Journal of Computers Vol. 34 No. 1, February 2023, pp. 131-144
doi: 10.53106/199115992023023401010

Financial Forecasting Method for Generative Adversarial Networks Based on Multi-model Fusion

Pei-Guang Lin*, Qing-Tao Li, Jia-Qian Zhou, Ji-Hou Wang, Mu-Wei Jian, Chen Zhang

Department of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China
llpwgh@sina.com, tao5125801@163.com, 244656732@qq.com, 1176238504@qq.com,
jianmuwei@ouc.edu.cn, jason88131@hotmail.com

Received 14 April 2022; Revised 13 July 2022; Accepted 11 August 2022

Abstract. To some extent, stock prices can reflect national economic development and residents' living standards. However, current stock forecasts are mainly analyzed using the stock's own price. It is found that the exchange rate and technical indicators are closely related to the fluctuations of stocks and stock indexes. In this study, a generative adversarial network (GAN) financial forecasting method based on multi-model fusion is proposed, which introduces the exchange rate and technical indexes of stock and stock index into the input data, and combines the characteristics of its own highest and lowest price, the closing price of stock or stock index is predicted by using the generated antagonistic network. The methods are as follows: firstly, the exchange rate features are de-noised by wavelet transform technology, the overall trend of exchange rate features is extracted, and the dimensionality of technical indicators is reduced by Principal Component Analysis (PCA). Then, a convolutional neural network (CNN) is used in the generator to extract the local features of the input data, and the attention mechanism is used to improve the prediction effect of the model. Finally, the prediction was made by long short-term memory network (LSTM). Convolutional neural network and multilayer fully connected neural network are used in the discriminant, and gelu is used as activation function in the hidden layer. This study selects the data of Zhuhai Port (000507.SZ), Xiamen International Trade (600755.SH), Orient Venture (600278.SH) and Shanghai Stock Exchange Index (SSE, 000001.SH) in the export industry from January 1, 2013 to December 31, 2019 as the experimental data set. The exchange rate price between RMB and US dollar at the corresponding time is selected as the exchange rate feature. Compared with the best in-depth learning financial prediction comparison model, the results show that the generation confrontation network proposed in this paper has improved by 3% to 15% in each experimental data set.

Keywords: financial forecasting, generating adversarial network, deep learning

1 Introduction

Stock is an important way of investment in People's Daily life, and the fluctuation of the stock price has received a great deal of attention from both investors and researchers, and how to accurately predict stock prices has attracted many scholars' research. However, stock prices are characterized by nonlinearity and multiple features, and it is difficult for traditional statistical learning methods to make accurate predictions of stock prices. In recent years, artificial neural networks have been developed rapidly, and neural networks (e.g., convolutional neural networks, recurrent neural networks (RNN) have been widely used in various fields of life [1-2]. In the field of financial prediction, because neural networks can handle large amounts, multidimensional, and nonlinear data, deep learning methods have started to replace the traditional statistical learning methods.

The characteristics used in traditional stock forecasting models are based on the closing price, opening price, etc. of the stock itself. However, in stock analysis, fundamental analysis and technical analysis are also factors that investors and investment personnel need to consider. In this paper, exchange rate indicators in fundamental analysis and MA5 and MACD, which are commonly used in technical analysis, are selected as characteristics of stocks to enhance the prediction effect of the model. The exchange rate as a macroeconomic indicator and the link between the exchange rate and stock prices have always been the focus of economists' research. Experimental results show that the change of exchange rate price affects the rise and fall of stock price, where stock price is more sensitive to the fluctuation of exchange rate price in the short term. Technical indicators are used to analyze stock price movements by constructing different mathematical analysis models, mainly to analyze the rise and fall of stock prices in the short term.

* Corresponding Author

In addition, some researchers start from the stock prediction model [3-5] and use the generative confrontation network model to predict the stock price instead of the traditional LSTM network. The experimental results show that the stock prices generated by the generative adversarial network have better prediction results compared with the traditional financial prediction models.

Exchange rate price changes as an important indicator of a country's economy have an important impact on the national economy, even affecting the trend of stock prices. Sensoy and Tabak [6] analyzed the dependence between the stock market and exchange rate prices, and the experimental results showed that exchange rate prices were correlated with stocks. Zhang et al. [7] used the SVM model to prove that the exchange rate price could be applied to the index prediction. In this study, the exchange rate price of RMB and USD for 7 years from January 1, 2013 to December 31, 2019 is chosen as the exchange rate characteristic. In stock and stock index price forecasting, technical indicators are often an important reference indicator, and investors will take measures according to the changes of technical indicators. This study uses Ta-lib financial database to generate various technical indicators for stocks and stock indexes. Considering the large effect of exchange rate on import and export enterprises, chooses to use the export trade companies Zhuhai Port (000507.SZ), Xiamen International Trade (600755.SH), and Dongfang Venture (600278.SH) from January 1, 2013 to December 31, 2019 as the experimental dataset, and SSE index (000001.SH) was selected to verify the generality of the model.

Aiming at the above problems of the financial forecasting model, this chapter proposes a generative adversarial network financial forecasting method based on multiple features. In the input data, the exchange rate characteristics in basic analysis and several technical indicators in technical analysis are used to form the multivariate characteristics of stocks, so as to take into account the commonly used stock analysis methods as much as possible. In the prediction model, generative adversarial network is used as the prediction model in this chapter. Different from previous financial prediction models based on generative adversarial network, hybrid network is used as the generator and discriminator model in this method, in which attention-CNN-LSTM model is used as the generator model. One-dimensional convolutional neural network and fully connected network are used as discriminator models. In the generator model, the convolutional neural network extracts the features of the input data, adds the Attention layer in the middle of the two-layer LSTM, and assigns different weights to the output of each timestep hidden layer in the LSTM network to improve the prediction effect of the model. In the stage of experimental results, this paper compares with the financial prediction model commonly used at present and two kinds of generative adversarial network prediction models to verify the prediction effect of this model.

