A New Deep Network Model for Stock Price Prediction

Min Liu, Hui Sheng, Ningyi Zhang, Yu Chen, and Longjun Huang(✉)

School of Software, Jiangxi Normal University, Nanchang, China
{liumin1,202140100843,zny,003721}@jxnu.edu.cn,
chenyu7825@163.com

Abstract. Over the recent past, many stock price prediction models that rely on deep neural networks have been developed. However, each has unique characteristics that cause variations in performance between models. With the existing deep neural network models, we propose a novel deep neural network-based stock price prediction model in this paper in order to predict stock prices more accurately. Specifically, this paper presents a method for extracting stock price series data based on the auto-encoder (AE) technique, which has strong non-smoothness and non-linear characteristics. Furthermore, the bi-directional long short-term memory (BiLSTM) module is imported as the primary unit structure in AE so that the historical and future important information of stock price series data can be sufficiently mined. Attention mechanisms are also investigated to make the extracted features more valuable for predicting stock prices. Lastly, the prediction is implemented by multi-layer, fully connected network work. The prediction results of the proposed method on two stock datasets are more prominent than other methods.

Keywords: Stock price prediction · Auto-encoder · Bi-directional LSTM · Attention mechanism

1 Introduction

The rapidly developing Internet technology has allowed a huge mass of financial information to be collected from various financial websites and has had a large impact on our daily lives. How to make full use of financial data to provide investors with valuable information for decision-making has become a significant task during the era of big data. The stock market is an important place in the emerging financial activities in the increasingly active financial market. It offers a high rate of return, attracts all types of investors, and allows for forecasting through financial data, thus reducing the risk of investment decisions [1]. Nevertheless, there are many factors that can affect the movement of stock prices, such as the macro-economic development of the country, the formulation of relevant laws and regulations, the operating and financial conditions of the company, the psychology of investors, the guidance of public opinion, and so on. The above reasons make stock prices have high irregularity and volatility, therefore there is a great uncontrollable risk for investors to invest in stocks. To enable investors

Y. Xu et al. (Eds.): ML4CS 2022, LNCS 13657, pp. 413–426, 2023.
https://doi.org/10.1007/978-3-031-20102-8_32

to reduce investment risks and make correct investment decisions, it is of great practical significance to study the technology or quantitative method [2] of stock price prediction.

At present, many stock price prediction models [2] have been developed and achieved excellent performance. There are many econometric statistical based traditional models, such as the autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), linear exponential smoothing (LES), generalized autoregressive conditional heteroscedasticity (GACH), which aim to explore the relationship among variables to find the best prediction result. Moreover, the above models still have some limitations. First, these methods rely heavily on the assumptions of linearly structured models, making it difficult to capture non-linear changes in stock prices. These methods assume that changes in the data are invariant, while financial time series contain features [3] such as high noise, time variation, and dynamics.

To address the above issues, this paper uses a machine learning approach to perform nonlinear analysis of financial time series. Artificial neural network (ANN) differs from econometric statistical models, which does not require a rigorous model structure and additional assumptions. In addition, it has been extensively applied to financial time series due to its powerful nonlinear mapping capability and generalization properties [3, 9, 10].

In latest years, long short-term memory (LSTM) networks are frequently utilized in time series prediction tasks [4, 5], and problems such as gradient disappearance or gradient explosion, which exist in recurrent neural networks (RNN) [6], can be overcome by LSTM using memory units and gating units. To make full use of historical and future information, a bi-directional LSTM (BiLSTM) [7] network is developed, which contains both forward LSTM and backward LSTM, and its final prediction result is better than LSTM.

The attention mechanism [8] can utilize finite resources to quickly filter out valuable target information from the vast amount of information. The attention mechanism originates from human vision, where the human eye can rapidly scan the target area to acquire helpful information and repress other meaningless details speedily. Currently, attention mechanisms have been broadly applied to computer vision, natural language processing, and speech recognition, and have succeeded in studies related to time series.

Due to the successful use of deep learning and attention mechanism in image analytics [11, 12], a prediction model based on time series is developed in this paper for predicting the stock closing price. First, BiLSTM is integrated into AE to extract the characteristics of the stock and combine them with the characteristics of the attention mechanism algorithm to finally obtain a deep density network for predicting stock prices. In the experimental section, the proposed model is compared with other models on two stock data sets, Shanghai Composite Index and CSI 300, to test the validity of the presented model.

