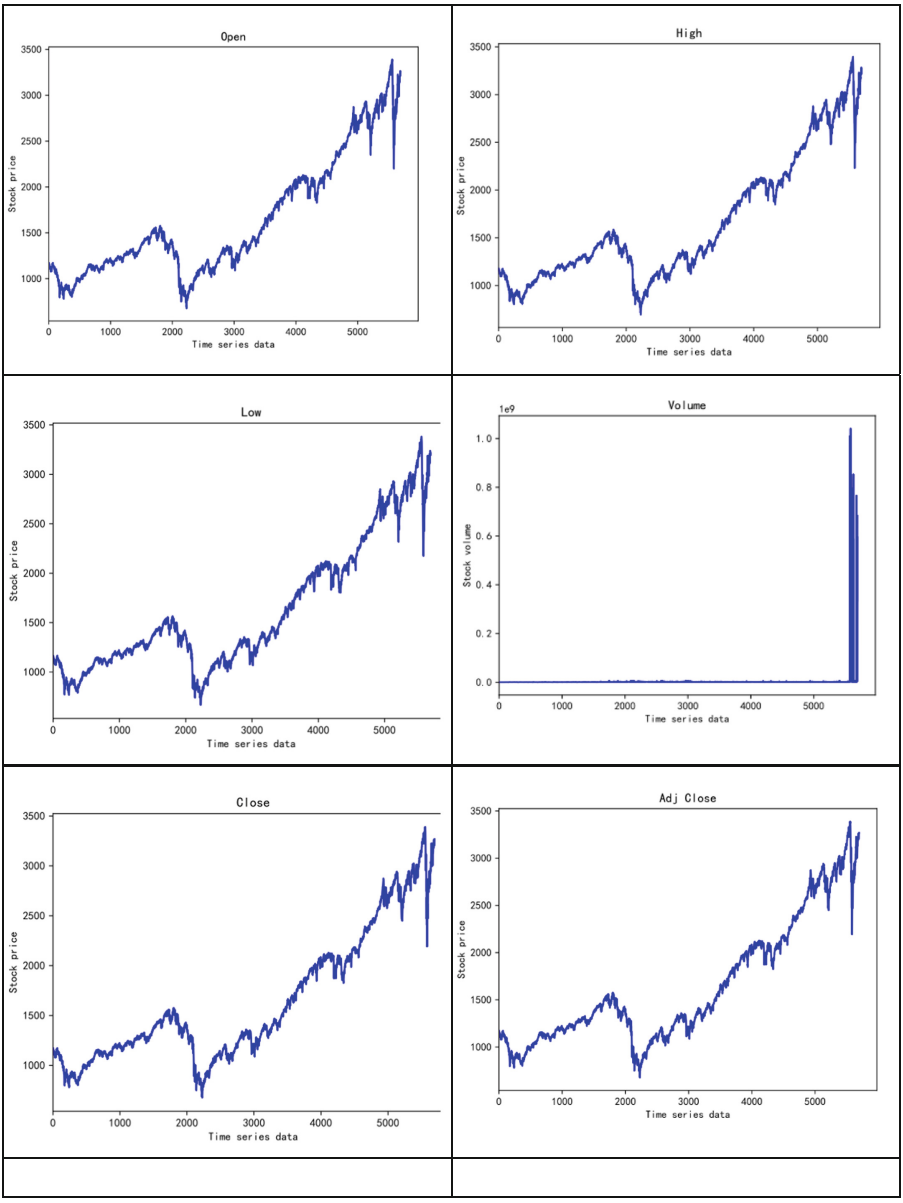


Fig. 2. LSTM-XGBoostModel building diagram

4 Simulation Experiment

In this paper, the experiment is carried out on Intel i7 3.1 ghz dual core four thread CPU and windows 8 operating system. The simulation platform is pycharm, and the python language is used for programming. Sklearn, pandas, numpy, keras and Matplotlib packages in Python are used respectively. The experimental data set was passed https://finance.yahoo.com Download, select the opening price, closing price, highest price, lowest price and trading volume of e-mini S & P500 (ES = F) from January 2002 to August 2020, with a total of 4725 data. Among them, 80% of ES = F stock data is used as training set, and 20% of data is used as test set. The changes of each attribute in ES = F stock data are shown in Table 1.

