

Predicting Chinese Stocks Using XGBoost-LSTM-Attention Model

Zhiyong Yang^{1,2,*}, Yuxi Ye², and Yu Zhou³

Abstract

Forecasting is a popular topic in the stock market. In recent years, many scholars have utilized machine- and deep-learning models in this field. However, many stock forecasting models suffer from problems of information overlap in stock trading data and a relatively simple structure of the prediction model. To overcome these issues, we built a stock forecasting model based on extreme gradient boosting (XGBoost), long short-term memory (LSTM), and attention (XGBoost-LSTM-Attention). XGBoost is used to extract important information from stock data, and the LSTM combined with the attention mechanism can enhance stock prediction performance. To verify the feasibility and effectiveness of XGBoost-LSTM-Attention, we selected 14 Chinese stocks from different industries for the prediction experiments and compared their performance with those of existing models. The experimental results showed that the average root-mean-square error value of the XGBoost-LSTM-Attention model for the different stock datasets was the smallest (0.012); the average R^2 value (0.96) and average accuracy (66.1%) were the highest.

Keywords

LSTM, Attention Mechanism, Extreme Gradient Progression Tree, Stock Forecasting

1. Introduction

The Internet has expanded the scale of economic activities and enhanced investment awareness [1]. Stock investing appeals to investors because of convenience, high risk, and high returns. Stock price forecasting, decision-making, and portfolio models provide useful tools for helping investors plan their investment strategies [2]. However, stock price trends are characterized by instability, randomness, and nonlinearity, making it difficult for investors and researchers to predict stock prices. Previous studies have attempted to improve the predictive performance of stock models, which is also the goal of this study.

Methods for predicting stock time series can be classified into statistical, traditional machine learning (ML), and deep learning (DL) models. Statistical methods include moving average, autoregressive moving average, and logistic regression. However, as stock market systems are complex and nonlinear, statistical models are unable to achieve accurate prediction results [3]. Researchers have attempted to use ML models to solve nonlinear stock time series, such as support vector machines, artificial neural networks, and decision trees [4-8]. Although the aforementioned methods have significantly progressed stock

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received September 28, 2022; first revision November 22, 2022; accepted December 11, 2022.

*Corresponding Author: Zhiyong Yang (zyy@cqvie.edu.cn)

¹ School of Big Data and Internet of Things, Chongqing Vocational Institute of Engineering, Chongqing, China (zyy@cqvie.edu.cn)

² School of Computer and Information Science, Chongqing Normal University, Chongqing, China (2020210516089@stu.cqnu.edu.cn)

³ School of Finance and Tourism, Chongqing Vocational Institute of Engineering, Chongqing, China (zy1982@cqvie.edu.cn)

market forecasting, these models have a highly complex training time and cannot efficiently handle large amounts of financial data.

In recent years, DL models have been utilized in stock prediction research because of their excellent data processing capabilities. The most common techniques include recurrent neural networks (RNN), long short-term memory (LSTM), and convolutional neural networks (CNN) [9-11]. These techniques can easily handle large amounts of financial time series data; however, single models can lead to low model generalization and cannot be applied to volatile stock markets. To overcome this issue, this study proposes a novel framework to enhance stock price prediction accuracy.

Compared with single models, multi-model integration approaches exhibit better predictive performance. In 2017, Cheng et al. [12] proposed an attention-based LSTM model for stock price movement prediction. In 2019, Kim and Kim [13] used LSTM and CNN to extract temporal and image features from SPDR S&P 500 exchange-traded fund (ETF) data and found that these methods can effectively reduce the prediction error. In 2020, Wu et al. [14] converted stock data and leading indicators into simulated images and predicted stock prices using a CNN-LSTM model. The prediction results were more accurate than those of single models. In 2021, Lu et al. [15] proposed a CNN-bidirectional LSTM (BiLSTM) attention mechanism (AM) model to predict stock index prices. The results show that the AM can improve prediction performance. In 2022, Lee [16] combined the gated recurrent unit (GRU) and AM to predict the rise and fall of future trading days. However, stock prices are affected by numerous factors because of the changing market conditions and economic environment, introducing noise into stock prices. However, the aforementioned models do not address the problem of noisy raw stock price data. To address this problem, we used extreme gradient boosting (XGBoost) to process the original data.

In this study, a novel XGBoost-LSTM-Attention model was used to predict Chinese stocks, including the Shanghai and Shenzhen Stock Exchanges. XGBoost can extract important features of the original stock information from the input data, and regarding network structure, the combination of LSTM and the AM can enhance the accuracy of stock predictions. The purpose of this study was to determine whether the XGBoost-LSTM-Attention model can accurately predict stocks in different industries and whether significant differences exist between different stock forecasting models in developing economies.

The remainder of this paper is organized as follows: Section 2 briefly introduces the techniques employed in this study. Section 3 introduces the training process for the XGBoost-LSTM-Attention model. Section 4 presents the experimental research process and an analysis of the experimental results using different models, followed by a conclusion in Section 5.