2 Related Works

Recently, a variety of stock price prediction methods have put forward, which includes traditional machine learning and deep learning methods. Multiple machine learning methods includes k-nearest neighbor (KNN), support vector machine (SVM), support vector regression (SVR), random forest (RF), general algorithm (GA) and neural network (NN) have been explored and fused for stock price prediction [13, 14]. For example, Nayak et al. [15] combined SVM and KNN to implement the Indian stock market forecasting. Then, the weighted SVM is combined with KNN for the development trend of Chinese stock market [16]. Zhang et al. [17] fused AdaBoost, GA, and probabilistic SVM to predict stock and obtained better prediction performance. Picasso et al. [18] integrates RF, SVM and NN to predict the stock trend.

Deep learning is an essential offshoot of machine learning that extracts higher-level abstract characteristics for data representation and is broadly adopted in image processing and computer vision [19]. Lately, financial time series such as stock price data are analyzed through deep learning methods [20, 21]. For instance, Rather et al. [22] combined the traditional machine learning methods (i.e., ARMA and LSE) and recurrent neural networks (RNN) to realize stock return prediction. Stock price prediction using the LSTM model can address the gradient vanishing and gradient exploding problems in RNN [23]. Subsequently, many methods have been developed that combine traditional machine learning techniques with LSTMs to analyze financial series data. For example, Kim et al. [24] combined LSTM with GACH model to determine the fluctuation of stock price. Li et al. [25] combined feature selection and LSTM method to build a prediction model. Zhao et al. [6] integrated AM into RNN to propose two different prediction models, which can select and pay attention to the key information of stock data.

Recently, many CNN-based methods have been proposed to predict stock trends and achieve good results [26, 27]. This is because of CNN has two characteristics, i.e., local perception and parameter sharing, which lead to the reduction of parameters. For example, Sezer et al. [28] first obtained the 2D images converted from stock indicators, and secondly designed a new method to make predictions through CNN. Moreover, 2D images were obtained from time series data using the gram angle field technology, and then the U.S. market trends were predicted through the integrated learning framework of CNN [29]. Since the temporal data contains noise, the sequence reconstruction method is used to denoise it first, and then in order to predict the stock price, the spatial structure is extracted from the denoised data through the CNN model [30]. The description of trading behavior patterns have been defined by utilizing three matrices, including transaction, sale and purchase quantity matrices. Then, in order to effectively extract the deep features of transaction behavior, a CNN model is explored [31].

3 Preliminary

The model presented in this paper is designed by a comprehensive analysis considering self-encoders, bi-directional long and short-term memory networks, and attention mechanisms. In that section, we will briefly outline the above approach.

Autoencoder is a category of artificial neural networks used in semi-supervised and unsupervised learning, and its function is to acquire representations of the input information by taking it as a learning target. The role of the encoder is to encode the high-dimensional input into low-dimensional hidden variables, thus forcing the neural network to learn the most informative features; the role of the decoder is to restore the hidden variables in the hidden layer to the initial dimension, and the optimal state of the autoencoder is that the output of the decoder can be ideally or approximately recovered from the original input.

LSTM is a kind of temporal recurrent neural network specially designed to solve the long-term dependence problem in recurrent neural networks. The LSTM has two kinds of gates: the input gate, the forgetting gate, and the output gate. The information of the previous state is input to the LSTM as input. First, the LSTM receives the input information and calculates the value of the input gate, which is used to control the current data input's effect on the memory unit's value. All gate calculations are influenced by the current input data and the last moment LSTM cell output value, in addition to the previous memory cell value. Then, the forgetting states information, which information coming from the previous cell needs to be discarded from the cell state. Then the state information is updated. It is decided which new information in the cell state needs to be stored. Finally, the new information is output through the output gate. Due to the unique design structure of LSTM, LSTM is suitable for processing and predicting important events with very long intervals and delays in the time series. The two-way LSTM is an extension of the LSTM model where two LSTMs are applied to process the input data, divided into a forward and a backward layer. In the forward layer, one LSTM is used for the input sequence, and in the backward layer, the reverse form of the input sequence is fed into the LSTM model. Two applications of the LSTM lead to improved learning of long-term dependencies, thus improving the model's accuracy.

The attention mechanism has two main aspects, deciding which part of the input needs to be attended to and allocating limited information processing resources to the essential parts. In this paper, we use channel attention. Firstly, the input feature map is subjected to a global average pooling operation, followed by a 1-dimensional convolution operation with a convolution kernel of size k, and the weights of each channel are obtained after the activation function. The weights are multiplied with the corresponding elements of the original input feature map to obtain the final output feature map. The attention mechanism used in this paper is effortless in thought and operation and has minimal impact on the network processing speed.

