

# Forecasting Stock Index Based On BP Neural Network Algorithm

Wanle Chi<sup>a</sup>

Department of information technology, Wenzhou Vocational & Technical College, ZheJiang, 325000, China;

<sup>a</sup>358455713@qq.com

**Keywords:** BP neural network, forecasting, stock index.

**Abstract.** BP neural network to forecast non-linear system, especially the financial data system, has a very strong capacity. A system model based on BP algorithm for forecast stock index was studied and shanghai complex index was forecasted in this paper. This paper has collected the index data of shanghai stock market (238 working days) during the period of 2017-01-08 to 2017-12-28. This paper was applying back propagation network to forecast shanghai complex index. The simulation result was shown that the BP algorithm was effective and feasible in stock index short-term forecast, and can achieve high accuracy.

## 1. Introduction

The stock market is a complex non-linear dynamic system, and the stock price trend is influenced by various unpredictable factors, such as political, economic and psychological factors. The volatility of stock prices often shows strong non-linear characteristics. Therefore, the linear analysis method to forecast the future trend of the stock price has limitations, and it is difficult to get a satisfactory result. Artificial neural network has strong mapping ability of non-linear system. It can realize mapping relationship between M dimension vector to N dimension vector. And it doesn't need to establish explicit analytic formula or mathematical model of complex non-linear system. BP neural network is the abbreviation of erroneous reverse transmission neural network, which is presently one of the most successful and widely applied neural network models. The result of this paper shows that the BP algorithm is effective and feasible in stock index short-term forecast, and can achieve high accuracy, and has certain practical value.

## 2. BP Neural Network & Data Processing

### 2.1 The Construction of Neural Network.

BP neural network generally adopts three layers of network structure, including input layer, hidden layer and output layer. The neurons in each layer are connected only with the adjacent layer neurons. And there is no connection between the neurons at the same layer. The adjacent layer neurons are connected by the weight value. The structure of BP neural network is shown in Fig.1.

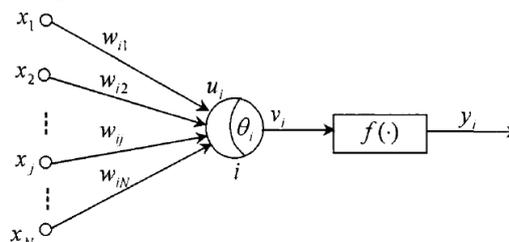


Fig. 1 The structure of three layers BP neural network

The characteristics of BP algorithm are the forward propagation of the signal and the reverse propagation of the error. In the forward propagation process, the input signal is propagated from the input layer to the hidden layer. After the action of the excitation function, the output of the hidden layer is transmitted to the output layer. The output result is obtained by the effect of the excitation function of the output layer.

If you do not get the desired output value, the backward propagation process should be carried on. The error between the real value and the output value is returned along the connected path. By modifying the connection weights of each layer of neurons, the error is reduced. And then the forward propagation process is transferred. Through the iteration for mentioned above, until the error is less than a given threshold value. By the iteration procedure mentioned above, the result of BP neural network can be obtained [1,2].

## 2.2 Data Processing.

This paper has collected the index data of shanghai stock market (238 working days) during the period of 2017-01-08 to 2017-12-28 ( [http://app.finance.china.com.cn/stock/quote/history.php?code=sh000001&begin\\_day=2017-01-09&end\\_day=2017-12-28](http://app.finance.china.com.cn/stock/quote/history.php?code=sh000001&begin_day=2017-01-09&end_day=2017-12-28) ). The stock closing index data of the last five working days made up a group as the input data p. The stock closing index data of the next working day is as the output data t. Some data are as Table 1.

Table 1 Some of original data

Date	Close-4	Close-3	Close-2	Close-1	Close-0	Next-close
2017/1/9	3135.92	3158.79	3165.41	3154.32	3171.24	3161.67
2017/1/10	3158.79	3165.41	3154.32	3171.24	3161.67	3136.75
2017/1/11	3165.41	3154.32	3171.24	3161.67	3136.75	3119.29
2017/1/12	3154.32	3171.24	3161.67	3136.75	3119.29	3112.76
2017/1/13	3171.24	3161.67	3136.75	3119.29	3112.76	3103.43

It,s necessary to normalize the input of sample in order to calculate conveniently and prevent partial neurons to supersaturate.

$$P = \frac{P - P_{\min}}{P_{\max} - P_{\min}} \quad (1)$$

In formula 1, input variable p of BP neural network is the stock index closing data. P is normalized factor of stock index data. To prevent neurons have reached saturation, the output flow is normalized as follow formula 2[3,4]:

$$T = \frac{t - t_{\min}}{t_{\max} - t_{\min}} \quad (2)$$

## 3. The Train of BP Neural Network

### 3.1 Determination of the Transfer Function and Training Function.

In this paper, the transfer function of output layer chooses purelin or tansig. Both convergent steps are similar, and both errors are same as above. The log - s - function has better fitting effect, so the transfer function of logsig is selected. Because the input and output values are normalized in [0,1] interval, this is in line with the actual comparison [5].