The main contributions of this study are as follows:

1) To solve the problem of poor prediction effect of single model, this study proposes a generative adversarial network model based on multi-model fusion. By comparing with other traditional financial forecasting models, the proposed model has better prediction effect in financial forecasting. In this study, we use generative adversarial network as a prediction model. The one-dimensional (1D) CNN network is used in the generator to extract stock features, and the LSTM network is used for prediction, and the attention mechanism is added behind the hidden layer of LSTM to improve the prediction effect. The discriminator model is composed of 1D CNN network and multiple fully connected networks, and *gelu* activation function is used in the hidden layer of each discriminator. Compared with other traditional activation functions, it is proved that *gelu* activation function has a positive effect on improving the prediction effect of the model.

2) This study introduces exchange rate characteristics and technical indicators into the input data. In this study, the input data are divided into two types: directly obtained data (open price, close price, high price, low price) and related data (exchange rate characteristics, stock technical indicators). For exchange rate indicators, this study uses wavelet denoising to extract the trend of exchange rate prices. For technical indicators, several commonly used technical indicators are generated through Ta-lib financial database, and PCA is used to reduce the dimensionality of technical indicators in order to extract the important features and reduce redundancy.

The remainder of this paper is organized as follows.

The first part introduces the relevant background of financial forecasting and the innovation of the model proposed in this paper.

The second part introduces the relevant methods in the field of financial forecasting.

The third part introduces the main structure of the model proposed in this paper.

The fourth part introduces the data set used in the experiment, the experimental setup and the evaluation indicators used in this paper, and analyzes and discusses the experimental results.

The fourth part is the summary.

2 Related Work

Financial data is a complex and high-dimensional form of time series. By dimensionality reduction of financial data, we can discover the essential features within the data, thus reducing data redundancy and improving computation speed. In the field of financial data, the traditional way of dimensionality reduction is mainly the principal component analysis method, which reduces the dimensionality of data in linear way. In 2017 Waqar et al. [8] used the New York Stock Exchange, London Stock Exchange and Karachi Stock Exchange to analyze the accuracy of the linear regression model after PCA reduction. The experiment showed that the PCA method could improve the model performance and prediction accuracy.

In the field of financial data prediction, prediction models can be broadly divided into traditional statistical models, machine learn-based models and deep learn-based models. Traditional statistical models are mainly based on Auto Regressive model (AR), Moving Average model (MA), Autoregressive Moving Average model (ARMA), and Autoregressive Integrated Moving Average model (ARIMA) models, which mainly to build linear models to fit the trend of a time series. In 2014 Ayodele et al. [9] proposed the use of ARIMA models to forecast stock data from the New York Stock Exchange and the Nigerian Stock Exchange, and the results showed that ARIMA models have great short-term forecasting potential. However, the traditional statistical models are mainly based on linear models, and financial data are characterized by multiple features and nonlinearity, it is difficult for traditional statistical models to make accurate forecasts of financial data.

Researchers are using machine learning methods to predict stock prices. Machine learning models are mainly based on support vector machines (SVM), random forests, etc. Mei et al. [10] in 2018 combined ARIMA and SVM models using IBM stock data from January 2, 1962, to November 10, 2017, for prediction. And the results showed that the accuracy of the model reached 96.10%, proving that the combination of machine learning and the traditional statistical model approach has good results in stock prediction. Xiao et al. [11] analyzed the shortcomings of current stock market prediction methods and support vector machines, and proposed a method to improve the prediction effect of the model by combining models, i.e., using ARI-MA-LS-SVM model for stock market prediction. Compared with the traditional single prediction model, the combined model based on ARI-MA-LS-SVM has better prediction fitting effect and better practical performance. With the development of deep learning, researchers found that deep learning has significantly improved the effect in prediction and classification compared with traditional machine learning models.

At present, the mainstream deep learning networks are mainly convolutional neural networks, recurrent neural networks, and long and short-term memory networks, and financial prices, as time-series data, need to consider the existence of a long-time dependence between prices in the prediction process, so one-dimensional (1D) convolutional neural networks are mainly used for short-term h prediction or extracting stock features [12-14]. Recurrent neural networks can save the information from previous stages and pass it to the later, so recurrent neural networks are widely used in the field of time-series forecasting. However, when the recurrent neural network processes long-time data, the gradient disappears and the gradient explodes. In order to solve these problems, Hochreiter and Schmidhuber [15] proposed the long short-term memory network. The LSTM network is a variant of recurrent neural network, which solves the above problems by three control units of forgetting gate, input gate, and output gate as well as cell states, and thus the long short-term memory network becomes the mainstream model in the field of temporal prediction. Selvin et al. [16] in 2017 compared three deep learning models with traditional statistical models by using convolutional neural networks, recurrent neural networks, and long short-term memory networks, and the results showed that the experimental results of the deep learning models were better than the traditional statistical models. However, a single neural network can suffer from poor fitting effect and low model prediction accuracy. To address this problem, researchers have started to use multi-model fusion to improve the prediction of models. Lu et al. [17] proposed a deep neural network model combining CNN and LSTM and achieved good results, which proved that neural network fusion can further improve the prediction of models.

Along with the rapid development of deep learning, in 2014 Goodfellow et al. [18] proposed a new model by two models against each other. Two models are trained simultaneously in the generative adversarial network: the generative model G, and the discriminative model D. The results are generated by adversarial between the two models (G and D). Generative adversarial networks are first used in the direction of image generation. In 2018, Zhang et al. [19] predicted multiple stock prices of the S&P S&P500 index by using the generative adversarial network, and by comparing the LSTM model, SVR model, and ANN model proved that generative adversarial networks have better results in stock prediction. Zhou et al. [20] predicted high-frequency stock market by using generative adversarial networks, and the experimental results showed that the stock prices generated by generative adversarial networks have better prediction results compared to traditional financial prediction models.

3 Model Network Structure

This section mainly introduces the prediction model of generative adversarial networks based on the multi-model fusion proposed in this study. Unlike most previous generative adversarial network financial forecasting models that use a single neural network as the generator. This study uses a one-dimensional convolutional neural network in the generator to extract stock features and assigns different weights to the output of the LSTM hidden layers through the attention mechanism, and finally predicts through a layer of LSTM network. The discriminant model consists of 1D convolutional neural network and fully connected network. The prediction results of *tanh*, *relu*, *leakyrelu* and *tanh* activation functions in each hidden layer of the discriminator are compared to find an activation function with higher prediction accuracy. Fig. 1 shows the model structure of this study.