Table 1. ES = F Change chart of each attribute of stock data



4.1 Evaluating Indicator

The performance evaluation indexes of stock forecasting model are root mean square error (RMSE), mean absolute error (RMSE), determination coefficient R2, accuracy and F1 score. RMSE and RMSE are used to measure the accuracy of the prediction model,

R2 is used to measure the fitting ability of the prediction model, and accuracy and F1 score are used to measure the accuracy of the prediction model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y^{(i)} - \widehat{y}^{(i)})^2} \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y^{(i)} - \widehat{y}^{(i)}\right| \tag{9}$$

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\widehat{y} - \overline{y})^2}{\sum_{i=1}^{n}(y - \overline{y})^2} = 1 - \frac{SSE}{SST} \tag{10}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

TP and FP are the number of correct and wrong prediction of stock price rise, TN and FN are the number of correct and wrong prediction of stock price.

$$f1\_score = \frac{2 * pre * recall}{pre + recall} \tag{12}$$

Equalization precision (pre) and recall rate (recall).

4.2 Experimental Comparison

In Experiment 1, the parameters of LSTM network were determined. Time window length is one of the most important variables affecting LSTM network. In order to find the optimal window length (W), six attributes of ES = f stock, including "open", "high", "low", "close", "volume" and "adj close", are used as experimental data. According to RMSE, Mae and P2 of each attribute prediction result, the window size of each attribute is determined. In order to better verify the performance of each window of LSTM model, epochs of LSTM model is set to 10, batch_ Then the windows were set to one day, one week, one month, three months, and half a year respectively, and RMSE, MAE, and P2 used the average of 10 training times in each window. The specific experimental comparison is shown in Table 2, 3, 4, 5, 6 and 7.

Table 2. Open comparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	**1270.98**	5002.03	8915.17	9669.21	9140.32
MAE	**24.3963**	41.3908	59.5100	68.6469	60.1048
R^2	**0.9821**	0.9295	0.8744	0.8638	0.8712

Table 3. High comparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	**768.74**	3636.71	7752.76	7962.88	7782.34
MAE	**17.6888**	38.0986	62.7833	50.9378	64.8920
R^2	**0.9892**	0.9491	0.8916	0.8886	0.8912

Table 4. Low comparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	**1066.91**	5245.60	9656.68	10757.64	9484.07
MAE	**20.7882**	44.4795	63.8474	57.8145	63.3581
R^2	**0.9848**	0.9253	0.8625	0.8469	0.8650

Table 5. Close comparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	**1273.77**	4476.19	8701.98	8342.23	8513.99
MAE	**22.3174**	43.6851	56.1627	61.6803	63.1581
R^2	**0.9820**	0.9367	0.8771	0.8821	0.8797

Table 6. Volume comparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	1.1000e+16	0.9678e+16	**0.9553e+16**	1.0021e+16	1.0057e+16
MAE	41285747.44	46530125.11	**38994040.05**	38822740.80	40332606.43
R^2	0.4713	0.5348	**0.5408**	0.5183	0.5166

Table 7. Adj CloseComparison of prediction results of attribute windows

Window size	1	7	30	90	180
RMSE	**1227.35**	4887.07	8398.61	8341.29	8202.48
MAE	**20.6443**	49.2456	60.9482	53.6753	53.2474
R^2	**0.9826**	0.9309	0.8813	0.8821	0.8841

According to Tables 2, 3, 4, 5, 6 and 7, when the window size is 1, the RMSE, MAE, and P2 of the "Open", "High", "Low", "Close" and "Adj Close" attributes are the best performance. When the sliding window is 30, the prediction performance of the "volume" attribute is the best compared with other window sizes. According to the comparison, the sliding window of "Open", "High", "Low", "Close" and "Adj Close" is set to 1, that is, the sliding window of "volume" attribute is set to 30. After determining the sliding window of each attribute, LSTM model is trained, and the prediction results are shown in Fig. 3. According to Fig. 3, it is obvious that the prediction fitting degree of each attribute of LSTM model is relatively good.

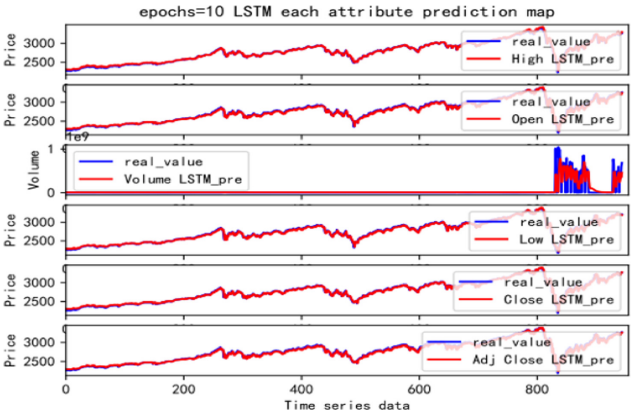


Fig. 3. LSTM prediction chart of each attribute of the model

In Experiment 2, after confirming the parameters of LSTM model, the XGBoost model was trained. The XGBoost parameters were the default parameters, and the training set was the first 80% data of ES = f stock. The XGBoost trained model was saved, and the prediction results of each attribute of LSTM model in Experiment 1 were integrated with time attributes to form a test training set. The XGBoost trained model is used to predict the new test set, and the final prediction results of ES = f stock are obtained. The comparative experimental data of LSTM-XGBoost model, LSTM and RNN models with epochs of 1, 10 and 30 are shown in Table 8, 9 and 10.

