Article

Stock Prediction via Linear Regression and BP Regression Network

Zhaoqin Yan [1]*, Zhangying Huang [1] , Minwei Liang [2]

[1] School of Science, Nanjing University of Science and Technology,Nanjing 210094, China
[2] School of Science, Nanjing University of Science and Technology,Nanjing 210094, China
* Corresponding author: Zhaoqin Yan(email:zhaoqin_yan@163.com)

Abstract: Stock prediction is an important and challenging problem. In this paper, two models, i. e., a multiple linear regression based model and a BP neural network based model were studied to predict the movements of the stock market. Through the analysis of the conceptual topic index of Shenzhen A-share "Gree Electric Appliances" (000651) and the prediction of stock price technical indicators on opening price of the next day, it was found that the results predicted by technical indicators on the opening price were more accurate than the conceptual topic index. At the same time, we compared the fitting degree, simulation and prediction ability of the multiple linear regression model and the BP neural network prediction model. The results show that the prediction model based on BP neural network works better than the multiple linear regression model. We trained and tested the index of "Gree Electric Appliances" in 220 trading days in this paper. Based on the experimental results, the conclusion was drawn that the BP neural network based model with the stock technical index as the characteristic variable can predict the movements of the stock market well.

Keywords: BP neural network; linear regression; stock price prediction; R language

1. Introduction

The stock market is a "high income with high risks" investment field. Countless experts and scholars are constantly looking for effective analytical methods and models to find out the laws of stocks and predict stock [1–5]. Nowadays, a popular research method is to apply a neural networks to prediction [6–9]. The efficiency of neural network model is often superior to multivariate statistics, which is relatively outdated [10–15]. Most previous work adopts various technical indicators of stocks as characteristic variables, while the price at a certain moment or the average price of a certain period of time is the response variable, and the accuracy of various forecasting results are compared.

However, each stock has its own "conceptual topic", and the trend of "conceptual topic" can also reflect the development and popularity of the stock. A stock generally belongs to several conceptual topics. *If the index of the conceptual topics are used as characteristic variables, can we predict the trend of stocks?* The problem has been rarely studied in this area. In this paper, we first use iterative algorithm to find out the best number of hidden layers applicable for a set of stock data. Then the multiple regression[16–19] and BP neural network [20–27] were used to predict the opening price of the second trading day, with technical indicators and the conceptual subject index as the characteristic variables. Finally, the pros and cons of four methods are shown through prediction results.

2. Preliminaries

2.1. Multiple linear regression

Multiple linear regression model is a popular regression analysis method exploiting the linear relationship between an observed variable and multiple input variables. The general form of multiple linear regression model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \tag{1}$$

where $y \in \mathbb{R}$ is the observed variable; x_i $(i = 1, 2, \cdots, p)$ are the input variables; $\beta_0, \beta_1, \cdots, \beta_p$ are the unknown parameters; and ε is a zero-mean random Gaussian noise with variance σ^2. For notation simplicity, let $\boldsymbol{x} = (x_1, \cdots, x_p)' \in \mathbb{R}^p$ be the collection of input variables and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)' \in \mathbb{R}^p$ be the coefficient vector to be estimated.

Suppose we have n observations $y_1, \cdots, y_n$ with input variables $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ and noise $\varepsilon_1, \ldots, \varepsilon_n$. Representing Eq. (1) in matrix form, we obtain

$$Y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

where $Y = (y_1, \cdots, y_n)' \in \mathbb{R}^n$ is the observation vector, $X = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)' \in \mathbb{R}^{n \times p}$ is the matrix of input variables, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)' \in \mathbb{R}^n$ is the noise vector.

In Eq. (2), if X is a column full rank matrix, by applying the least-squares estimation, the estimation of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \left(X'X\right)^{-1} X'Y \tag{3}$$

With the estimation $\hat{\boldsymbol{\beta}}$, and the regression result $y_{\mathbf{new}}$ of a new input vector $\boldsymbol{x}_{\mathbf{new}}$ is given by the sample regression function as follows:

$$y_{\mathbf{new}} = \boldsymbol{x}'_{\mathbf{new}} \hat{\boldsymbol{\beta}} \tag{4}$$

The sample regression function (4) needs to be further statistically tested to determine the reliability of the estimation, including the F-test of the regression equation, the regression coefficient t-test, and test of fitness of the regression.