The model of BP neural network has many training functions, such as traingd, traingdm, traingcf, traingp, traingcb, trainscg. These functions have a great influence on BP neural network performance, some converging quickly, some converging slowly; some occupying large memory, some occupying small memory. MAPE (mean absolute percentage error) is used to test the fitting result of train.

Table 2 The training function test & MAPE value

Training Function	MAPE(with different number neurons)										
	10	11	12	13	14	15	16	17	18	19	20
traingda	6.2169	6.0023	5.7205	5.4038	6.2709	5.6570	5.6559	5.3558	5.6237	6.2352	5.9860
traingdm	5.4604	5.4861	6.7818	5.4819	7.7659	5.6351	5.5177	5.2673	5.8467	5.5235	5.8926
traingda	5.2923	5.4611	5.2923	5.6112	5.6360	5.2240	5.9082	5.4427	5.5115	5.2978	5.3763
traingcf	5.1741	5.1468	5.1929	5.3255	5.1150	4.9687	5.0586	4.9644	4.9774	4.9658	4.9706
traingcp	5.2220	5.1116	5.5540	5.3159	5.1372	5.2065	4.9939	5.0749	5.0142	4.9561	5.1502
traingcb	5.1181	5.7697	5.1795	5.0485	5.2487	5.7379	5.0756	5.3965	5.0105	5.2106	5.0574
trainscg	4.6598	4.7757	4.5051	4.5346	4.1941	4.3819	4.2953	4.2353	3.9714	4.2161	4.1931
trainbfg	3.9415	4.1789	4.2811	3.4005	3.7032	3.5962	3.6143	3.6466	3.5577	2.6650	3.3539
trainsfg	4.7984	4.7244	4.6302	4.5739	4.7972	4.9256	4.4858	4.5154	4.4741	4.3021	4.8295
trainlm	3.6834	3.7760	3.7471	3.2393	3.4090	3.3304	2.4831	2.8819	2.7831	2.0077	2.0801

In this paper, the heuristics is the method to determine the number of neurons in hidden layer. The number of hidden layer neurons is designed to 10~20. Write a program to test more than once with different number neurons and different training functions, results in Table 2.

Based on the same transfer function, Table 2 is obtained by changing the training function on the premise of target error 0.000001. According to the minimum MAPE criterion, the function of trainlm is the best training function. In this paper, the function of trainlm is selected for the training function of BP network based on the above analysis.

**3.2 Determination of the Number of Neurons in the Hidden Layer.**

The heuristics is use to determine the final number of neurons in hidden layer. In this paper, the number of hidden layer neurons is tested to 10~60. The best number is selected with the minimum MAPE value. The maximum number of iterations is identified as 10000 times. The goal error of test is 0.0000001. The learning rate of test is set to 0.05. The momentum factor is set to 0.9. The results of test are shown as Table 3.

Table 3. The number of neurons & MAPE value

neuron number	11	12	13	14	15	16	17	18	19	20
MAPE value	3.7760	3.7471	3.2393	3.4090	3.3304	2.4831	2.8819	2.7831	2.0077	2.0801
neuron number	21	22	23	24	25	26	27	28	29	30
MAPE value	2.0946	1.8653	1.6903	1.3014	1.4017	1.3618	1.4040	1.0608	0.6484	0.5814
neuron number	31	32	33	34	35	36	37	38	39	40
MAPE value	0.5344	0.2938	0.1296	0.1035	0.1114	0.0940	0.0848	0.0966	0.0859	0.0870
neuron number	41	42	43	44	45	46	47	48	49	50
MAPE value	0.0935	0.0846	0.0863	0.0801	0.0231	0.0918	0.0922	0.0933	0.0856	0.0912
neuron number	51	52	53	54	55	56	57	58	59	60
MAPE value	0.0745	0.0759	0.0863	0.0818	0.0610	0.0891	0.0915	0.0819	0.0893	0.0835

The result is shown the MAPE value of 45 neurons is lest. So, the final selection of the paper is the model of 45 neurons in hidden layer with the trainlm training function and the tansig transfer function.