2. Methodology

2.1 XGBoost

The XGBoost algorithm is an extensible end-to-end enhancement tree system [17] that improves the gradient boosting decision tree (GBDT) objective function by adding a regular term to the original function and expanding the loss function according to the second-order Taylor's formula. In Eq. (1), $L(y_i, \hat{y})$ is the squared difference loss function between the predicted value $\hat{y}$ and real value y_i . In Eq. (2), T represents the number of leaf nodes; W_j is the value of the j -th leaf node; G_j is the sum of the first partial derivative; and H_j is the sum of the second partial derivative.

$$obj = \sum_{i=1}^n (L(y_i, \hat{y})) + \sum_{i=1}^t \Omega(f_i), \quad (1)$$

$$obj^{(t)} = \sum_{j=1}^T \left[G_j W_j + \frac{1}{2} (H_j + \lambda) W_j^2 \right] + \gamma T. \quad (2)$$

The objective function can be simplified by calculating the first-order W_j as follows:

$$obj = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T. \quad (3)$$

XGBoost calculates the split revenue for each node using a greedy algorithm. Eqs. (4) and (5) are the objective functions before and after splitting, respectively, and “Gain” is the basis for calculating feature importance scores and performing feature selection, as expressed in Eq. (6).

$$obj_1 = -\frac{1}{2} \left[\frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] + \gamma, \quad (4)$$

$$obj_2 = -\frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} \right] + 2\gamma, \quad (5)$$

$$\text{Gain} = \frac{1}{2} \left[\frac{G_L^2}{H_L^2 + \lambda} + \frac{G_R^2}{H_R^2 + \lambda} - \frac{(G_L + G_R)^2}{(H_L + H_R)^2 + \lambda} \right] - \gamma. \quad (6)$$

2.2 LSTM

The structure of LSTM is a variant of RNN. LSTM adopts recursive network architecture and gradient based learning algorithm [18] to select the retention and forgetting of data through three “Gates” in the cell structure and solves the problem of gradient disappearance caused by the long input sequence of RNN model, as shown in Fig. 1.

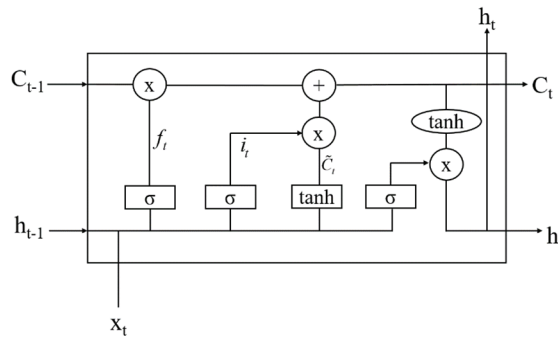


Fig. 1. Structure of LSTM cell.

LSTM states are updated as follows: first, the forgetting gate determines whether to keep the data based on the output of the previous time point and the current input data, as expressed in Eq. (7), where f_t is the forgetting gate, and h_{t-1} and x_t are the outputs of the previous and current moments, respectively.

$$f_t = \sigma(W_f * h_{t-1} + W_f * x_t + b_f). \quad (7)$$

Second, the input gate saves the updated data and stores them in the cell state using Eq. (8), where i_t is the input gate.

$$i_t = \sigma(W_i * h_{t-1} + W_i * x_t + b_i). \quad (8)$$

Third, the cell state is updated using Eq. (9), where C_t is the new cell.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \quad (9)$$

Finally, the output gate determines the data output of the cell state, as expressed in Eqs. (10) and (11).

$$o_t = \sigma(W_o * h_{t-1} + W_o * x_t + b_o), \quad (10)$$

$$h_t = o_t * \tanh(C_t). \quad (11)$$

Here, the parameter set $\{W_o, W_f, W_i\}$ corresponds to the weight matrix of the different gates; $\{b_o, b_f, b_i\}$ represents the corresponding offset items; $\tanh$ and σ are activation functions.

2.3 Attention Mechanism

In 1980, Treisman and Gelade [19] proposed the attention mechanism. Based on the human brain, AM stems from the ability of the human vision to quickly focus on key areas with high attention weights. Similarly, AM calculates the importance of each element in the input sequence and assigns weights to different features according to their importance. The calculation principle of the AM is illustrated in Fig. 2.

First, value (V) of each key (K) is obtained by calculating the similarity between each query (Q) and key. Second, the softmax function is used to normalize the weights of the values to obtain the weight coefficients. Third, the weight coefficient and corresponding value are weighted and summed to obtain the final attention value, as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)v. \quad (12)$$

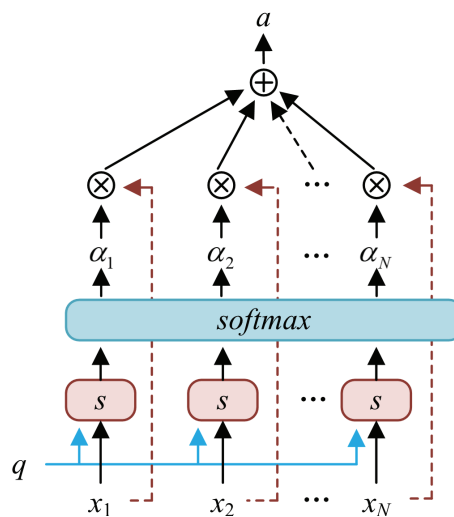


Fig. 2. Attention mechanism.