- The F test of the significance of the regression equation

  H_0:Original hypothesis is tenable

  H_1:Original hypothesis is not tenable

  If the assumption of is tenable, then

  $$F = \frac{\frac{SSR}{p}}{\frac{SSE}{(n-p-1)}} \tag{5}$$

  where $F \sim F(p, n-p-1)$, $SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$, $SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$. Given the level of significance α,check the table to get the critical value $F_{\alpha(p,n-p-1)}$,and calculate the value of F from a sample. If $F>F_{\alpha(p,n-p-1)}$,then reject (or accept) in the opposite case H_0,to determine whether the linear relationship of the original equation is significant.
- t-test of regression coefficient

  Structure t Statistics

  $$t_i = \frac{\hat{\beta}_i}{\hat{\sigma}\sqrt{c_{ii}}} \sim t(n-p-1) \tag{6}$$

  where $\hat{\sigma} = \sqrt{\frac{SSE}{n-p-1}}$. Given a significant level α,get the critical value $t_{\frac{\alpha}{2}}(n-p-1)$.According to $|t_i| > t_{\frac{\alpha}{2}}(n-p-1)$ to decide (or accept) the null hypothesis H_0,then to determine whether the corresponding explanatory variable should be included in the model.

- Test of fitness of the regression

$$R^2 = \frac{SSR}{SST} = \frac{SSR}{SSR + SSE} \tag{7}$$

In the formula,R^2 is the determined coefficient of the sample. The closer the statistic is to 1, the higher the goodness of fit of the model.

2.2. BP neural network

The BP algorithm[28] was proposed by a group of scientists led by Rumelhart and McCelland in 1986. It is a multi-layer feedforward network trained by error inverse propagation algorithm and is one of the most widely used neural network models. The BP algorithm consists of two parts:the forward propagation of the signal and the back propagation of the error. In the forward propagation process, the signal is transmitted from the input layer of the network to the output layer through the hidden layer layer by layer, and the actual output of the network is obtained. If the actual output does not match the expected output, then the error back propagation phase is entered. In the back propagation phase, the output error is transmitted back to the input layer via the hidden layer,thereby obtaining an error signal of each unit of each layer, and the network connection weight is adjusted according to the signal. The two processes of forward propagation of the signal and back propagation of the error are repeatedly performed until the network output error is less than a predetermined threshold or until a predetermined study number of times is performed.

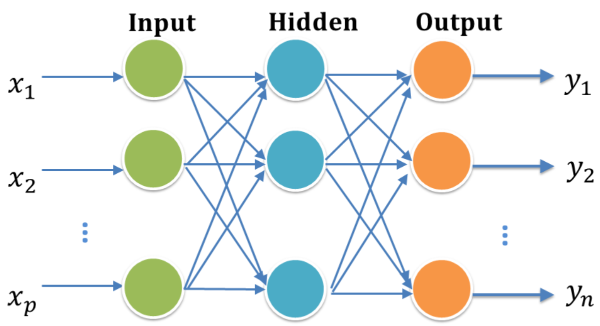

Figure 1. Structure of BP network

Suppose that we study the lth layer of neurons j, i represent the (l-1)th layer of neurons and k indicates l+1 layer of Then o_j representing the output of neurons j, ω_{ji} representing the connection weight between neurons i and j, ω_{jk} representing the connection weight between neurons j and k.

The steps of BP neural network algorithm :

1. Initialize all network weights $\omega_{ki}(t)$ (including bias) is a small random number, where t indicates the number of learning steps and is initialized to 0.
2. Repeat the follow process until convergence (calculated each training sample).