**3.3 Result of Train.**

The results of train of BP neural network are follows:

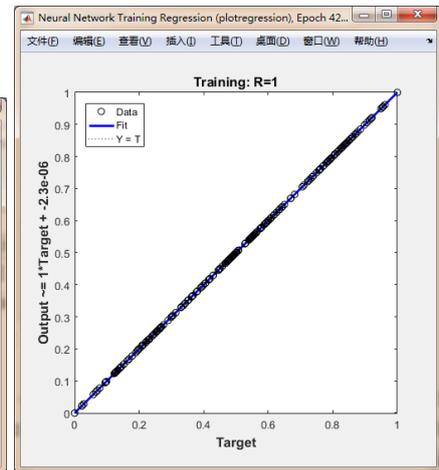
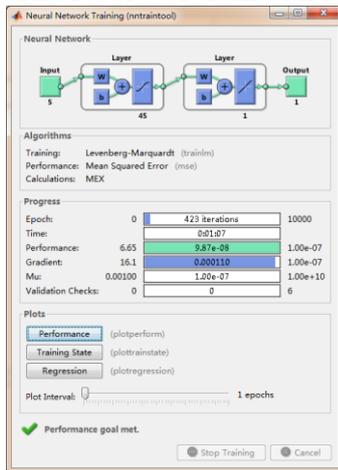


Fig. 2 The graph of training Fig. 3 The training performance Fig. 4 The training regression

From Fig. 2 and Fig.3, the number of iterations is 423 .It shows that the convergence rate of the algorithm is high. The computational performance of the algorithm is excellent. From Fig.4. the fitting result is closer to the real value. The algorithm has good effect.

**4. The Forecasting of BP Neural Network**

**4.1 Test Data.**

The test data for forecasting are the index data of shanghai stock market (17 working days) during the period of 2017-12-29 to 2018-1-23.Same as training data, the index data of the last five working

days are the input data. Some of test data are shown as Table 4. The trained model is used to forecast the index value of the next day.

Table 4. Some of test data

Date	Close-4	Close-3	Close-2	Close-1	Close-0	Next-close
2017/12/29	3306.12	3275.78	3296.39	3307.17	3307.17	3348.33
2018/1/2	3275.78	3296.39	3307.17	3307.17	3348.33	3369.11
2018/1/3	3296.39	3307.17	3307.17	3348.33	3369.11	3385.71
2018/1/4	3307.17	3307.17	3348.33	3369.11	3385.71	3391.75
2018/1/5	3307.17	3348.33	3369.11	3385.71	3391.75	3409.48

#### 4.2 Forecasted Result.

As shown in Table 15, the prediction results of numerical simulation show a good agreement with the real values. The MAPE of forecasting is 0.95%. Considering the limit of the stock index, the maximum of error rate is 10%. It shows that the forecasting of BP neural network is not excellent. this prediction performance may be due to the lack of training data. The training data is not added to volume data and related industry data. Perhaps the effect of training data into volume data and related industry data will be better.

Table 5 Result of forecasting

Date	2017/12/29							2018/1/10	
Forecast	3334.63	3319.97	3378.72	3353.05	3373.18	3439.55	3444.83	3330.63	
Real	3348.33	3369.11	3385.71	3391.75	3409.48	3413.9	3421.83	3425.34	
Difference	13.70	49.14	6.99	38.70	36.30	-25.65	-23.00	94.71	
PAE	0.41	1.46	0.21	1.14	1.06	0.75	0.67	2.77	
Date	2018/1/11							2018/1/23	
Forecast	3429.84	3445.64	3445.71	3418.74	3446.16	3431.53	3444.15	3561.40	3600.78
Real	3428.94	3410.49	3436.59	3444.67	3474.75	3487.86	3501.36	3546.51	3559.47
Difference	-0.90	-35.15	-9.12	25.93	28.59	56.33	57.21	-14.89	-41.31
PAE	0.03	1.03	0.27	0.75	0.82	1.62	1.63	0.42	1.16

## 5. Summary

The BP neural network is applied to stock index prediction. And a stock index prediction system model is established, which achieves good prediction effect and has certain reference value for practical application. If the size of the sample can be expanded, and add more parameters in the stock market, such as the opening price, the highest price, the lowest price, the volume, the DMA, the RSI, etc., the BP neural network will be able to make more accurate predictions.

## Acknowledgments

This work was financially supported by Wenzhou vocational & technical college scientific research project fund(wzy2016019).

## References

- [1]. Xin Yao, Yong Liu. A new evolutionary system for evolving artificial neural networks. *Transaction on Neural Networks*. Vol. 8 (2007) No. 3, p. 34-37.
- [2]. Valian E, Mohanna S, Tavakoli S. Improved cuckoo search algorithm for feed forward neural network training. *International Journal of Artificial Intelligence , Applications*. Vol. 2 (2011) No. 3, p. 36-43.
- [3]. Jin Xu. A new algorithm for feed forward neural network learning and its simulation. *Journal of Commercial University: Natural Science Edition*. Vol. 4 (2007) No. 1, p.1-3.

- [4]. YongDong Zhang, QiuXiang Bi. An empirical comparison of the Volatility Prediction Model in Shanghai stock market. *Journal of Management Engineering*. Vol. 2 (2003) No. 3, p.16-19.
- [5]. JiGang Zhang, Na Liang, YuMin Shen. An improvement of prediction of stock prices based on feed forward neural network. *Journal of Xianning College*. Vol. 27(2007) No. 3, p.1-3.