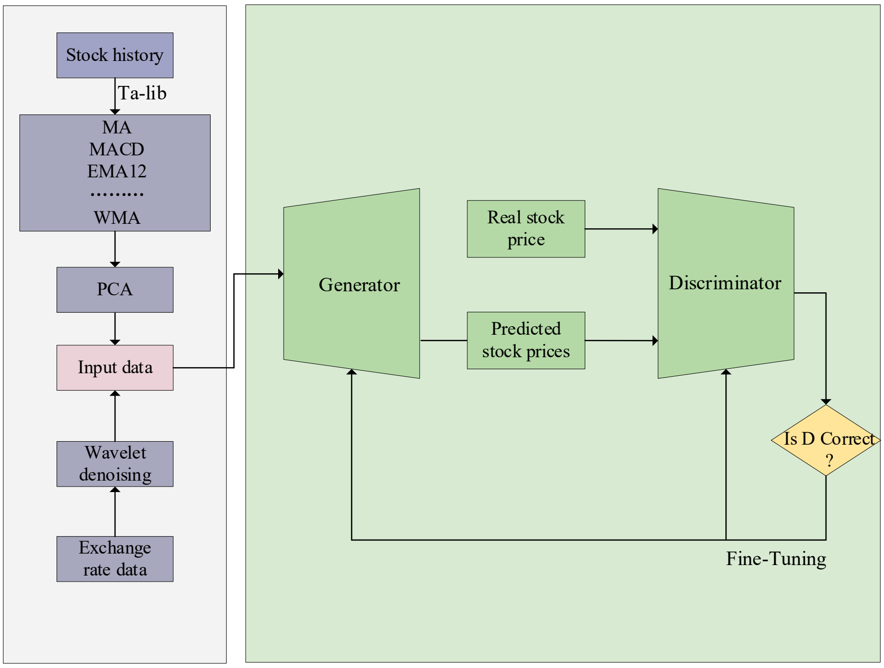


Fig. 1. Forecast model of generative countermeasure network finance based on multi model fusion

3.1 Generate Adversarial Networks

The generated adversarial network mainly trains the model through the game confrontation between the generator G and the discriminator D. In the generation adversarial network, the generator model is used to generate the forecast price of the stock, and the discriminator model is used to judge the input data. When the discriminator model is unable to determine whether the input data is the forecast data generated by the generator model or the real stock data, it means that the model is trained successfully. The input data in this study is the stock time series data $X = \{x_1, x_2, x_3, \ldots\ldots, x_k\}$, where x is the matrix of $t*n$, t is how many days the model predicts the next day, in this study, the data of the previous 5 days is used to predict the next day, and n represents the characteristic number of the stock.

Generator Model. In the field of stock prediction, traditional generator models are dominated by LSTM networks. Considering that a single deep learning model can have overfitting or underfitting phenomena in the prediction process, this study chooses to use a hybrid neural network based on CNN and LSTM in the generator and adds an attention mechanism to improve the prediction effect of the model. Prediction result $\hat{Y}_{t+1}$ is output through the fully connected layer.

$$\hat{Y}_{t+1} = G(X) . \tag{1}$$

Firstly, the input data $X=\{x_1,x_2,x_3,\ldots\ldots,x_k\}$ is fed into the 1D convolutional neural network to perform feature extraction on the input financial data, and the 1D convolutional neural network can extract short-term features of financial data through convolutional kernel. Secondly, in this study, a two-layer LSTM network is constructed after the one-dimensional convolutional network and an attention mechanism is added in the middle [21]. The formulation of the LSTM network is shown in Equations 2-7.

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f). \tag{2}$$

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i). \tag{3}$$

$$\overline{c}_t=tanh(W_c\cdot[h_{t-1},x_t]+b_c). \tag{4}$$

$$c_t=f_t*c_{t-1}+i_t*\overline{c}_t. \tag{5}$$

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o). \tag{6}$$

$$h_t=o_t*tanh(c_t). \tag{7}$$

The attention mechanism used in this study is to weight the hidden layer vectors h_t output from the first LSTM layer to assign different weights and improve the prediction of the model. First, a fully connected layer is constructed, and the input is the hidden layer vector h_t output from the first LSTM layer; Second, the hidden layer vector h_t is input to the fully connected layer to find the weight score s_i of each hidden layer vector, and the weight coefficient a_i of each feature is obtained by normalizing the weight score s_i using the $softmax$ function; Finally, the weight coefficient a_i is weighted with the hidden layer vector h_t to obtain the final output c_i. The formulas of the attention mechanism are shown in Equations (8) and (9).

$$a_i=softmax(s_i)=\frac{exp(s_i)}{\sum_{j=1}^{N}exp(s_j)}. \tag{8}$$

$$c_i=\sum_{i=1}^{k}a_ih_i. \tag{9}$$

Discriminator Model. The task of the discriminator model is to distinguish the real data from the false data as much as possible and output the result by the *Sigmoid* activation function. When the discriminator model judges that the input data is real data, the output result is 1, and when the input data is judged to be the predicted data generated by the generator, the output result is 0. In this study, the data generated by the generator are spliced with the real data of the previous 5 days as the predicted data ($\mathbf{Y}=\{Y_{t-4},Y_{t-3},Y_{t-2},Y_{t-1},Y_t,\hat{Y}_{t+1}\}$) and the real results of the day are spliced with the data of the previous 5 days as the real data ($\mathbf{Y}=\{Y_{t-4},Y_{t-3},Y_{t-2},Y_{t-1},Y_t,Y_{t+1}\}$) to be input to the discriminator model. The discriminator network, it consists of 1D convolutional neural network and fully connected neural networks. The 1D convolutional neural network can extract the local features of the input data well, a multilayer fully connected network follows the convolutional neural network to improve the classification ability of the discriminator, and the last layer is used as the output layer to output the results. Fig. 2 shows the network structure of the discriminator model.