4 The Proposed Method

Deep Neural Networks are used in combination with a large number of methods to predict stock prices, but there are some differences on the performance of different models because the stock prices are affected by many factors. This paper presents a new deep network prediction model to forecast stock prices more accurately. Initially, the Auto-Encoder is introduced in the model to extract the effective features of the stock data to deal with the stock price series due to its strong non-stationary and non-linear characteristics. Moreover, the BiLSTM module is considered as the main unit structure of the auto-encoder, which can fully mine the important information of historical data and future data for stock price series data. Then, attention theory is introduced to feature extraction from stock series data, which allows for a more accurate prediction of stock prices. Finally, a multi-layer fully connected network is employed to accomplish the prediction. The general structure of the model is shown in Fig. 1.

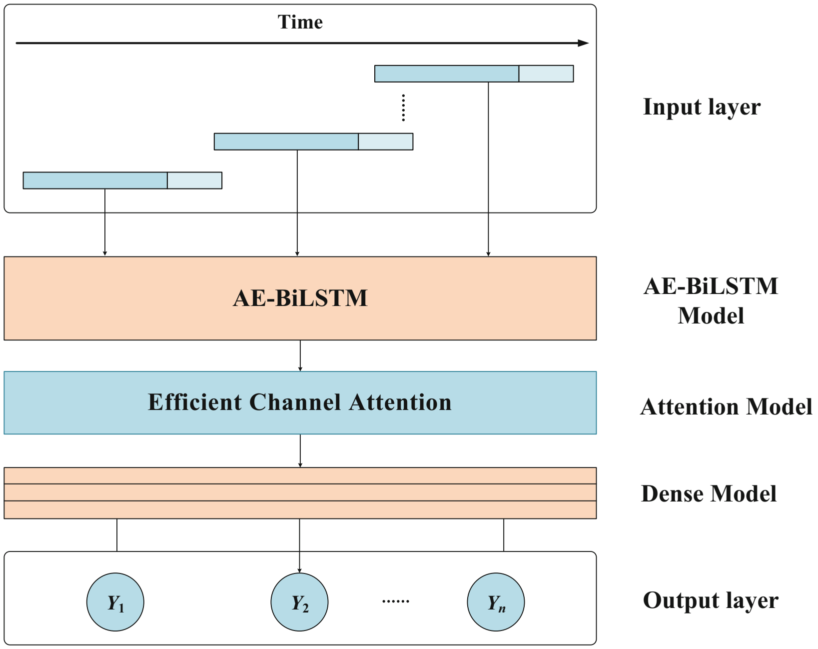

Fig. 1. The general structure of the proposed model

The auto-encoder (AE) is an unsupervised neural network model that consists of two parts: encoding and decoding [32]. The encoder is mainly responsible for learning the hidden features from the input sample data, while the decoder is mainly responsible for reconstructing the original input data from the hidden features. Assuming D-dimensional samples $x^{(n)} \in R^D,\ n = 1, 2, ..., N$, the AE maps the data to the feature space and obtains an encoding $z^{(n)} \in R^d,\ n = 1, 2, ..., N$ for each sample. It is expected that this set of encodings can reconstruct the original samples. The objective function can be defined as follows:

$$\varphi = \min \sum_{n=1}^{N} ||x^{(n)} - g(f(x^{(n)}))||_F^2 \tag{1}$$

where $f(\bullet)$ and $g(\bullet)$ denote as activation function. The three-layer neural network is the simplest auto-encoder and shown in Fig. 2, consisting of an input layer, a hidden layer, and an output layer. The amount of neurons in both input and output layers is identical. The encoding process is performed between the input layer and the hidden layer, the decoding process is performed between the hidden layer and the output layer, and the layers are fully connected to each other.

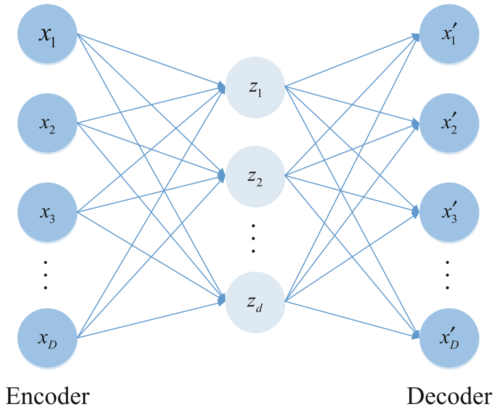

Fig. 2. The structure of auto-encoder with three-layer network

Considering the characteristics of stock time series data, BiLSTM as a unit is introduced into the auto-encoder, and the structure of BiLSTM is shown in Fig. 3.