3. XGBoost-LSTM-Attention

The XGBoost algorithm can extract important features of the input data; consequently, it is widely used for data processing. The LSTM can handle long-term dependence information and is widely used in time-series analysis. The AM can capture the internal correlations of the data and information at different points in the input sequence. Thus, based on the advantages of XGBoost, LSTM, and AM, an XGBoost-LSTM-Attention model for stock prediction is constructed. The network structure is illustrated in Fig. 3. The fundamental structures are the XGBoost, and LSTM attention modules. First, the XGBoost module extracts important feature subsets by computing the important feature scores for each stock dataset. Second, the LSTM-Attention module, which includes LSTM, attention, and BiLSTM layers, improves the generalization ability and accuracy of the prediction model.

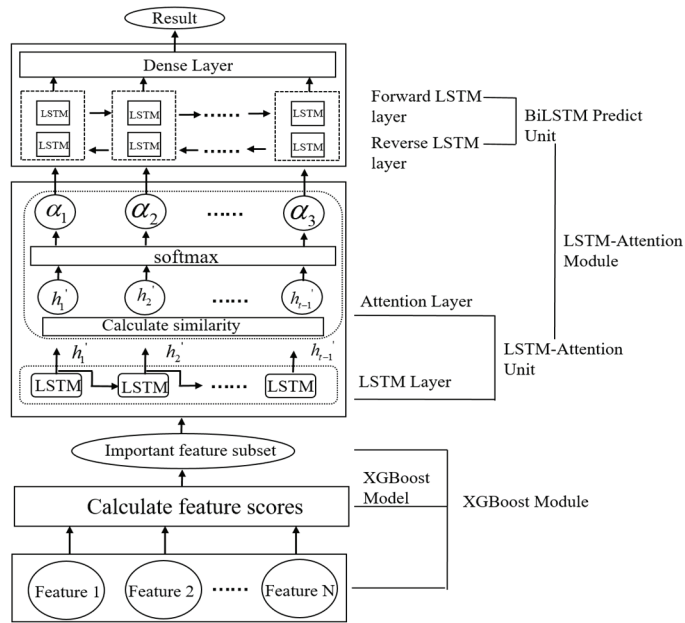


Fig. 3. XGBoost-LSTM-Attention model structure diagram.

3.1 Raw Stock Data Processing

The stock dataset is the key to ensuring that the forecasting model has good generalization ability. The XGBoost algorithm is used to extract important information from the original stock data and obtain important feature subsets. The XGBoost algorithm is outlined in Table 1.

Table 1. Data processing steps

XGBoost Algorithm Processing Data Steps.	
Input:	The original stock data
Output:	Important feature subset
Step 1:	Divide the original stock data into training and validation datasets at a ratio of 7:3.
Step 2:	Build the XGBoost model for training and calculate the feature scores and accuracy of different dimensions.
Step 3:	Arrange the features in descending order of score and calculate the accuracy of different dimensions.
Step 4:	Fuse highest scoring features.

We considered Kweichow Moutai (SH600519) as an example. The original dataset includes the stock data for each trading day from 2012 to 2022. Each piece of data contains nine items: open, high, low, close, previous close, volume, amount, change, and percentage change. The original data are presented in Table 2.

Table 2. Original stock data

date	open	highest	lowest	close	pre_close	volume	amount	change	pct_change
20120104	191.50	192.77	185.00	185.27	193.30	33878.28	637232.81	-8.03	-4.15
20120105	184.00	185.20	182.36	183.15	185.27	30122.64	553347.01	-2.12	-1.14
...	...	...	...	...	...	...	...	...	...
20220609	1872.0	1888.3	1849.0	1853.0	1865.0	26329.02	4897066.6	-12.60	-0.67
20220610	1845.0	1907.0	1835.0	1900.0	1853.0	47344.62	8882462.5	47.6	2.56

The specific steps of the XGBoost data processing are:

First, the XGBoost method is used to build a model. Each feature score of the original stock data is obtained through model training; the scores are then arranged in descending order, as shown in Fig. 4.

Second, the feature dimension is increased using the recursive method, and the accuracies of the models under different dimensions are compared, as shown in Fig. 5. The results show that the accuracy rate is the highest when extracting six-dimensional features (87.5%).

Third, the first six feature items are used as an important feature subset, which includes volume, open, change, amount, change, and close.

Finally, the important feature subsets and technical indicators are used as input data for the LSTM-Attention module. The technical indicators primarily included the moving average (5 days) and Bollinger Bands.