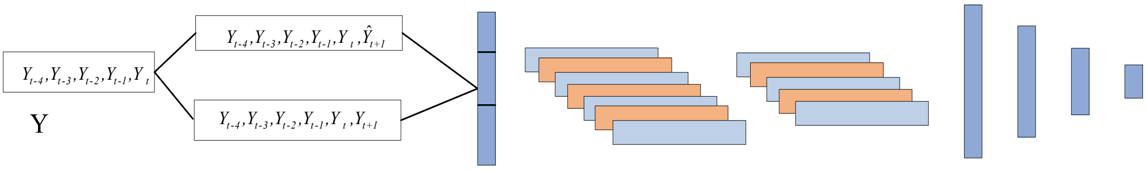


Fig. 2. Discriminator model structure

To improve the nonlinear capability of the model, this study introduces *gelu* as the activation function in the hidden layer of the discriminator [22]. *Gelu* activation function introduces the idea of random regularity, which is realized by multiplying the input by mask, which is random and depends on the input value in the range of 0-1. The *gelu* activation function can combine nonlinearity and stochastic regularization. In the experimental section of this study, we compare the prediction results of the model under the use of other activation functions (e.g. *tanh*, *relu*, *leakyrelu*) to verify the superiority of the *gelu* activation function in the present model.

$$Gelu(x) = xP(X <= x) = x\,\sigma(x)\ . \tag{10}$$

3.2 Loss Function

The optimization objective function for the generative adversarial network model is as follows:

$$\min_G \max_D V(D,G) = E_{x\sim P_{data}(x)}[\log(D(x))] + E_{z\sim P_z(z)}[\log(1-D(G(z)))]\ . \tag{11}$$

In this study, the loss function selection of generator refers to the loss function of most current generation countermeasure network time series prediction models. That is, the generator loss function is combined with the mean square error, and the generator is adjusted by the generator loss function and the mean square error. Where g_{loss} is the loss function of the generator in the generation adversarial network, g_{mse} is the mean square error loss function. λ_1 and λ_2 are user-defined hyperparameters. The formula is shown in Equation (12-14):

$$g_{loss} = \frac{1}{m}\sum_{i=1}^{m}\log\left(1-D\left(\mathbf{X}^{\mathbf{i}}_{fake}\right)\right). \tag{12}$$

$$g_{mse} = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{\mathbf{x}}^{i}_{t+1} - \mathbf{x}^{\mathbf{i}}_{t+1}\right)^2\ . \tag{13}$$

$$G_{loss} = \lambda_1 g_{mse} + \lambda_2 g_{loss}\ . \tag{14}$$

4 Data Analysis and Experiment

4.1 Data Processing

This section mainly deals with the financial data and exchange rate data of stocks and indexes in the data preprocessing stage. Firstly, the financial data of stock and index are aligned with the exchange rate data of corresponding time according to the date. Because some days of the obtained data are missing, this study chooses to delete the days with missing data; Then wavelet denoising is used for exchange rate data to remove small fluctuations in exchange rate data, Ta-lib financial database is used for financial data of stocks and indexes to generate a plurality of commonly used technical indicators, and PCA method is used for dimensionality reduction of the generated technical indicators to reduce data redundancy and improve the operation effect and prediction accuracy of the model; Finally, the input data were normalized to eliminate the dimensional differences existing between each feature to improve the model's prediction results [23].

In this study, the data of Zhuhai Port (000507.SZ), Xiamen International Trade (600755.SH), Oriental Venture (600278.SH) and Shanghai Composite Index (000001.SH) from January 1, 2013 to December 31, 2019 are selected as experimental data sets.

Wavelet Denoising. This study uses wavelet transform to extract the price trend of exchange rate features. Compared with Fourier transform denoising, wavelet transform has better characterization ability in both frequency and time domains. Fig. 3 shows the wavelet denoising results for the first 100 data in the exchange rate price data. As can be seen from Fig. 3, the use of wavelet denoising allows for the removal of smaller fluctuations in exchange rate prices while retaining the overall trend in exchange rate prices.

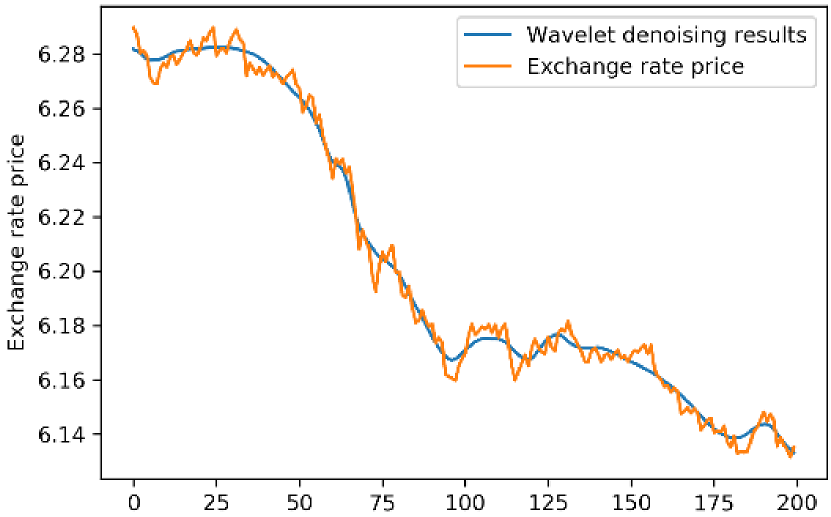


Fig. 3. The result of exchange rate price wavelet denoising

Normalization of Data. The input data are normalized to eliminate the dimensional differences that exist between individual features. Equation (15) is the maximum-minimum normalization formula.

$$\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} . \quad \textbf{(15)}$$

The Evaluation Index. In this study, MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error) and R^2_Score were selected as evaluation indicators. The formula is shown in Equations 16-19:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| . \quad \textbf{(16)}$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)^2 . \quad \textbf{(17)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} . \quad \textbf{(18)}$$

$$R^2 = 1 - \frac{\sum_i \left(\hat{y}^{(i)} - y_i\right)^2}{\sum_i \left(\bar{y} - y_i\right)^2} . \tag{19}$$