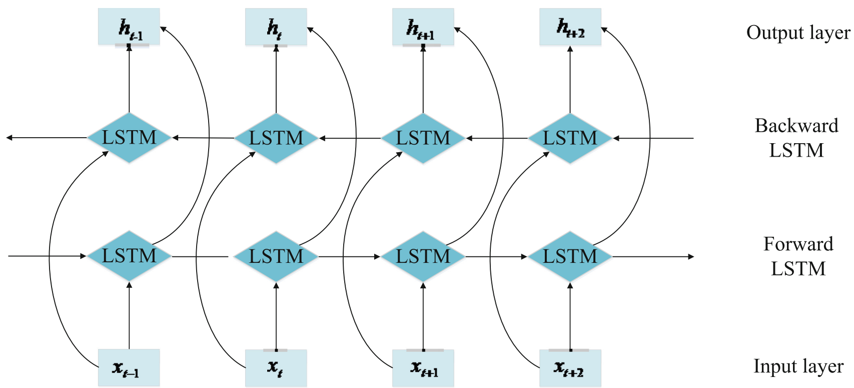

Fig. 3. The structure of BiLSTM

The channel attention mechanism (CAM) is one of the most widely used methods in computer vision, such as image classification and image segmentation, and achieves better performance. Nevertheless, in order to achieve better model performance, most methods are devoted to designing more complex attention models, which will inevitably increase the computational complexity of the model. To prevent model overfitting and reduce computational effort, Efficient Channel Attention (ECA) as a lightweight and low-complexity module is integrated into our proposed method [33]. ECA learns the relevance of each channel and generates different weights according to the magnitude of the correlation. In multidimensional stock price statistics, ECA assigns more weight to the most crucial components and generates less weight in the irrelevant components. This allows the network model to focus on more valuable information and enhances the sensitivity to key functions. The ECA structure is shown in Fig. 4.

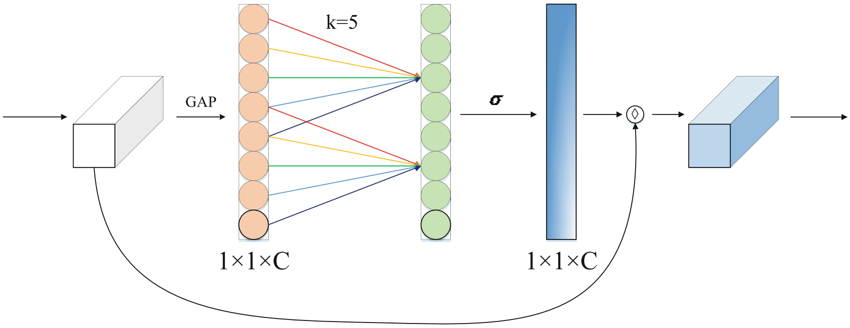

Fig. 4. Architecture of ECA model

From the Fig. 4, after channel global average pooling (GAP), local cross-channel interactions are captured by each channel of ECA and its *k* neighbors. At the same time, the channel weights are generated using a one-dimensional fast convolution of size *k*. To avoid manually adjusting *k* by cross-validation, the mapping of channel dimension C adapts to determine the value of *k*, that is:

$$\boldsymbol{\omega} = \boldsymbol{\sigma}(\boldsymbol{C}1\boldsymbol{D}_{\boldsymbol{k}}(\boldsymbol{y})) \tag{2}$$

where C1D represents 1-D convolution. Since the kernel diameter *k* of the one-dimensional convolution is in proportion to the channel dimension *C*, the correspondence is as follows:

$$\boldsymbol{C} = \boldsymbol{\varphi}(\boldsymbol{k}) = 2^{(\boldsymbol{\gamma} \times \boldsymbol{k} - \boldsymbol{b})} \tag{3}$$

Thus, the kernel size *k* can be adaptively determined according to Eq. (4) for a given channel dimension *C*.

$$\boldsymbol{k} = \boldsymbol{\psi}(\boldsymbol{C}) = \left|\frac{\log_2(\boldsymbol{C}) + \boldsymbol{b}}{\boldsymbol{\gamma}}\right|_{odd} \tag{4}$$

where $|t|_{odd}$ denotes the odd number closest to t. In all experiments, γ and b are set to 2 and 1, respectively. Markedly, non-linear mapping increases the interaction range of higher dimensional channels and shortens the interaction range of lower dimensional channels

Finally, using the fully connected (Dense) model, the correlations among the features are obtained and mapped to the outcome space. In the proposed AE-BiLSTM-ECA's network model, to solve the nonlinear problem better, a multi-layer full connection layer is introduced, which can achieve accurate prediction according to many different factors. The structure diagram of the fully connected layer with three layers is represented in Fig. 5.