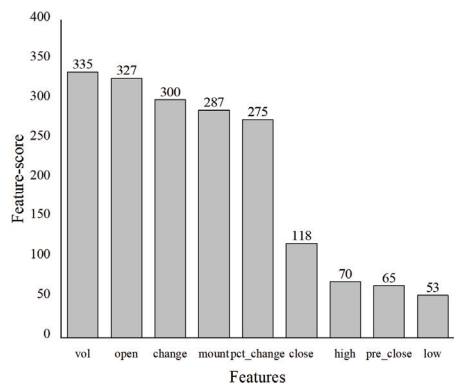


Fig. 4. Stock feature important score.

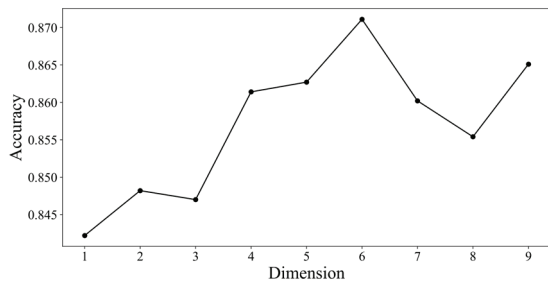


Fig. 5. Accuracy of different dimension features.

3.2 LSTM-Attention Prediction Process

The LSTM attention module predicts stock prices as follows (Fig. 6):

- Input processed data is the dataset obtained after the XGBoost module processing.
- Data standardization: The discrepancies in the stock dataset necessitate that the stock dataset be normalized to better train the model. The normalization method for processing the dataset is shown in Eq. (13), where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the stock dataset, respectively.

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}}. \quad (13)$$

- Network initialization: The weights of each network layer are initialized to a fixed mean or variance.
- LSTM-Attention calculation: The model output value is obtained by calculating each network layer of the LSTM-Attention module.
- Calculation error: The output and actual values are compared to determine the corresponding error.
- End condition corresponds to the maximum number of iterations reached or to a predicted error that is lower than the set threshold.
- Output results: The results are output when the prediction process is completed and the neural network model is trained.

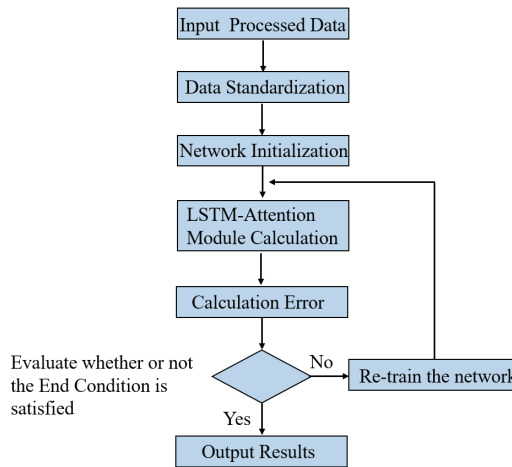


Fig. 6. Training process of LSTM-Attention.

4. Experiments

In this section, we verify the validity of XGBoost-LSTM-Attention on Chinese stocks in different industries, and the performance of different models is compared in developing stock markets [14-16].

4.1 Datasets

We selected Chinese stocks from different industries for the experimental data, which are presented in Table 3. Transaction data for 2,357 trading days from January 2012 to June 2022 were obtained from the Tushare database [20]. To evaluate the generalization error, the stock dataset was split into training and testing datasets at a ratio of 8:2.

Table 3. Stock picking list

Industry	Leading stocks	Small & medium-sized stocks
Finance	China Merchants Bank (SH600036)	Gemdale Group (SH600383)
Real estate	China Vanke (SZ000002)	Chongqing Yukaifa (SZ000514)
Coal	China Shenhua Energy (SH601088)	Shanghai Energy (SH600508)
Steel	Baoshan Iron & Steel (SH600019)	Hunan Valin Steel (SZ000932)
Metals	Zijin Mining (SH601899)	Tibet Mining (SZ000762)
Petrochemical	Petro China (SH601857)	Yueyang Xingchang (SZ000819)
automobile	BYD (SZ002594)	Jiangling Motors (SZ000550)

4.2 Model Implementation

All models were built on the Windows 10 operating system, and the GPU version of the TensorFlow framework was used. The experimental hardware environment was GeForce RTX2060 with 16 GB of RAM. The model settings are listed in Table 4.

Table 4. LSTM-Attention setting

Parameter	Value
Number of hidden units in BiLSTM layer	32
Number of hidden units in LSTM layer	64
BiLSTM layer activation function	tanh
LSTM layer activation function	tanh
Attention layer activation function	softmax
Dense layer	1

The Adam optimizer method was used in the experiment with model training parameters as follows: batch size of 5, time step of 5, epoch of 50, mean absolute error (MAE) as the loss function, and a learning rate of 0.001.

To evaluate the prediction effect of experience level agreements (XLA), root-mean-square error (RMSE), R-square (R^2), and accuracy were used as model evaluation criteria.

The RMSE equation is:

$$RMSE = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}, \tag{14}$$

where $\hat{y}_i$ is the predicted value, and y_i is the actual value. The smaller the RMSE value, the smaller is the prediction deviation of the model.