4.2 Experimental Results

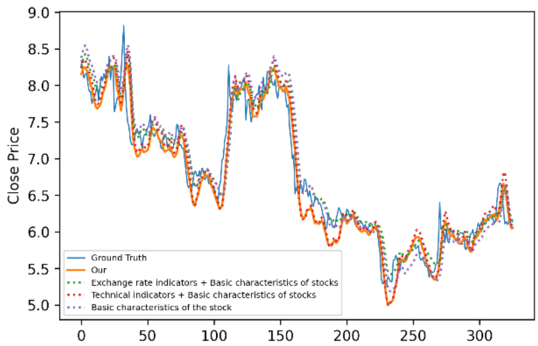
In this chapter, three sets of comparison experiments are designed to verify the prediction effect of the proposed model: in the first part, the effect of using different input features on the prediction effect of the model is explored, in this part, different features are chosen as the input data of the model and the prediction effect of each model is analyzed; in the second part, the superiority of the *gelu* activation function in the discriminator hidden layer is explored, three common deep learning activation functions *relu*, *tanh*, and *leaky-relu* are compared with the *gelu* activation function. In the third part, the predictive metrics of the prediction results are compared with the commonly used CNN, LSTM, CNN-LSTM, SVR financial prediction models and two generative adversarial network financial prediction models, respectively, to demonstrate the superiority of the proposed multivariate feature-based generative adversarial network financial prediction method in this paper.

(1) Input feature analysis

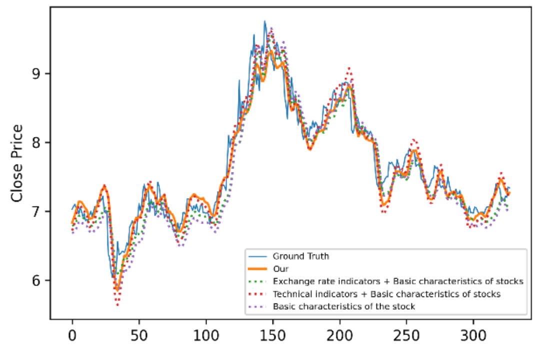
This paper designs multiple sets of experiments to verify the effects of different input features on the prediction effect of the model: 1) the input data consists of basic features of stocks, exchange rate features and technical indicators; 2) the input data consists of basic features of stocks and exchange rate data; 3) the input data consists of basic features of stocks and technical indicators of stocks; 4) the input data consists of basic features of stocks only. The prediction model is a multivariate feature-based generative adversarial network financial prediction model proposed in this paper, and Zhuhai Port, Xiamen International Trade, Eastern Venture, and SSE Index are selected as the experimental data sets. The prediction indexes of each model are shown in Table 1, and Fig. 4 shows the prediction result graph of each model.

Table 1. Predictors of different input feature models

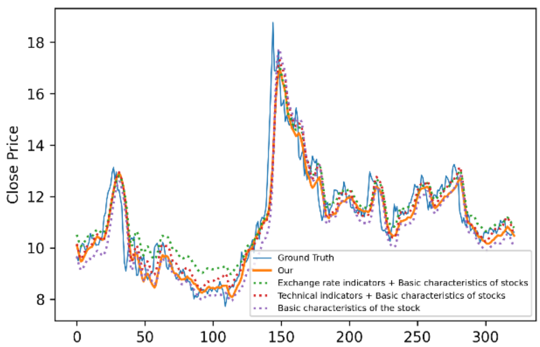
		MAE	MSE	RMSE	R^2_Score
Zhuhai Port dataset	**Our**	**0.182**	**0.060**	**0.245**	**0.926**
	Exchange rate characteristics	0.183	0.071	0.265	0.909
	Technical specifications	0.203	0.072	0.268	0.916
	Basic characteristics of stocks	0.211	0.088	0.296	0.910
Xiamen International Trade data set	**Our**	**0.172**	**0.052**	**0.228**	**0.905**
	Exchange rate characteristics	0.203	0.070	0.264	0.894
	Technical specifications	0.211	0.069	0.263	0.912
	Basic characteristics of stocks	0.240	0.090	0.300	0.885
Eastern Venture data set	**0.515**	**0.592**	**0.769**	**0.816**	**0.515**
	Exchange rate characteristics	0.605	0.743	0.862	0.711
	Technical specifications	0.525	0.648	0.805	0.801
	Basic characteristics of stocks	0.673	0.870	0.933	0.780
SSE Index Dataset	**Our**	**49.561**	**4157.197**	**64.476**	**0.836**
	Exchange rate characteristics	51.666	4320.783	65.733	0.869
	Technical specifications	52.403	4655.949	68.235	0.815
	Basic characteristics of stocks	58.909	5471.890	73.972	0.847



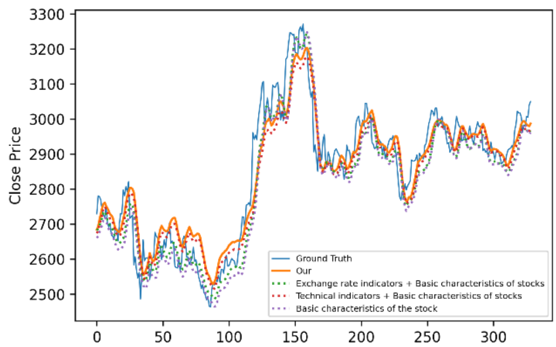
(a) Zhuhai Port data set



(b) Xiamen International Trade data set



(c) Orient Venture



(d) SSE data set

Fig. 4. Prediction results of different input feature models

As can be seen from Table 1, when only the basic characteristics of the stock itself are used as the input features of the model, the predictive index of the model is the worst result among the comparison models of all data sets. When the input characteristics are the basic characteristics of the stock itself plus the exchange rate characteristics or the technical indicators of the stock, the prediction results of the model are better than those of the comparison model using only the basic characteristics of the stock as the input. Based on the predictors of the four data sets in Table 1, it can be concluded that using the basic characteristics of the stock itself and the exchange rate characteristics and using the basic characteristics of the stock itself and the technical indicators characteristics have their own advantages and disadvantages in prediction under different data sets. When the input data are the multivariate feature inputs combining the stock's own features with exchange rate features and technical indicators as proposed in this paper, the prediction model achieves the optimal predictors and effects. The experimental results show that the introduction of both exchange rate indicators and technical indicators in the input features can improve the prediction effect of the model, and the model achieves the best prediction effect when the input data is a combination of the stock's own features with exchange rate features and technical indicators as proposed in this paper.