- Forward propagation of the signal.Calculate the output o_j of each hidden layer and output layer neurons layer by layer:

$$net_j = \sum_i \omega_{ji}(t)\, o_i \tag{8}$$

$$o_j = f\left(net_j\right) \tag{9}$$

In the formula,$f\left(\cdot\right)$ is an activation function. Use the Sigmoid function as the activation function:

$$o=f\left(net\right)=\frac{1}{1+e^{-net}} \tag{10}$$

- The direction of the error propagates.

  First, calculate the neurons in the output layer.

$$\delta_j=\left(d_j-o_j\right)o_j\left(1-o_j\right) \tag{11}$$

In the formula,δ_j is the error signal,d_j is the desired output,and o_j is the actual output.

Then calculate the neurons of each hidden layer along the network.

$$\delta_j=o_j\left(1-o_j\right)\sum_k \omega_{kj}\delta_k \tag{12}$$

If j is the first l layer of neurons, in the above formula $\sum$ is the summary of all neurons k of l+1 layer.

- Calculate the weight correction amount:

$$\Delta\omega_{ji}\left(t\right)=\eta\delta_j o_i \tag{13}$$

- Update each network weight:

$$\omega_{ji}\left(t+1\right)=\omega_{ji}+\Delta\omega_{ji}\left(t\right) \tag{14}$$

3. Data Acquisition

"Gree Electric Appliances" is a typical stock representing China's manufacturing industry. It is a conceptual stock with a lot of topics.This experiment predicts the opening price of Gree Electric Appliances by analyzing the concept theme index and stock price technical indicators of it. For the sake of generality, make use of the 220 sets of data from 5 August 2017 to 11 July 2018 when exploring the next day's opening index.80% of them were randomly selected as network training data, and the restwas used as test data to verify the accuracy of the BP network. The training data of groups is 176 and the test data is 44.

The main conceptual topics belonging to "Gree Electric" include the ten concept themes of state-owned enterprise reform, smart home, MSCI China, new energy, low-carbon economy, new energy vehicles, margin trading, electrical machinery, industry 4.0, and social security. The indices of the ten subjects are set to x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, and x_{10} as the first set of characteristic variables respectively. In addition, the traditional stock technical indicators include the volume, the number of transactions, K, D, J, MACD, BIAS1, BIAS2, BIAS3, and the average volume of the transaction volume. Similarly, these ten indicators are set to x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10} which are used as the second set of characteristic variables. Among them, the transaction volume refers to the total number of shares traded on the same day (1 lot = 100 shares), the number of transactions refers to the number of pending orders of all traders, and the KDJ indicator is calculated based on the highest price, the lowest price and the closing price. The K value, D value and J value, the three can accurately reflect the price fluctuation trend and its strength, generally used to predict the short- and medium-term market. The MACD is an exponentially smooth moving average that characterizes the current long-short state and possible trends in stock prices. BIAS is a technical indicator that reflects the degree to which the stock price deviates from the moving average during the volatility process. It judges the timing of buying and selling by calculating the percentage of the stock price that deviates from the moving average. These technical indicators are the most basic and most representative of the stock price trend.

In order to compare the predictive performance of multiple regression and BP neural network on stock price, the next-day opening price of Gree Electric Appliance is taken as the one-dimensional response variable of the final output.

Normalization and anti-normalization To eliminate the difference in magnitude between data, the input methodshould first be normalized. The normalization formula is as follows:

$$x' = \frac{x - min\left(x\right)}{max\left(x\right) - min\left(x\right)} \tag{15}$$

However, the computing result (predicteddata) is all between 0 and 1,anti-normalizationmethod should be applied to reset the results to the magnitude of actual data.Theanti-normalization formula is as follows:

$$x' = x_{result}\left(max\left(x\right) - min\left(x\right)\right) + min\left(x\right) \tag{16}$$

The hidden layer number of BP neural network has a great influence on the accuracy of network prediction. In this experiment, the change of prediction accuracy is observed by changing the number of hidden layers. We found out that when the network has 5 or 7 hidden layers, the correlation coefficient between predicted and actual data reach the highest point, which means that prediction accuracy is relatively higher. However, a network with too many hidden layers may cause "over-fitting" phenomenon, so we only established a BP network with 5 layers.