The R^2 equation is:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \tag{15}$$

where $\hat{y}_i$ is the predicted value, and y_i is the actual value. The closer the R^2 value is to 1, the better is the fitting of the model.

The accuracy (ACC) equation is:

$$ACC = \frac{TP}{TP + TN} * 100\%, \tag{16}$$

where TP and TN denote the number of predicted successes and failures, respectively. The higher the accuracy, the more accurate is the stock price prediction.

4.3 Results

The stock datasets were provided as input to XGBoost-LSTM-Attention for training, and the test datasets were predicted using a trained model. The actual and predicted values were compared, and the experimental results were divided based on the movement of the stock price into the following three groups: volatile, upward, and downward trends, as shown in Figs. 7–9.

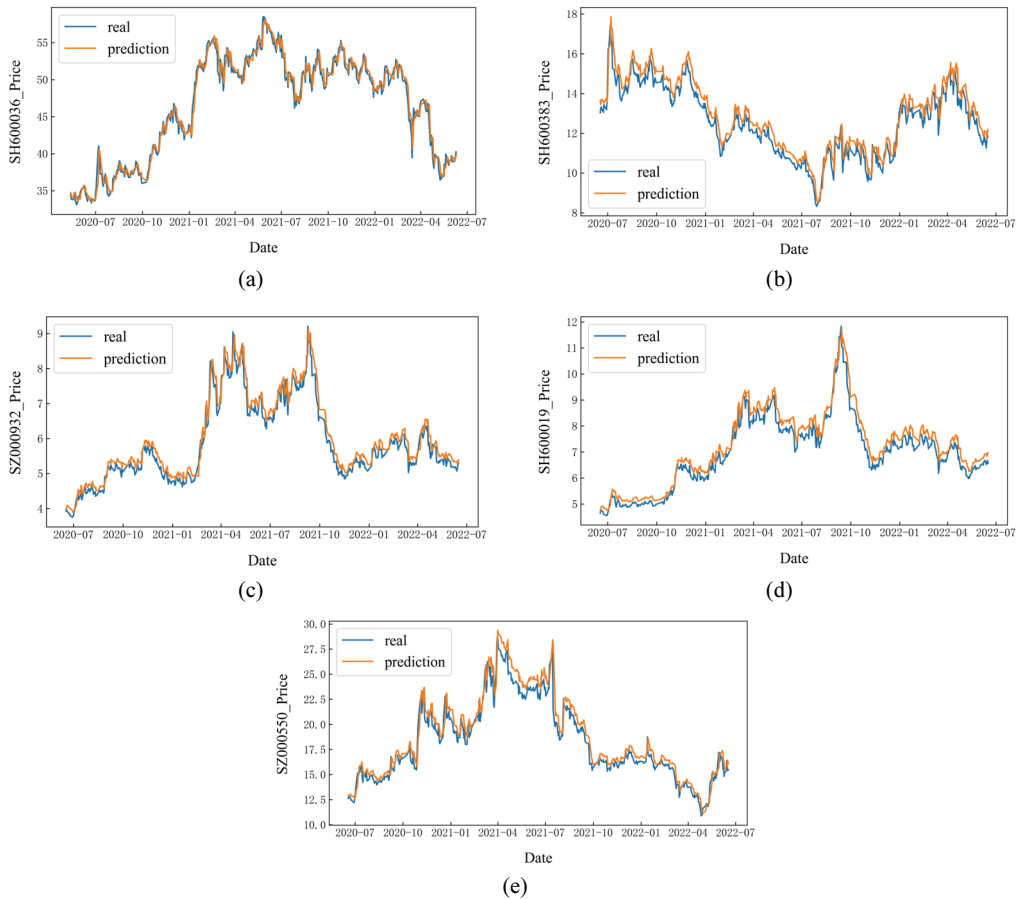


Fig. 7. Stock price forecast results under a volatile trend: (a) China Merchant Bank, (b) Gemdale Group, (c) Human Valin Steel, (d) Baoshan Iron & Steel, and (e) Jiangling Motors.

In a volatile trend, many factors influence the price, leading to rises and falls, making it difficult to predict. In Fig. 7(a)–7(e), the prediction results of XGBoost-LSTM-Attention are relatively stable in terms of volatility.

In an upward trend, stocks are in short supply, causing stock prices to keep rising. In Fig. 8(a)–8(h), the minimum and maximum stock price increases are 40% and 1000%, respectively, and the model-predicted prices follow the actual prices well.

In a downward trend, stocks are sold off and prices continue to fall. Fig. 9 shows that the model predicted price results are relatively close to the actual prices.

In Figs. 7–9, the prediction curve of XGBoost-LSTM-Attention fits well with the actual price curve and can reflect the actual price trend. The stock prediction experiment demonstrates that our model can accurately predict stocks in different industries, and the prediction results are relatively stable.