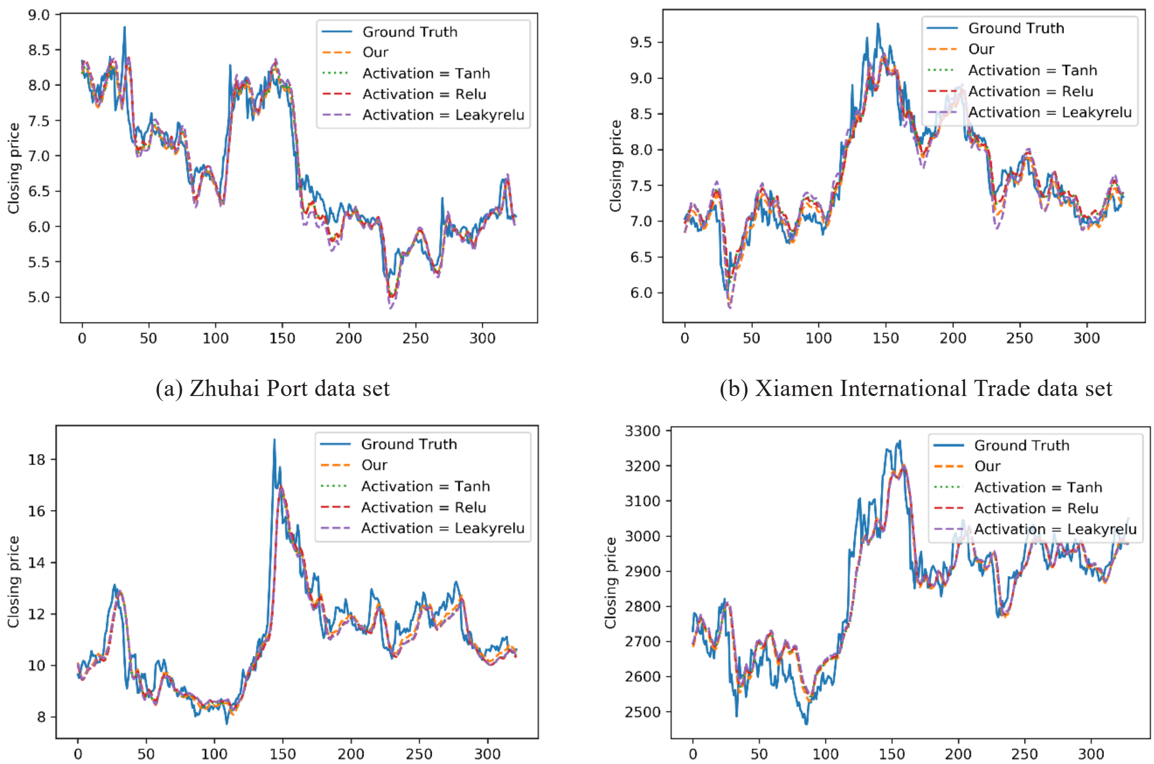
(2) Activation function analysis

Activation functions can handle nonlinear data in deep learning models. However, different activation functions have different ability to process data. To verify the superiority of the *gelu* activation function in the hidden layer of the discriminator, this paper selects Zhuhai Port, Xiamen International Trade, Eastern Venture, and SSE as the experimental data sets, uses the *relu*, *leaky-relu*, *tanh*, and *gelu* activation functions in the hidden layer of the discriminator, respectively, and analyzes the result values of MAE, MSE, RMSE, and R^2 _Score for each model. Table 2 shows the results of the predictive metrics for each model.

As can be seen from Table 2, compared to other activation functions, the metrics of MAE, MSE, and RMSE of the model under the use of the *gelu* activation function are lower than those of other comparison models, and the R^2 _Score is higher than that of other comparison models. Through Fig. 5, it can be intuitively seen that when different activation functions are used in the hidden layer of the discriminator, the prediction effect of the model will be different, and the prediction results using the *gelu* activation function are closer to the curve of the true value compared to other activation functions. The analysis of Table 2 and Fig. 5 proves that different activation functions have certain influence on the prediction performance of the model. However, it can be seen from Table 2 that different activation functions have less influence on the prediction results of the model, and the prediction index values of each activation function do not differ much. The experimental results show that the *gelu* activation function can improve the prediction accuracy of the model to a small extent compared with the traditional *relu*, *tnah* and *leaky-relu* activation functions.

Table 2. Predictors of different activation function models

		MAE	MSE	RMSE	R^2_Score
Zhuhai Port dataset	***Gelu***	**0.181**	**0.060**	**0.245**	**0.926**
	Leaky-Relu	0.219	0.082	0.287	0.910
	Relu	0.193	0.066	0.256	0.924
	Tanh	0.183	0.062	0.248	0.925
Xiamen International Trade data set	***Gelu***	**0.172**	**0.052**	**0.228**	**0.905**
	Leaky-Relu	0.204	0.069	0.262	0.868
	Relu	0.184	0.059	0.244	0.868
	Tanh	0.186	0.060	0.245	0.867
Eastern Venture data set	***Gelu***	**0.515**	**0.592**	**0.769**	**0.816**
	Leaky-Relu	0.559	0.636	0.798	0.789
	Relu	0.574	0.673	0.820	0.784
	Tanh	0.558	0.643	0.802	0.789
SSE index dataset	***Gelu***	**49.561**	**4157.197**	**64.476**	**0.836**
	Leaky-Relu	50.410	4316.175	65.698	0.833
	Relu	50.229	4309.145	65.644	0.833
	Tanh	50.311	4283.513	65.449	0.835



(a) Zhuhai Port data set

(b) Xiamen International Trade data set

(c) Orient Venture data set

(d) SSE data set

Fig. 5. Prediction results of different activation function models

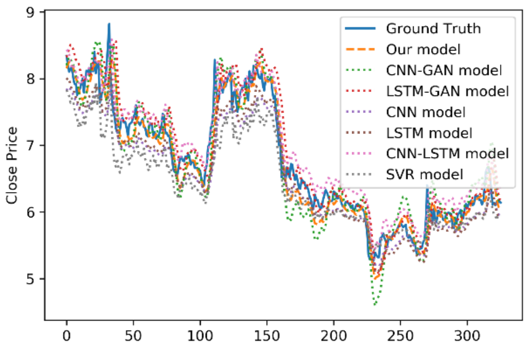
(3) Analysis of comparative experimental results

To demonstrate the prediction effect of the multivariate feature-based generative adversarial network financial forecasting method proposed in this paper, Zhuhai Port, Xiamen Guomao, Orient Venture, and SSE Index are used as experimental data sets, and the currently used financial forecasting models LSTM, CNN, CNN-LSTM, SVR and two generative adversarial network forecasting models LSTM-GAN, CNN-GAN are used as comparison models, and MAE, MSE, RMSE, and R^2 _Score are selected as experimental evaluation metrics.