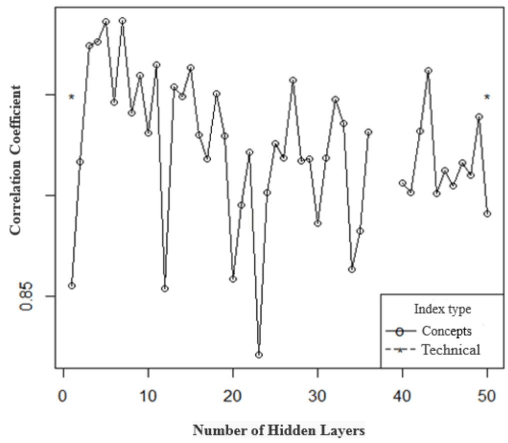

Figure 2. Correlation coefficient between predicted and actual data

This paper uses the nnet package in R to carry out the simulation experiments, and the BP neural network is created with 10 inputs and one output. The maximum times of training is 1000 and the target precision is 10^{-5}.

4. Analysis and Results

4.1. Use BP regression neural network to predict the opening price of the next day

After the BP neural network training is completed, the test data is used to test the predictive power of the neural network. The criterion is the relative error between the output value and the actual value calculated by the neural network from the feature vector:

$$RelativeErrorr = \frac{\left|PredictedData - ActualData\right|}{ActualData} \tag{17}$$

Table 1. BP nerual network by using R language

	BP nerual network by using R language
	preindex<-read.table("D:nextdayopen.txt",header=T)
step1	normalize<-function(x)return((x-min(x))/(max(x)-min(x)))
	pre.norm<-as.data.frame(lapply(preindex,normalize))
	#Import and normalize the data.
	n<-dim(pre.norm)[1]
	set.seed(13)
step2	train_index<-sample(1:n,round(n*0.8))
	train<-pre.norm[train_index,]
	test<-pre.norm[-train_index,]
	#Set 80% of the trading day data for training, while the remainder to test the accuracy of the network.
	library(nnet)
	r<-1/max(abs(train[,1:10]))
	for(i in 1:50)
	set.seed(101)
step3	model<-nnet(Y∼ X1+X2+X3+X4+X5+X6+X7+X8+X9+X10,data=train,size=i,rang=r,decay=1E-5,maxit=1000)
	pred_test<-predict(model,test[,1:10])
	cov<-cor(test[,11],pred_test)
	cov[i]<-cov
	#The i represents the number of hidden layers in the network.
	#The correlation coefficient between the predicted data and the test data helps us get the best i.
	set.seed(101)
step4	model<-nnet(Y∼X1+X2+X3+X4+X5+X6+X7+X8+X9+X10,data=train,size=5,rang=r,decay=1E-5,maxit=1000)
	pred_test<-predict(model,test[,1:10])
	#We found that 5 hidden layers works best.
	preindex$Y
	maxY<-max(preindex$Y)
	minY<-min(preindex$Y)
step5	test.real<-test[,11]*(maxY-minY)+minY
	t.r<-as.matrix(test.real)
	pred_test.real<-pred_test*(maxY-minY)+minY
	pt.r<-as.matrix(pred_test.real)
	#Anti-normalize the predicted data.
	wucha<-abs(t.r-pt.r)/t.r
	cbind(t.r,pt.r,wucha)
step6	cor(t.r,pt.r)
	mean(wucha)
	#Calculate the relative error and correlation coefficient.