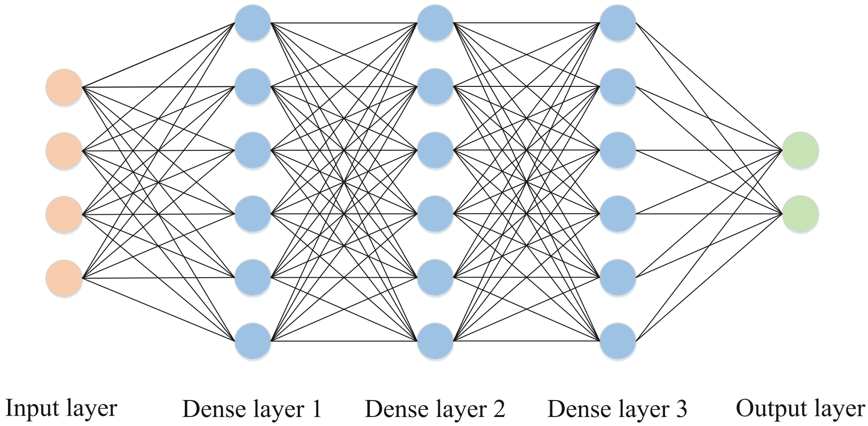

Fig. 5. The structure diagram of dense model

5 Experiments

5.1 Dataset Description and Pre-processing

The experimental data were obtained from NetEase Finance, including Shanghai Stock Composite Index (referred to as SSCI, stock code: 000001), and CSI 300 (stock code: 399300).The data of each stock contains seven characteristics, such as closing price (CP), highest price (HP), lowest price (LP), opening price (OP), previous day's closing price (PCP), up or down amount (UDA), and up or down rate (UDR). Since these seven attributes are obtained from actual stock market transactions and can reflect the most basic information about stock prices, they are selected as the inputs to the model. The specific information of the two stock data is summarized in Table 1. Tables 2 and 3 displays selected data from each stock as well as descriptive statistical information about the data, where descriptive statistics include mean, median, plurality, standard deviation, variance, maximum value, minimum value, and quantity for each attribute. We can clearly see that there are large variances and fluctuations in the stock data.

Table 1. Specific information on the two types of stock data

Stock	The time periods	Total data/group
SSCI	1990/12/20–2020/11/23	7304
CSI 300	2002/01/07–2021/03/17	4657

5.2 Evaluation Functions

The mean squared error (MSE) stands for the expected value of the square of the difference between the parameter estimate and the true value of the parameter. The degree of

Table 2. Partial information on SSCI stock data

Date	CP	HP	LP	OP	PCP	UDA	UDR
1990/12/20	104.390	104.390	99.980	104.300	99.980	4.410	4.411
1990/12/21	109.130	109.130	103.730	109.070	104.390	4.740	4.541
...	...	...	...	...	...	...	...
2000/12/28	2053.704	2061.051	2047.742	2054.517	2058.244	−4.540	−0.221
2000/12/29	2073.477	2073.878	2055.505	2055.828	2053.704	19.773	0.963
...	...	...	...	...	...	...	...
2020/11/20	3377.727	3380.149	3356.309	3359.597	3363.088	14.639	0.435
2020/11/23	3414.490	3431.653	3377.986	3384.104	3377.727	36.763	1.088
Average	1986.648	2004.091	1965.175	1985.333	1986.221	0.442	0.071
Median	1909.716	1927.174	1892.712	1911.063	1909.408	0.855	0.067
Mode	134.240	134.740	134.190	125.270	134.240	−0.460	0.997
Standard deviation	1072.439	1081.367	1059.720	1071.066	1072.504	40.963	2.439
Variance	1150125	1169355	1123006	1147182	1150266	1677.959	5.948
Minimum	104.390	104.390	99.980	104.300	99.980	−354.684	−16.394
Maximum	6092.057	6124.044	6040.713	6057.428	6092.057	649.500	105.269
Quantity	7304	7304	7304	7304	7304	7304	7304

variation of the data can be evaluated by calculating the MSE. If the MSE is smaller, it means that the prediction model describes the experimental data with better accuracy, and is calculated as below:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(X_t - X_t')^2 \tag{5}$$