Table 8. Epochs = 1 Comparison of model prediction results

	LSTM-XGB	LSTM	RNN
RMSE	612.5332	613.0891	**611.2626**
MAE	**15.5**	15.7	15.6
Accuracy	**0.60**	0.59	0.58
$f1$_score	**0.7524**	0.7358	0.7248

Table 9. Epochs = 10 Comparison of model prediction results

	LSTM-XGB	LSTM	RNN
RMSE	612.6305	656.2361	**610.3129**
MAE	**15.5**	17.0039	**15.5**
Accuracy	**0.60**	0.39	0.57
$f1$_score	**0.7524**	0.0238	0.1554

Table 10. Epochs = 30 Comparison of model prediction results

	LSTM-XGB	LSTM	RNN
RMSE	615.5513	**610.3231**	610.6260
MAE	**15.6**	**15.6**	**15.6**
Accuracy	**0.60**	**0.60**	**0.60**
$f1$_score	**0.7524**	0.7524	0.7508

It can be seen from Table 8 that when epochs = 1, the prediction accuracy and fitting degree of LSTM-XGBoost model are better than those of LSTM and RNN models. According to Table 9, when epochs = 10, the prediction performance of LSTM-XGBoost model is significantly higher than that of LSTM-XGBoost model. Although RNN model is slightly better than LSTM-XGBoost model in fitting degree, the prediction accuracy of LSTM-XGBoost model is higher than that of RNN model. Table 10 shows the comparison results of each model when epochs = 30. Although the prediction performance of the models is relatively close, the LSTM-XGBoost model proposed in this paper has reached the prediction performance of LSTM and RNN models when epochs = 1 and epochs = 30. No matter what value epochs is used, the LSTM-XGBoost model shows good prediction performance, which indicates that LSTM-XGBoost model is more stable and practical.

For the first mock exam, we further verify the stability of LSTM-XGBoost model in the prediction of stock price fluctuation, and compare it with the single model LSTM, and compare with the RNN model which performs well in stock price prediction. In order to verify the generalization ability and prediction performance of the models, three models were applied to mini Dow Jones Indus. - $5 Sep 20 (YM = f), AAPL, silver Sep 20 (Si = f) and crude oil SEP, respectively 20 (CL = f), the data sets of four stocks are time series data of "open", "high", "low", "close", "volume" and "adj close" from February 2002 to mid August 2020, in which the first 80% of the data set is the training data set, and the remaining 20% is the test data set. After the model is trained and tested, the experimental results are compared and analyzed. The experimental results are shown in Tables 11, 12, 13 and 14.

From Tables 11, 12, 13 and 14, we can see that LSTM-XGBoost model has good prediction performance in the rise and fall forecast of the four stocks. According to

Table 11. "YM = F" Comparison of model prediction results

Model	Epochs = 1			Epochs = 10			Epochs = 30		
	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN
RMSE	**58552**	70056	59087	**58522**	65292	85784	**59882**	71790	85134
MAE	**162**	197	168	**162**	185	234	**166**	202	233
Accuracy	**0.58**	0.41	0.42	**0.58**	0.41	0.41	**0.58**	0.41	0.41
$f1$_score	**0.74**	0.0	0.07	**0.74**	0.0	0.0	**0.74**	0.0	0.0

Table 12. "APPL" Comparison of model prediction results

Model	Epochs = 1			Epochs = 10			Epochs = 30		
	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN
RMSE	**9.52**	10.8	15.24	**9.56**	14.06	13.15	**9.56**	13.39	13.91
MAE	**2.02**	2.30	3.03	**2.03**	2.88	2.74	**2.04**	2.76	2.85
Accuracy	**0.54**	0.45	0.45	**0.54**	0.45	0.45	**0.53**	0.45	0.45
$f1$_score	**0.71**	0.0	0.01	**0.70**	0.0	0.01	**0.69**	0.0	0.01

Table 13. "SI = F" Comparison of model prediction results

Model	Epochs = 1			Epochs = 10			Epochs = 30		
	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN
RMSE	**0.04**	1.44	1.29	**0.04**	0.76	1.13	**0.04**	1.44	1.29
MAE	**0.13**	1.18	1.12	**0.12**	0.85	1.04	**0.13**	1.18	1.12
Accuracy	**0.59**	**0.59**	**0.59**	**0.59**	**0.59**	**0.59**	**0.59**	**0.59**	**0.59**
$f1$_score	**0.74**	**0.74**	**0.74**	**0.74**	**0.74**	**0.74**	**0.74**	**0.74**	**0.74**