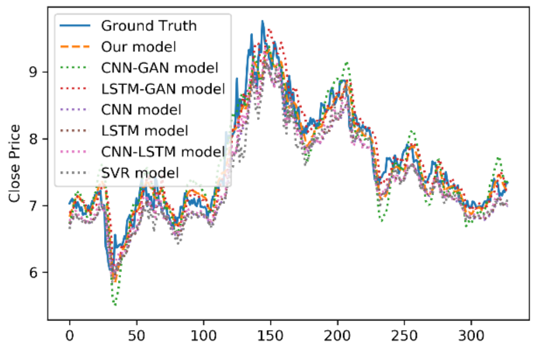
Among the comparison models, CNN model and LSTM model are consistent with Chapter 3, SVR model as one of the representative models in machine learning, some scholars have verified that SVR model can be applied in the field of financial data prediction. CNN-LSTM as a hybrid model effectively combines the advantages of CNN network to extract the spatial characteristics of data and LSTM in time series prediction, through the hybrid model the LSTM-GAN model and CNN-GAN model are the generative adversarial network model using the attention-based mechanism of LSTM network as a generator and the one-dimensional convolutional neural network, respectively. Table 3 shows the prediction index results of each dataset on different models, and Fig. 6 shows the prediction results of each model.

Table 3. Comparison of model predictors

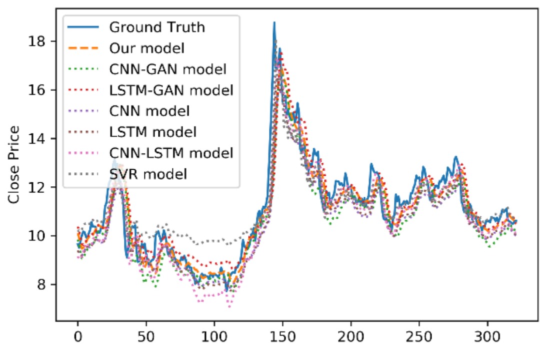
		MAE	MSE	RMSE	R^2 _Score
Zhuhai Port dataset	**Our model**	**0.181**	**0.060**	**0.245**	**0.926**
	LSTM-GAN	0.231	0.103	0.321	0.885
	CNN-GAN	0.263	0.110	0.332	0.881
	LSTM	0.251	0.088	0.297	0.872
	CNN	0.270	0.109	0.330	0.817
	LSTM-CNN	0.213	0.064	0.254	0.906
	SVR	0.341	0.167	0.408	0.636
Xiamen International Trade data set	**Our model**	**0.172**	**0.052**	**0.228**	**0.905**
	LSTM-GAN	0.191	0.066	0.257	0.897
	CNN-GAN	0.235	0.089	0.298	0.861
	LSTM	0.259	0.095	0.308	0.815
	CNN	0.273	0.105	0.325	0.793
	LSTM-CNN	0.221	0.071	0.267	0.865
	SVR	0.277	0.113	0.337	0.784
Eastern Venture data set	**Our model**	**0.515**	**0.592**	**0.769**	**0.816**
	LSTM-GAN	0.608	0.816	0.903	0.739
	CNN-GAN	0.730	0.933	0.966	0.721
	LSTM	0.631	0.640	0.800	0.793
	CNN	0.682	0.774	0.880	0.719
	LSTM-CNN	0.555	0.493	0.702	0.877
	SVR	0.592	0.592	0.769	0.670
eSSE index dataset	**Our model**	**49.561**	**4157.197**	**64.476**	**0.836**
	LSTM-GAN	52.588	4667.956	68.322	0.835
	CNN-GAN	68.448	7861.437	88.665	0.722
	LSTM	57.325	5104.094	71.443	0.848
	CNN	58.228	4884.184	69.887	0.789
	LSTM-CNN	57.876	4694.331	68.515	0.818
	SVR	81.249	9032.486	95.039	0.672



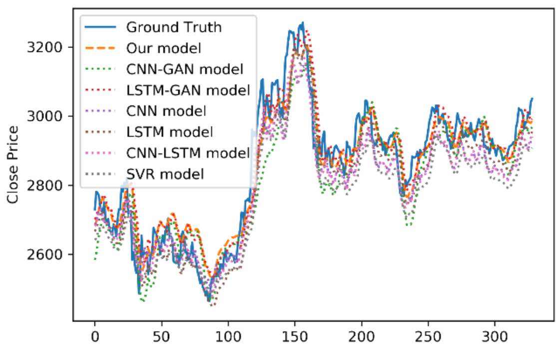
(a) Zhuhai Port data set



(b) Xiamen International Trade data set



(c) Orient Venture



(d) SSE data set

Fig. 6. Graph of prediction results of each model

As can be seen from Table 3, the proposed method in this paper has lower MAE, MSE, and RMSE predictors and higher R^2 _Score predictors than the comparison models in the Zhuhai Port dataset, Xiamen International Trade dataset, Eastern Venture dataset, and SSE index dataset.

As traditional financial prediction models, SVR, CNN, and LSTM models are often used as comparison models in the field of stock prediction. Experimental results show that the proposed multivariate feature-based generative adversarial network financial prediction method in this paper is better than traditional financial prediction models in all prediction indexes.