The results are as follows:

Table 2. The results by technical indicators

Primary Data	Predicted Data	Relative Error
39.11	38.94711	0.004164887
39.1	38.77864	0.008218959
39.62	38.91383	0.017823466
37.02	37.72703	0.01909854
36.7	37.48038	0.021263794
38.5	39.54442	0.027127728
...	...	...
48.7	47.38575	0.026986734
49.93	49.78173	0.002969631
47.5	48.12087	0.013071005
47.36	46.24841	0.023471145
43.79	43.85666	0.00152232
45.15	44.21843	0.020632842

Table 3. The results by conceptual topic index

Primary Data	Predicted Data	Relative Error
39.11	38.21203	0.022960176
39.1	39.07308	0.00068847
39.62	39.26665	0.00891839
37.02	37.97765	0.02586847
36.7	37.47494	0.021115416
38.5	38.09592	0.010495533
...	...	...
48.7	48.51834	0.003730131
49.93	49.52229	0.008165664
47.5	47.12746	0.007842977
47.36	45.38064	0.041793917
43.79	43.06277	0.016607228
45.15	45.84767	0.015452365

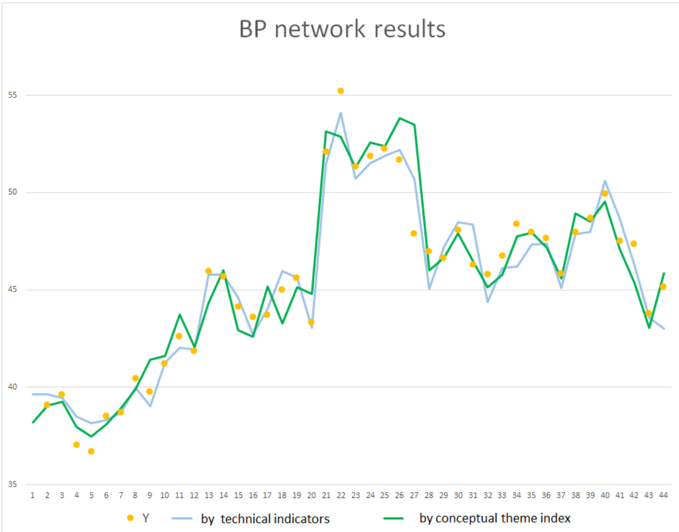

Figure 3. BP neural network results

BP neural network has a wide range of applications in nonlinear systems because of its powerful mapping and learning ability. In this prediction, the role of the BP neural network is to approximate the function. That is, to train a network approximate to the objective function through a large amount of historical data. From the plot, the prediction results of technical indicators are in a close race with that of the conceptual topic index, but the former are more accurate and consistent with the original data.

4.2. *Use regression analysis to predict the opening price of the next day*

Through the method of multiple linear regression, observe how the concept sector index and technical indicators affect the opening price of the next day.

Using regression prediction and the concept sector index to predict the opening price of the second day is more accurate, and the predicted data is closer than the original data.

Table 4. The regression results of technical indicators

Variable	Regression Coefficients	Standard Error	t-Stat	P-value
β_0	7.759591664	3.290300309	2.358323234	0.019280566
x_1	-0.009210939	0.003165654	-2.909648366	0.004009927
x_2	0.007026998	0.002899701	2.423352805	0.016229487
x_3	0.002765903	0.007028866	0.3935063	0.694346706
x_4	-0.006011439	0.002115299	-2.841885902	0.00492901
x_5	0.01754617	0.001778021	9.868371199	4.22955E-19
x_6	0.001234999	0.001905114	0.648254973	0.517531614
x_7	0.001336016	0.00818905	0.163146615	0.870560641
x_8	0.000720765	0.001473969	0.488996116	0.625357199
x_9	-0.004464724	0.002895185	-1.542120331	0.124557077
x_{10}	0.000132717	0.001989701	0.066702106	0.946882621

Table 5. The regression results of conceptual index

Variable	Regression Coefficients	Standard Error	t-Stat	P-value
β_0	7.798106051	3.296737438	2.365401005	0.018930063
x_1	-0.009236689	0.003171453	-2.912446822	0.003977502
x_2	0.007211801	0.002925265	2.465349343	0.014497831
x_3	0.002265616	0.00710312	0.318960607	0.750076463
x_4	-0.006069737	0.002121746	-2.860727428	0.004657993
x_5	0.017396936	0.001802912	9.649356268	1.92266E-18
x_6	0.001320036	0.00191503	0.689303192	0.491400722
x_7	0.002086015	0.00832273	0.250640656	0.802339353
x_8	0.000755001	0.001477893	0.510863446	0.609988469
x_9	-0.004560978	0.002905762	-1.569632534	0.118020602
x_{10}	-4.38761E-05	0.002020424	-0.021716262	0.98269511