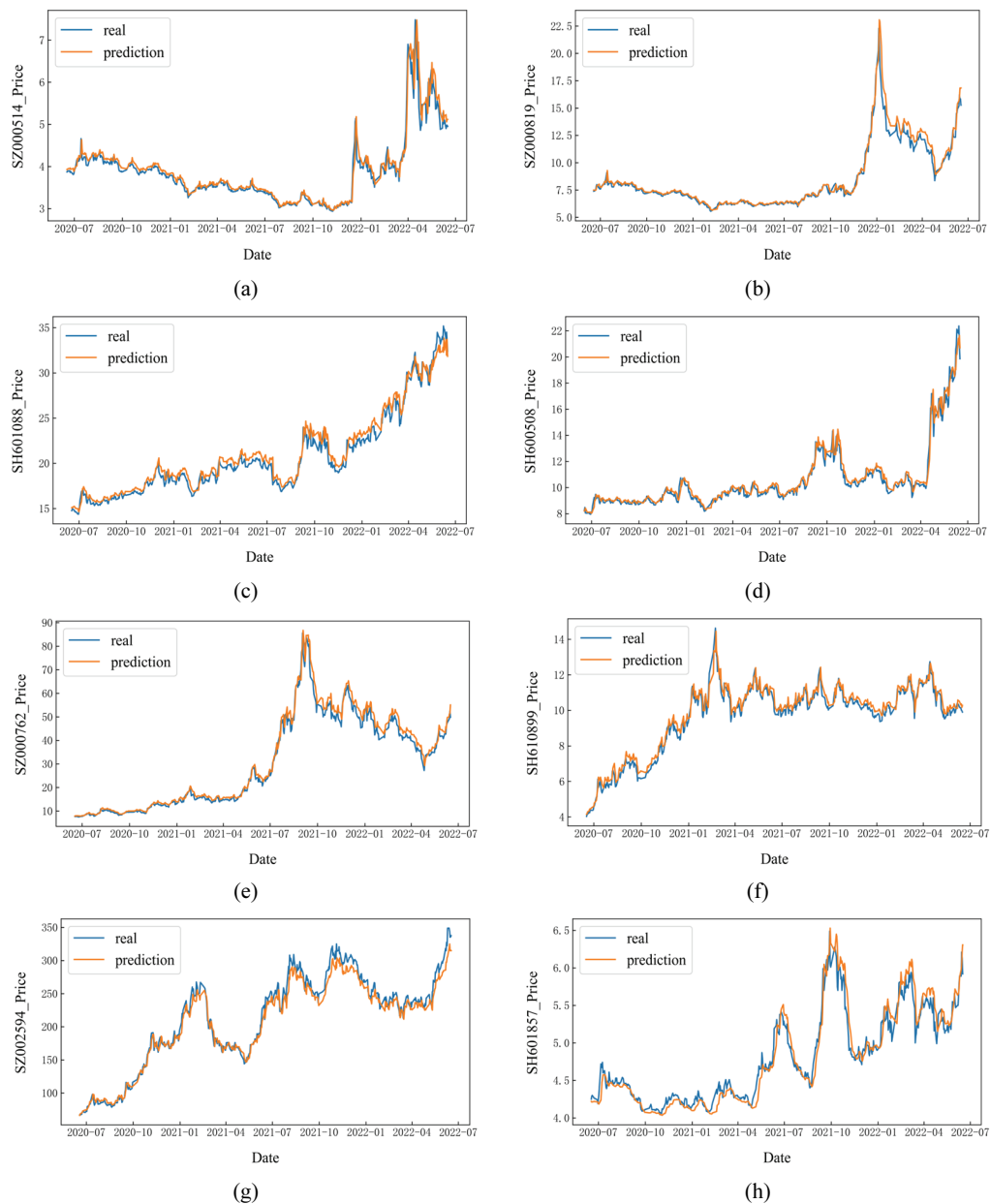


Fig. 8. Stock price forecast results show upward trend: (a) Chongqing Yukaifa, (b) Yueyang Xingchang, (c) China Shenhua Energy, (d) Shanghai Energy, (e) Tibet Mining, (f) Zijin Mining, (g) BYD, and (h) Petro China.

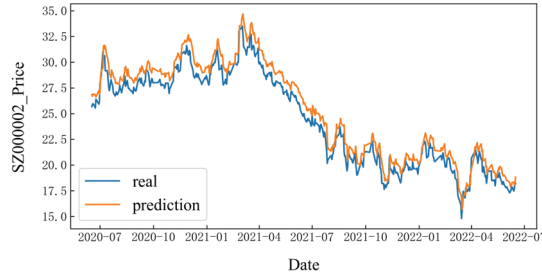


Fig. 9. Stock price forecast result of China Vanke shows a downward trend.