Table 14. "CL = F" Comparison of model prediction results

Model	Epochs = 1			Epochs = 10			Epochs = 30		
	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN	LSTM-XGB	LSTM	RNN
RMSE	**1.48**	22.74	54.81	**1.51**	36.14	66.74	**1.91**	56.05	62.35
MAE	**0.76**	4.63	7.32	**0.77**	5.91	8.10	**0.93**	7.41	7.82
Accuracy	**0.54**	0.46	0.46	**0.51**	0.46	0.46	**0.53**	0.46	0.46
$f1$_score	**0.63**	0.0	0.0	**0.64**	0.0	0.0	**0.69**	0.0	0.0

RMSE, MAE, accuracy and F1_ In the score evaluation index, it can be clearly seen that the lstm-xgboost model is significantly better than the LSTM model in the prediction of the rise and fall of the stock YM = F, APPL and Cl = F, and has a comparative

advantage compared with the RNN model. The accuracy and F1 of the three models are shown in Table 13_ In RMSE and Mae, lstm-xgboost model is better than LSTM and RNN models. It further shows that the performance of LSTM xgboost model is stable and efficient.

5 Conclusion

This paper proposes a hybrid financial time series model based on LSTM and XGBoost, namely LSTM_ XGBoost model. The XGBoost model is used to train the opening price, closing price, highest price, lowest price, trading volume, up and down range, adjusted closing price and converted time data information of the processed stock historical data set, and the training results are saved. Then the opening price, closing price, highest price, lowest price, trading volume and adjusted closing price are separately input into LSTM model for prediction. The prediction results of each attribute are used as the test set of XGBoost training prediction, and the parameters of each model are continuously adjusted to obtain the optimal stock rise and fall prediction model. The model is verified by the ups and downs of ES = F, YM = F, AAPL, SI = F and Cl = F. Through the evaluation indexes of root mean square error RMSE, mean absolute error MAE, prediction accuracy and F1_ LSTM was found by comparing score_ XGBoost model in the prediction of five stocks show stable and good prediction performance, that is, the LSTM XGBoost model proposed in this paper is stable and feasible in the stock price forecast. But the overall prediction performance needs to be further improved. In the later stage, we will consider the influence of network public opinion and the correlation between various attributes on the rise and fall of stocks to improve the performance of the model for stock price prediction. So as to bring more valuable reference for people to master the overall rise and fall of stock price.

Acknowledgement. The paper is funded by Dongguan social science and technology development (general) project (No. 2020507154645) in 2020 and the key platform construction leap up program project of Guangdong University of science and technology: network engineering application technology research center of Guangdong University of science and technology.This work is supported by Natural Science Foundation of Guangdong Province of China with No. 2020A1515010784.

References

1. Kohara, K., Ishikawa, T., Fukuhara, Y.: Stock price prediction using prior knowledge and neural networks. Intell. Syst. Account. Finan. Manag. **6**(1), 11–22 (2015)
2. Nelson D., Pereira A., Oliveira R.: Stock market's price movement prediction with LSTM neural networks. In: International Joint Conference on Neural Networks. IEEE (2017)
3. Kim, H.Y., Won, C.H.: Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models. Expert Syst. Appl. **103**(8), 25–37 (2018)
4. Althelaya K.A., El-Alfy E.S.M., Mohammed, S.: Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU). In: 21st Saudi Computer Society National Computer Conference (2018)

5. Sun, S., Wei, Y., Wang, S.: AdaBoost-LSTM ensemble learning for financial time series forecasting. In: Shi, Y., et al. (eds.) ICCS 2018. LNCS, vol. 10862, pp. 590–597. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93713-7_55
6. Chen, W., Yeo, C.K., Lau, C.T.: Leveraging social media news to predict stock index movement using RNN-boost. Data Knowl. Eng. **118**(11), 14–24 (2018)
7. Mehtab S., Sen J.: A robust predictive model for stock price prediction using deep learning and natural language processing. Papers (2019)
8. Xiao, C., Xia, W., Jiang, J.: Stock price forecast based on combined model of ARI-MA-LS-SVM. Neural Comput. Appl. **32**(10), 5379–5388 (2020). https://doi.org/10.1007/s00521-019-04698-5
9. Greff, K., Srivastava, R.K., Jan, K.: LSTM: A Search Space Odyssey. IEEE Trans. Neural Netw. Learn. Syst. **28**(10), 2222–2232 (2016)
10. Chen T., Guestrin C.: XGBoost: a scalable tree boosting system. In: ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2016)
11. Sundermeyer, M., Schlüter, R., Ney H.: LSTM neural networks for language modelling. In: Interspeech, pp. 601–608 (2012)
12. Miao, Y., Gowayyed, M., Metze, F.: EESEN: end-to-end speech recognition using deep RNN models and WFST-based decoding (2016)