The results are as follows:

Table 6. The results by technical indicators

Primary Data	Predicted Data	Relative Error
38.48	40.20426773	0.044809452
38.01	39.87262843	0.049003642
37.89	39.69626394	0.047671257
38.32	39.57951755	0.032868412
39.11	41.23412985	0.054311681
39.1	40.94146589	0.047096314
...	...	...
43.8	44.17434939	0.00854679
43.79	43.16430863	0.014288453
44.78	43.75722162	0.022840071
46.02	46.5088528	0.010622616
44.8	43.98332356	0.018229385
45.15	41.80486062	0.074089466

Table 7. The results by conceptual topic index

Primary Data	Predicted Data	Relative Error
38.48	37.84681531	0.016454903
38.01	37.75276641	0.006767524
37.89	37.92744439	0.000988239
38.32	37.44328359	0.022878821
39.11	38.14067425	0.024784601
39.1	39.17662843	0.001959806
...	...	...
43.8	44.44597188	0.014748216
43.79	43.55431506	0.005382163
44.78	43.31751376	0.032659362
46.02	43.8753903	0.046601688
44.8	44.95678364	0.003499635
45.15	44.07358296	0.023840909

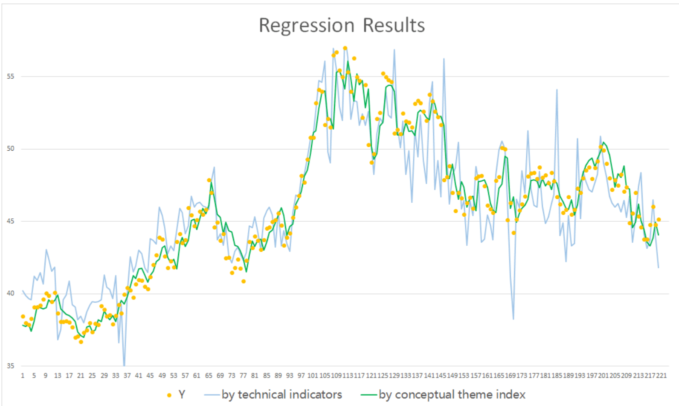

Figure 4. Regression Results

After multivariate regression analysis is performed by the data of 220 trading days, the multivariate regression polynomial (ie, the objective function) is obtained, and the characteristic variables are substituted into the polynomial to figure out the prediction result. The points represent the actual data, and the endpoints of the polylines represent the predicted data. We can clearly see that when using technical indicators as characteristic variables, the prediction results are quite different from the actual values, and the adjacent points are not as coherent and smooth as the stock trend. By contrast, when it comes to the conceptual topic index to which the stock belongs, the forecast result is apparently more suitable.

4.3. Analysis and Results

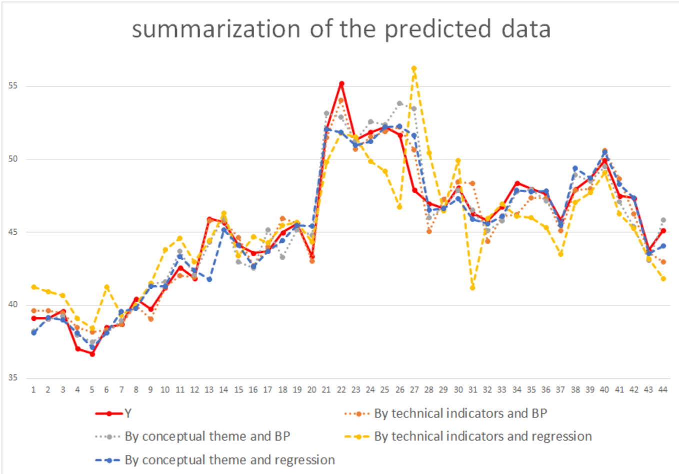

Figure 5. Summarization of the predicted data

Table 8. Accuracy analysis of four prediction methods

Prediction Methods	Correlation Coefficient	Relative Error(%)
BP network by technical indicators	0.9862409	1.29088
Regression by technical indicators	0.9012992	3.715985814
BP network by conceptual topic index	0.9590527	1.928325
Regression by conceptual topic index	0.969870007	1.869738773