In the comparison experiments, we trained GRU-Attention, CNN-LSTM, CNN-BiLSTM-AM, and XGBoost-LSTM-Attention using the leading stock dataset. Based on the experimental results of all the models, the generalization errors of each model can be calculated. The results of the evaluation criteria of each model are listed in Tables 5 and 6, where relatively excellent results are observed for three of the four test datasets. Among the four models, the R^2 of the XGBoost-LSTM-Attention model is the highest; RMSE is relatively small; and ACC is relatively high. Similarly, the GRU-Attention model also exhibits good performance. The R^2 , RMSE, and ACC of the GRU-Attention model increased by 0.3%, 5%, and 9% on the Baoshan Iron & Steel test datasets, respectively, compared to the XGBoost-LSTM-Attention model. Comparing CNN-BiLSTM-AM with CNN-LSTM, the RMSE decreased by 0.082, 0.108, 0.101, and 0.086 for the four test datasets. Thus, CNN-BiLSTM-AM, GRU-Attention, and XGBoost-LSTM-Attention exhibit better performance than the CNN-LSTM model. This is because the neural network models based on attention can effectively study the correlation between stock information, thus reducing forecast error, which improves prediction performance.

Table 5. Evaluation metrics of different models on China Shenhua and Baoshan Iron & Steel datasets

Model	SH601088			SH600019		
	R^2	RMSE	ACC (%)	R^2	RMSE	ACC (%)
CNN-LSTM	0.927	0.111	60.86	0.876	0.141	59.24
GRU-Attention	0.924	0.014	60.38	0.944	0.018	67.58
CNN-BiLSTM-AM	0.888	0.029	57.34	0.891	0.033	61.84
Proposed	0.978	0.014	61.07	0.941	0.019	61.95

Table 6. Evaluation metrics for different models on the Zijin Mining and Petro China datasets

Model	SH601899			SH601857		
	R^2	RMSE	ACC (%)	R^2	RMSE	ACC (%)
CNN-LSTM	0.898	0.115	62.31	0.766	0.117	59.42
GRU-Attention	0.933	0.007	66.80	0.928	0.027	63.35
CNN-BiLSTM-AM	0.437	0.014	53.41	0.866	0.031	61.49
Proposed	0.962	0.008	67.70	0.931	0.021	66.87

In Tables 7 and 8, for the same input conditions, the evaluation results of our model are significantly different from those of the other three models. For the three test datasets, the R^2 , RMSE, and ACC values of our model were 0.955, 0.008, and 65.21%; 0.974, 0.003, and 67.08%; and 0.984, 0.014, and 73.08%, respectively. This demonstrates that the prediction performance of the model can be improved by

processing the original stock data.

In general, the XGBoost-LSTM-Attention model had the best generalization ability among the four models, with an average RMSE of 0.012, an average R^2 of 0.96, and an average ACC of 66.1%, indicating that the forecasting performance can be improved by processing the original data and by combining LSTM with the AM to accurately predict stocks in different industries. For the other three models, based on the results, they may not be applicable to stock markets in developing countries.

Table 7. Evaluation metrics for different models on China Merchants Bank and BYD datasets

Model	SH600036			SZ002594		
	R^2	RMSE	ACC (%)	R^2	RMSE	ACC (%)
CNN-LSTM	0.287	0.015	51.96	0.902	0.117	59.21
GRU-Attention	0.359	0.088	49.06	0.402	0.336	55.27
CNN-BiLSTM-AM	0.324	0.015	50.93	0.854	0.085	54.03
Proposed	0.955	0.008	65.21	0.974	0.003	67.08

Table 8. Evaluation metrics for different models on the China Vanke datasets

Model	SH600036		
	R^2	RMSE	ACC (%)
CNN-LSTM	0.927	0.111	59.42
GRU-Attention	0.957	0.134	63.76
CNN-BiLSTM-AM	0.945	0.029	62.73
Proposed	0.984	0.014	73.08

5. Conclusion

In this study, we focused on two tasks: first, we proposed a new model, XGBoost-LSTM-Attention, for stock prediction in the China A-share market. Subsequently, the XGBoost-LSTM-Attention model was used to predict stocks in different industries, and the results were compared with those of existing models.

Specifically, in terms of data input, the XGBoost algorithm was used to extract the important features of the original data, thereby reducing market noise and redundant feature information. In terms of the network structure, an LSTM-based attention unit was combined with the BiLSTM prediction unit to improve prediction accuracy. In the experiments comparing the CNN-LSTM, GRU-attention, and CNN-BiLSTM-AM models with our XGBoost-LSTM-Attention model, the proposed model had the highest average R^2 , lowest average RMSE, and highest ACC, indicating that the model is effective for stock price prediction. XGBoost-LSTM-Attention can be used as an important reference for investors when investing in stocks. In the future, we will introduce sentiment analysis to predict the stock market more comprehensively.

Conflict of Interest

The authors declare that they have no competing interests.

Funding

None.