The Root Means Squared Error (RMSE) represents the expected value of the squared error, which is the deviation between the predicted and actual values in the range of [0, $+\infty$). If the prediction result is consistent with the actual situation, the RMSE is 0, and the model can be called ideal. Thus the numerical result of RMSE varies with the predicted result. The formula for calculating RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(X_t - X_t')^2} \tag{6}$$

The Mean Absolute Error (MAE) refers to the average distance between the model predicted value and the true value of the sample. MAE can avoid the problem of errors canceling each other out and is used to express the severity of the deviation of the predicted values compared to the true values. A more petite MAE means that the deviation

Table 3. Partial information on CSI 300 stock data

Date	CP	HP	LP	OP	PCP	UDA	UDR
2002/01/07	1302.080	1302.080	1302.080	1302.080	1316.460	−14.380	−1.092
2002/01/08	1292.710	1292.710	1292.710	1292.710	1302.080	−9.370	−0.720
...	...	...	...	...	...	...	...
2012/01/04	2298.753	2365.988	2298.298	2361.499	2345.742	−46.989	−2.003
2012/01/05	2276.385	2316.657	2272.153	2290.780	2298.753	−22.368	−0.973
...	...	...	...	...	...	...	...
2021/03/16	5079.362	5084.309	5009.951	5054.409	5035.544	43.818	0.870
2021/03/17	5100.858	5123.545	5020.126	5062.771	5079.362	21.496	0.423
Average	2762.711	2785.604	2734.602	2760.160	2761.898	0.813	0.043
Median	2851.915	2888.093	2818.248	2848.155	2850.829	1.339	0.069
Mode	1221.760	1221.760	1221.760	1221.760	1221.760	−17.370	−0.487
Standard deviation	1187.877	1201.201	1171.218	1187.380	1187.571	52.682	1.653
Variance	1411051	1442883	1371753	1409871	1410325	2775	2.731
Minimum	818.033	823.860	807.784	816.546	818.033	−391.866	−9.240
Maximum	5877.202	5930.912	5815.609	5922.071	5877.202	378.179	9.390
Quantity	4657	4657	4657	4657	4657	4657	4657

of the predicted value from the actual value is minor. The formula of MAE is shown as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|X_t - X_t'\right| \tag{7}$$

Mean Absolute Percentage Error (MAPE), which represents the sum of each absolute error divided by the actual value, is one of the most commonly used metrics to assess prediction accuracy. In fact, it is the average of the error percentages. The formula of MAPE is shown as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|X_t - X_t'|}{|X_t'|} \tag{8}$$

where X_tX_t represents the predicted value and $X_t'X_t'X_t'$ represents the true value. The smaller the value of the above evaluation index, the more accurate the prediction result.

5.3 Results and Discussions

In this section, we will perform a calibration verification of the model described in this paper, CNN, LSTM, BiLSTM, CNN-LSTM, AE-LSTM, CNN-BiLSTM, AE-BiLSTM,

BiLSTM-ECA, AE-LSTM-ECA, CNN-LSTM-ECA are used and compared with the AE-BiLSTM-ECA model proposed.

Table 4. Forecast results of different models on the SSCI stock data

Model	MSE	RMSE	MAE	MAPE
CNN	8447.149	91.908	79.914	2.426
LSTM	4222.102	64.978	50.585	1.805
BiLSTM	2603.726	51.027	35.419	1.239
CNN-LSTM	2603.726	51.027	35.419	1.239
AE-LSTM	2590.940	50.901	37.692	1.296
CNN-BiLSTM	2321.235	48.179	32.289	1.129
AE-BiLSTM	2455.008	49.548	36.147	1.249
BiLSTM-ECA	2184.278	46.736	32.737	1.116
AE-LSTM-ECA	2025.349	45.004	29.624	1.025
CNN-LSTM-ECA	2200.705	46.911	32.353	1.142
AE-BiLSTM-ECA	**1935.398**	**43.993**	**28.940**	**1.019**