According to the summary table, when the BP regression network is used for prediction, the prediction error of the technical indicator is only 1.29%, while that of the conceptual topic index is close to 2%. When the multiple regression model is used for prediction, the prediction error of the technical indicator is high up to 3.7%, while that of the conceptual topic index is 1.87%. The technical indicators are flexible and complicate calculations on the stock price, so it is applicable to the neural network that can make nonlinear prediction. However, the conceptual topic index describes the aspect and industry of a stock. We come to the conclusion that the trend of a stock can be observed by industry and macro-economy conditions. On the whole, by introducing the technical indicators and the combination of technical indicators to do predictions, it seems that BP neural network do the job better, the results are the closer to the original data.

5. Conclusions

Through applying BP neural network and multiple linear regression to predict the second day opening price of 'Gree Electric Appliance', we found that both models can predict the data accurately, but, relatively speaking, BP neural network made the prediction more precisely. A stock is an unstable system with nonlinear dynamic changes, so, with its nonlinear mapping ability, BP neural network works better than multivariate linear regression models. Therefore, it had obvious advantages in the fitness, simulation and prediction of data.

In the aspect of choosing characteristic variables, the technical indicators are flexible and complicate calculations on the stock price, so it is applicable to the neural network which can make nonlinear prediction. However, neural network model needs complex computer programs to achieve desired results. Comparatively speaking, the multivariate linear regression operation is simpler to operate. To reach high-precision and stable prediction results, selecting the conceptual topic index as the characteristic variable is very worth recommending.

With the expanding market capacity of China's stock market, there are varieties of technical means on trading stocks. It is an effective means of investment to choose and predict the stock by using the conceptual topics. The whole stock market needs the guidance and promotion of the conceptual topics. Whether they have good prospects of development, profitability and competitiveness reveals the trend of a stock more intuitively, for the conceptual topics describes the development of the industry a stock belongs to. Therefore, the conceptual topics of a stock cannot be ignored when we are to predict it. Though this method may not be the most accurate, it is wise and safe to choose sustainable industries, this is the point of 'value investment'.

Author Contributions: Zhaoqin Yan,Zhangying Huang and Minwei Liang made equal contributions.

Funding: This research received funding by Nanjing University of Science and Technology undergraduate research training "100 million" plan project funding (national level), project number 201710288065.