CNN-LSTM as a hybrid model effectively combines the advantages of CNN network that can extract effective features and LSTM in time series prediction, and it is also demonstrated in Chapter 3 that the prediction effect of the model can be improved by fusing multiple neural networks, and hybrid model prediction is also one of the mainstream prediction methods at present. By comparing and analyzing with CNN-LSTM models, the multivariate feature-based generative adversarial network financial prediction method proposed in this paper improves the prediction effect of the model through the discriminator's feedback for the generator. In addition, in the generator model, this paper adds an Attention layer in the middle of the two-layer LSTM based on the CNN-LSTM model, and assigns different weights to the output of each time-step hidden layer in the LSTM network, and the experimental results show that the introduction of both GAN and Attention mechanisms can improve the fitting effect and prediction performance of the model.

LSTM-GAN and CNN-GAN are used as ablation experiments to verify the prediction effectiveness of the proposed models in this chapter. Compared with LSTM-GAN and CNN-GAN, the multivariate feature-based generative adversarial network financial forecasting method proposed in this paper has better prediction metrics and prediction effects than these two generative adversarial network-based forecasting models. LSTM-GAN is more suitable for handling financial time series data compared with CNN-GAN because LSTM network is more suitable than CNN network, therefore, the prediction effect of LSTM-GAN is better than CNN-GAN in each dataset.

5 Conclusion

This study proposed a novel generative adversarial network financial forecasting model for predicting stock prices. The model uses CNN in the generator to extract stock features and assigns different weights to the hidden layers of the LSTM output through the attention mechanism to improve the prediction results, and finally generates the prediction results through the LSTM network; the discriminator consists of a convolutional neural network and a fully connected network, and uses *gelu* as the activation function in each hidden layer of the discriminator to crank up the prediction ability of the model. Considering that stock price is affected by many factors, exchange rate price and stock technical index are introduced in this study. In our experiments using datasets for Zhuhai Port, Xiamen International Trade, Eastern Venture, and SSE Index, the proposed model outperformed four other models (i.e., SVR, CNN, LSTM, CNN-LSTM, LSTM-GAN and CNN-GAN) in predicting stock prices.

Although this study considers the influence of several characteristics on stock prices, exchange rate and stock technical indicators are only a part of the many characteristics that affect stock prices, and more stock characteristics will be considered in the following study.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 61802230).

References

[1] R. Zhang, Z. Yuan, X. Shao, A new combined CNN-RNN model for sector stock price analysis, in: Proc. of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2018.
[2] R. Achkar, F. Elias-Sleiman, H. Ezzidine, N. Haidar, Comparison of BPA-MLP and LSTM-RNN for stocks prediction, in: Proc. of the 2018 6th International Symposium on Computational and Business Intelligence (ISCBI), 2018.
[3] X. Li, P. Wu, Stock price prediction incorporating market style clustering, Cognitive Computation 14(1)(2022) 149-166.
[4] Y. Chen, J. Wu, Z. Wu, China's commercial bank stock price prediction using a novel K-means-LSTM hybrid approach, Expert Systems with Applications 202(2022) 117370.
[5] M. Arya, H.G. Sastry, Stock indices price prediction in real time data stream using deep learning with extra-tree ensemble optimization, International Journal of Computational Science and Engineering 25(2)(2022) 140-151.
[6] A. Sensoy, B.-M. Tabak, Dynamic efficiency of stock markets and exchange rates, International Review of Financial Analysis 47(2016) 353-371.
[7] J. Zhang, Y.-H. Shao, L.-W. Huang, J.-Y. Teng, Y.-T. Zhao, Z.-K. Yang, X.-Y. Li, Can the exchange rate be used to predict the Shanghai composite index? IEEE Access 8(2019) 2188-2199.
[8] M. Waqar, H. Dawood, P. Guo, M.-B. Shahnawaz, M.-A. Ghazanfar, Prediction of stock market by principal component analysis, in: Proc. of the 2017 13th International Conference on Computational Intelligence and Security (CIS), 2017.
[9] A.A. Adebiyi, A.O. Adewumi, C.K. Ayo, Comparison of ARIMA and artificial neural networks models for stock price prediction, Journal of Applied Mathematics 2014(2014) 614342.
[10] W. Mei, P. Xu, R. Liu, J. Liu, Stock price prediction based on ARIMA-SVM model, in: Proc. of the 2018 International Conference on Big Data and Artificial Intelligence (ICBDAI), 2018.
[11] C. Xiao, W. Xia, J. Jiang, Stock price forecast based on combined model of ARI-MA-LS-SVM, Neural Computing and Applications 32(10)(2020) 5379-5388.
[12] R. Zhang, Z. Yuan, X. Shao, A new combined CNN-RNN model for sector stock price analysis, in: Proc. of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2018.
[13] Y. Wei, V. Chaudhary, The representation of wave trend direction in stock price time series via RNN, in: Proc. of the 15th International Conference on Machine Learning and Data Mining, MLDM, 2019.
[14] H. Li, Y. Shen, Y. Zhu, Stock price prediction using attention-based multi-input LSTM, in: Proc. of the Asian Conference on Machine Learning, 2018.
[15] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9(8)(1997) 1735-1780.
[16] S. Selvin, R. Vinayakumar, E.A. Gopalakrishnan, V.-K. Menon, K.-P. Soman, Stock price prediction using LSTM, RNN and CNN-sliding window model, in: Proc. of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017.
[17] W. Lu, J. Li, Y. Li, A.-J. Sun, J.-Y. Wang, A CNN-LSTM-based model to forecast stock prices, Complexity 2020(2020) 6622927.
[18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y.-S. Bengio, Generative adversarial nets, in: Proc. of the Advances in Neural Information Processing Systems, 2014.
[19] K. Zhang, G. Zhong, J. Dong, S.-K. Wang, Y. Wang, Stock market prediction based on generative adversarial network, Procedia Computer Science 147(2019) 400-406.

[20] X. Zhou, Z. Pan, G. Hu, S.-Q. Tang, C. Zhao, Stock market prediction on high-frequency data using generative adversarial nets, Mathematical Problems in Engineering 2018(2018) 4907423.
[21] D. Hendrycks, K. Gimpel, Gaussian error linear units (gelus). <https://arxiv.org/abs/1606.08415>, 2016.
[22] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.-N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in: Proc. of the Advances in Neural Information Processing Systems, 2017.
[23] Timing strategy of CSI 300 index based on wavelet denoising and random forest algorithm. Wu Han: Huazhong University of Science and Technology 2019.