Table 5. Forecast results of different models on the CSI 300 stock data

Model	MSE	RMSE	MAE	MAPE
CNN	6218.092	78.855	63.981	1.843
LSTM	5809.153	76.217	58.679	1.662
BiLSTM	5091.610	71.356	52.119	1.446
CNN-LSTM	4905.472	70.039	52.457	1.431
AE-LSTM	4908.142	70.058	53.112	1.525
CNN-BiLSTM	4643.541	68.144	51.143	1.410
AE-BiLSTM	4288.830	65.489	48.640	1.335
BiLSTM-ECA	4161.203	64.507	46.453	1.289
AE-LSTM-ECA	3225.251	56.791	37.473	1.037
CNN-LSTM-ECA	4568.808	67.593	51.061	1.395
AE-BiLSTM-ECA	**3158.452**	**56.200**	**36.681**	**1.020**

Tables 4 and 5 shows the prediction results of two stock datasets using different methods. Because of the different amounts of data contained in each dataset and the volatility of the stock data, which leads to considerable differences in the data, the results obtained by applying the same evaluation metrics to the two datasets also differ significantly. According to the experimental results, we can draw the following conclusions:

1) For a single model (i.e., CNN, LSTM, and BiLSTM), initially, the CNN model has the lowest prediction performance because it does not utilize the historical stock price information more effectively. Thne, the BiLSTM utilizes bi-directional information from historical data, so the BiLSTM model is able to further improve the prediction ability compared to the LSTM.
2) Since CNNs and AEs are good at extracting effective features from data, combining them with traditional recurrent neural networks (LSTM and BiLSTM networks) can improve the predictive performance of the models. However, since AE combined with a single model is using the recurrent neural network as the core unit in the AE model, the model performance of AE combined with recurrent neural is higher than that of CNN combined with the recurrent neural network model.
3) The sensitivity of this network to the main features is improved due to the addition of the attention module ECA. Therefore the prediction effect of the network model incorporating the attention module is significantly better than the prediction effect of the original network model. The results show that the introduction of the attention module ECA in this paper helps the model to focus more on the impact on the stock price.
4) Of all methods compared, the AE-BilSTM-ECA model was optimal on all four indicators. In particular, the results of the proposed model are significantly better than the traditional single model (i.e. CNN, LSTM and BiLSTM). A recurrent neural network model that combines feature extraction with attention mechanism outperforms the model that contains only features or attention. The results show that the prediction of stock data using AE-BiLSTM-ECA method is correct.

6 Conclusions

In this paper, a novel deep network forecasting model, that is AE-BilSTM-ECA model, is presented to predict stock prices. The network which introduces the bi-directional LSTM module as the primary cell in the autoencoder can sufficiently mine the historical and future critical information of the stock price series data. And the ECA module is also implemented in this network to weight the extracted characteristics, which enables more reliable prediction of stock prices. The experiments in this paper are conducted on two stocks, SSE Composite Index, China Unicom and CSI 300. Among the ten comparative methods, the AE-BilSTM-ECA model performs the most optimally, and in particular, the results of this model are remarkably superior to the conventional single models (i.e., CNN, LSTM, and BiLSTM). The future research can explore whether the introduction of sparse self-attentiveness can enhance the prediction performance and effectiveness of this model.

Acknowledgment. This work is supported in part by grants from the National Natural Science Foundation of China (No. 62062040), the Outstanding Youth Project of Jiangxi Natural Science Foundation (No. 20212ACB212003), the Jiangxi Province Key Subject Academic and Technical Leader Funding Project (No. 20212BCJ23017).