References

- [1] X. Liu, D. Xiong, Z. Latif, and Y. Lu, "A comprehensive research of models between Internet development and China's economic growth," in *Proceedings of 2020 International Signal Processing, Communications and Engineering Management Conference (ISPCEM)*, Montreal, Canada, 2020, pp. 58-62. <https://doi.org/10.1109/ISPCEM52197.2020.00017>
- [2] X. Liu, Z. Latif, D. Xiong, S. K. Saddozai, and K. U. Wara, "Mean-VaR portfolio: an empirical analysis of price forecasting of the Shanghai and Shenzhen stock markets," *Journal of Information Processing Systems*, vol. 15, no. 5, pp. 1201-1210, 2019. <https://doi.org/10.3745/JIPS.04.0135>
- [3] X. Liu, Z. Latif, and Y. Lv, "A risk measurement by using mean-variance-kurtosis hybrid multi-objective portfolio optimization model," in *Proceedings of 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, Nanjing, China, 2018, pp. 843-847. <https://doi.org/10.1109/CSCWD.2018.8465263>
- [4] Y. Kara, M. A. Boyacioglu, and O. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul Stock Exchange," *Expert systems with Applications*, vol. 38, no. 5, pp. 5311-5319, 2011. <https://doi.org/10.1016/j.eswa.2010.10.027>
- [5] Y. Lin, H. Guo, and J. Hu, "An SVM-based approach for stock market trend prediction," in *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, USA, 2013, pp. 1-7. <https://doi.org/10.1109/IJCNN.2013.6706743>
- [6] C. S. Vui, G. K. Soon, C. K. On, R. Alfred, and P. Anthony, "A review of stock market prediction with artificial neural network (ANN)," in *Proceedings of 2013 IEEE International Conference on Control System, Computing and Engineering*, Penang, Malaysia, 2013, pp. 477-482. <https://doi.org/10.1109/ICCSCE.2013.6720012>
- [7] M. Qiu, Y. Song, and F. Akagi, "Application of artificial neural network for the prediction of stock market returns: the case of the Japanese stock market," *Chaos, Solitons & Fractals*, vol. 85, pp. 1-7, 2016. <https://doi.org/10.1016/j.chaos.2016.01.004>
- [8] H. Yu, R. Chen, and G. Zhang, "A SVM stock selection model within PCA," *Procedia Computer Science*, vol. 31, pp. 406-412, 2014. <https://doi.org/10.1016/j.procs.2014.05.284>
- [9] N. Naik and B. R. Mohan, "Intraday stock prediction based on deep neural network," *National Academy Science Letters*, vol. 43, no. 3, pp. 241-246, 2020. <https://doi.org/10.1007/s40009-019-00859-1>
- [10] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in *Proceedings of 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Udupi, India, 2017, pp. 1643-1647. <https://doi.org/10.1109/ICACCI.2017.8126078>
- [11] K. Chen, Y. Zhou, and F. Dai, "A LSTM-based method for stock returns prediction: a case study of China stock market," in *Proceedings of 2015 IEEE International Conference on Big Data (Big Data)*, Santa Clara, CA, USA, 2015, pp. 2823-2824. <https://doi.org/10.1109/BigData.2015.7364089>
- [12] L. C. Cheng, Y. H. Huang, and M. E. Wu, "Applied attention-based LSTM neural networks in stock prediction," in *Proceedings of 2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA, 2018, pp. 4716-4718. <https://doi.org/10.1109/BigData.2018.8622541>

- [13] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," *PLOS One*, vol. 14, no. 2, article no. e0212320, 2019. <https://doi.org/10.1371/journal.pone.0212320>
- [14] J. M. T. Wu, Z. Li, N. Herencsar, B. Vo, and J. C. W. Lin, "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators," *Multimedia Systems*, vol. 29, no. 3, pp. 1751-1770, 2023. <https://doi.org/10.1007/s00530-021-00758-w>
- [15] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Computing and Applications*, vol. 33, no. 10, pp. 4741-4753, 2021. <https://doi.org/10.1007/s00521-020-05532-z>
- [16] M. C. Lee, "Research on the feasibility of applying GRU and attention mechanism combined with technical indicators in stock trading strategies," *Applied Sciences*, vol. 12, no. 3, article no. 1007, 2022. <https://doi.org/10.3390/app12031007>
- [17] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785-794. <https://doi.org/10.1145/2939672.2939785>
- [18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [19] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, no. 1, pp. 97-136, 1980. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- [20] S. Wang, F. Xu, W. Ge, X. Chen, and M. Kong, "The study on stock selection based on deep learning and the Billboard data," in *Proceedings of 2021 2nd International Conference on Big Data Economy and Information Management (BDEIM)*, Sanya, China, 2021, pp. 209-212. <https://doi.org/10.1109/BDEIM55082.2021.00049>



Zhiyong Yang <https://orcid.org/0000-0003-4399-5557>

He received his M.S. and Ph.D. degrees from Chongqing University in 2009 and 2013, respectively. He is currently a full professor and vice president at the Chongqing Vocational Institute of Engineering, Chongqing, China. His current research interests include artificial intelligence and big data processing and analysis.



Yuxi Ye <https://orcid.org/0000-0003-3801-9404>

He is currently an undergraduate student at the School of Computer and Information Science at Chongqing Normal University. His research interests include artificial intelligence and deep learning.



Yu Zhou <https://orcid.org/0000-0001-8329-8532>

She received her B.S. from Lanzhou University of Finance and Economics in 2005 and M.S. degree from Chongqing University in 2011. She is currently a professor at the Chongqing Vocational Institute of Engineering, Chongqing, China. Her current research interests include data processing and analysis.