Acknowledgments: We thank Professor Peibiao Zhao for his guidline and support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ilaria, B.; Stefano, B.; Guido, C.; Matthieu, C.; Antti, U.; Ingmar, W. Web Search Queries Can Predict Stock Market Volumes. *Plos One* **2012**, *7*, e40014.
2. Hiemstra, Y. Linear regression versus back propagation networks to predict quarterly stock market excess returns. *Computational Economics* **1996**, *9*, 67–76.
3. Safer, A.M. A comparison of two data mining techniques to predict abnormal stock market returns. *Intelligent Data Analysis* **2003**, *7*, 3–13.
4. Meng, D. A Neural Network Model to Predict Initial Return of Chinese SMEs Stock Market Initial Public Offerings. IEEE International Conference on Networking, Sensing and Control, 2008, pp. 394–398.
5. Ilambarathy, R. An Analysis of technical data and develop a model to predict all share price index in Colombo Stock Market by using artificial neural network. *Renewable Energy* **2010**, *35*, 405–411.
6. Guresen, E.; Kayakutlu, G.; Daim, T.U. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications* **2011**, *38*, 10389–10397.
7. Bing, Y.; Hao, J.K.; Zhang, S.C. Stock Market Prediction Using Artificial Neural Networks. *Information Technology for Manufacturing Systems III* **2012**, *6-7*, 1055–1060.
8. Zhao, Y.; Zhang, Y.; Qi, C. Prediction Model of Stock Market Returns Based on Wavelet Neural Network. IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008, pp. 31–36.
9. Shachmurove, Y.; Witkowska, D. Utilizing Artificial Neural Network Model to Predict Stock Markets. *Penn Caress Working Papers* **2000**, pp. 423–430.
10. Torabi-Kaveh, M.; Naseri, F.; Saneie, S.; Sarshari, B. Application of artificial neural networks and multivariate statistics to predict UCS and E using physical properties of Asmari limestones. *Arabian Journal of Geosciences* **2015**, *8*, 2889–2897.
11. Sujatha, K.V.; Sundaram, S.M. Neural Network and PLS Regression Models for Predicting Stock Prices. *Artificial Intelligent Systems & Machine Learning* **2011**.
12. Sousa, S.I.V.; Martins, F.G.; Alvim-Ferraz, M.C.M.; Pereira, M.C. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software* **2007**, *22*, 97–103.
13. Refenes, A.N.; Azema-Barac, M.; Zapranis, A.D. Stock ranking: neural networks vs multiple linear regression. IEEE International Conference on Neural Networks, 1993, pp. 1419–1426 vol.3.
14. Pan, J.; Zhuang, Y.; Fong, S. *The Impact of Data Normalization on Stock Market Prediction: Using SVM and Technical Indicators*; Springer Singapore, 2016; pp. 72–88.
15. Chen, R.C.; Lin, C.M. Forecasting Stock Prices with Technical Indicators Based on Support Vector Machines. *Asian Journal of Information Technology* **2005**.
16. Han, L.S.; Nordin, M.J. Regression techniques for the prediction of stock price trend. International Conference on Statistics in Science, Business, and Engineering, 2012, pp. 1–5.

17. Lei, L.U.; Sciences, S.O. Combinational Stock Price Forecasting Based on Multiple Regression and Technical Analysis. *Journal of Shanghai Institute of Technology* **2014**.
18. Han, L.S.; Nordin, M.J. Integrated Multiple Linear Regression-One Rule Classification Model for the Prediction of Stock Price Trend. *Journal of Computer Science* **2017**, *13*, 422–429.
19. Ahangar, R.G.; Yahyazadehfar, M.; Pournaghshband, H. The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange. *International Journal of Computer Science & Information Security* **2010**, *7*, 248–255.
20. Zhang, J.; Yang, Y. BP Neural Network Model Based on the K-Means Clustering to Predict the Share Price. Fifth International Joint Conference on Computational Sciences and Optimization, 2012, pp. 181–184.
21. Fei, Y.; Xiao, J.; Chen, Y.; Cao, W. *The Combined Stock Price Prediction Model Based on BP Neural Network and Grey Theory*; Springer Netherlands, 2012; pp. 1479–1486.
22. Yang, L.; Wang, Q. Predicting the stock price based on BP neural network and big transaction. International Conference on Fuzzy Systems and Knowledge Discovery, 2012, pp. 2957–2960.
23. Ma, W.; Wang, Y.; Dong, N. Study on stock price prediction based on BP Neural Network. IEEE International Conference on Emergency Management and Management Sciences, 2010, pp. 57–60.
24. Huo, L.; Jiang, B.; Ning, T.; Yin, B. A BP Neural Network Predictor Model for Stock Price. International Conference on Intelligent Computing, 2014, pp. 362–368.
25. Liu, P.; Ren, Y. Bp neural network model for prediction of listing Corporation stock price of Qinghai province. International Conference on Logistics, Informatics and Service Sciences, 2016, pp. 1–6.
26. Xiang, L.I. Prediction of the price of stock based on the BP neutral network. *Journal of Dalian Maritime University* **2008**.
27. Zou, K.; Liu, S.; Xu, L. New Fuzzy BP Neural Network Algorithm and Its Application in Stock Market. International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1–4.
28. Xiaolin, L.; Jie, S. *Data mining and machine learning*; China Renmin University Press, 2016.