References

1. Cavalcante, R.C., Brasileiro, R.C., Souza, V.L.F., et al.: Computational intelligence and financial markets: A survey and future directions. Expert Syst. Appl. **55**, 194–211 (2016)
2. Thakkar, A., Chaudhari, K.: Fusion in stock market prediction: a decade survey on the necessity, recent developments, and potential future direction. Information Fusion **65**, 95–107 (2021)
3. Sezer, O.B., Gudelek, M.U., Ozbayoglu, A.M.: Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Appl. Soft Comput. **90**, 106181 (2020)
4. Ding, G., Qin, L.: Study on the prediction of stock price based on the associated network model of LSTM. Int. J. Mach. Learn. Cybern. **11**(6), 1307–1317 (2019). https://doi.org/10.1007/s13042-019-01041-1
5. Baek, Y., Kim, H.Y.: ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. Expert Syst. Appl. **113**, 457–480 (2018)
6. Zhao, J., Zeng, D., Liang, S., Kang, H., Liu, Q.: Prediction model for stock price trend based on recurrent neural network. J. Ambient. Intell. Humaniz. Comput. **12**(1), 745–753 (2020). https://doi.org/10.1007/s12652-020-02057-0
7. Lu, W., Li, J., Wang, J., et al.: A CNN-BiLSTM-AM method for stock price prediction. Neural Comput. Appl. **33**(10), 4741–4753 (2020). https://doi.org/10.1007/s00521-020-05532-z
8. Niu, Z., Zhong, G., Yu, H.: A review on the attention mechanism of deep learning. Neurocomputing **452**, 48–62 (2021)
9. Hu, J., Zheng, W.: A deep learning model to effectively capture mutation information in multivariate time series prediction. Knowl.-Based Syst. **203**, 106139 (2020)
10. Torres, J.F., Hadjout, D., Sebaa, A., et al.: Deep learning for time series forecasting: a survey. Big Data **9**(1), 3–21 (2021)
11. Chong, E., Han, C., Park, F.C.: Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Syst. Appl. **83**, 187–205 (2017)
12. Nti, I.K., Adekoya, A.F., Weyori, B.A.: A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review **53**(4), 3007–3057 (2020)
13. Ahmed, N.K., Atiya, A.F., Gayar, N.E., et al.: An empirical comparison of machine learning models for time series forecasting. Economet. Rev. **29**(5–6), 594–621 (2010)
14. Henrique, B.M., Sobreiro, V.A., Kimura, H.: Literature review: Machine learning techniques applied to financial market prediction. Expert Syst. Appl. **124**, 226–251 (2019)
15. Nayak, R.K., Mishra, D., Rath, A.K.: A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. Appl. Soft Comput. **35**, 670–680 (2015)
16. Chen, Y., Hao, Y.: A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. Expert Syst. Appl. **80**, 340–355 (2017)
17. Zhang, X., Li, A., Pan, R.: Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine. Appl. Soft Comput. **49**, 385–398 (2016)
18. Picasso, A., Merello, A., Ma, Y., et al.: Technical analysis and sentiment embeddings for market trend prediction. Expert Syst. Appl. **135**, 60–70 (2019)
19. Shrestha, A., Mahmood, A.: Review of deep learning algorithms and architectures. IEEE access **7**, 53040–53065 (2019)
20. Han, Z., Zhao, J., Leung, H., et al.: A review of deep learning models for time series prediction. IEEE Sens. J. **21**(6), 7833–7848 (2019)
21. Li, A.W., Bastos, G.S.: Stock market forecasting using deep learning and technical analysis: a systematic review. IEEE access **8**, 185232–185242 (2020)

22. Rather, A.M., Agarwal, A., Sastry, V.N.: Recurrent neural network and a hybrid model for prediction of stock returns. Expert Syst. Appl. **42**(6), 3234–3241 (2015)
23. Bathla, G., Rani, R., Aggarwal, H.: Stocks of year 2020: prediction of high variations in stock prices using LSTM. Multimedia Tools and Applications, pp. 1–17 (2022)
24. 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**, 25–37 (2018)
25. Li, H., Hua, J., Li, J., et al.: Stock forecasting model FS-LSTM based on the 5G Internet of things. Wireless Communications and Mobile Computing, 2020 (2020)
26. Hoseinzade, E., Haratizadeh, S.: CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert Syst. Appl. **129**, 273–285 (2019)
27. Chen, Y., Fang, R., Liang, T., et al.: Stock price forecast based on CNN-BiLSTM-ECA Model. Scientific Programming, 2021 (2021)
28. Sezer, O.B., Ozbayoglu, A.M.: Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. Appl. Soft Comput. **70**, 525–538 (2018)
29. Barra, S., Carta, S.M., Corriga, A., et al.: Deep learning and time series-to-image encoding for financial forecasting. IEEE/CAA Journal of Automatica Sinica **7**(3), 683–692 (2020)
30. Wen, M., Li, P., Zhang, L., et al.: Stock market trend prediction using high-order information of time series. Ieee Access **7**, 28299–28308 (2019)
31. Long, J., Chen, Z., He, W., et al.: An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. Appl. Soft Comput. **91**, 106205 (2020)
32. Mohanty, D.K., Parida, A.K., Khuntia, S.S.: Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine. Appl. Soft Comput. **99**, 106898 (2021)
33. Guo, C., Szemenyei, M., Hu, Y., et al.: Channel attention residual u-net for retinal vessel segmentation. In: IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp. 1185–1189 (2021)