ATLANTIS PRESS
Advances in Engineering, volume 126
5th International Conference on Machinery, Materials and Computing Technology (ICMMCT 2017)

A New Employment Forecast Model and Empirical Study Based on BP Neural Network

Huang Rui[1], Chang Xi[1], Zhao Danni*[1,2]
1. Behavioral Finance and Regional Laboratory of Guangdong University of Finance
2. Guangzhou Key Research Bases of Regional Financial Policy of Guangdong University of Finance (Correspondence author)

Abstract:

Since 2006, the labor market in China continued to increase a lot. Based on the analysis of main factors which affecting the labor market, this paper uses BP neural networks based on BFGS to forecast the labor market in China. First of all, dealing with the initial data, try the best to meet the requirements of BP neural network. And then, it is required to accumulate an appropriate BP neural network model, by using the actual data to verify this model. After that, comparing it with traditional statistical models, proving that the prediction model of BP neural network based on BFGS has a higher precision and practicability.
Keywords: BFGS; BP neural networks; Forecast; Countermeasures

1. Introduction

Since the reform and opening up for 20 years, China's economy has maintained a high rate annual growth. It is not only unique in transition countries, but is rare all around the world. Unfortunately, since 1998, the rate of employment goes decline, the employment situation is becoming depressing increasingly. However, around 1988, China appeared the new growth of labor supply, at the same time, the national economy entered a period of adjustment, with the deepening of reform, and many institutions began to save labor and surplus staff, which resulting the current new employment pressure. Therefore, a reasonable perdition of employment decisions becomes the foundation of a stable economic development and important policy for social responsibilities. Thus, the forecast of

the size of the tertiary industry has been paid attention to the academia.

Feed-forward neural network occupies a very important point in the theory of neural network research and modeling of nonlinear systems. The present researches are focus on learning algorithm, error function, the network structure and the convergence and stability of the network weights. It is obvious that the optimization theory can give a lot of weight learning algorithm, including the steepest descent method, which is widely used in recent years. As well as Gauss Newton algorithm and its derivation of Levernberg-Marquardt weight learning algorithm. Many reviews improved the feed-forward neural network based on Gauss Newton method, and achieved good results.

2. BP neural network of Quasi Newton method

The content of labor market is very important to the economic development, if the current labor market has few young people, then it is needed to re-allocation of resources. For example, change the family planning policy to obtain more workers, or migrant workers from abroad, or eliminate those aged. In order to avoid the countries in trouble or facing difficulties, it is believed that future the labor market can be regarded as an inevitable trend to the national future, because outstanding workforces can much national prosperity. Otherwise, labor quality greatly reduced then the country must decline. In recent years, along with the economic development, the overall situation of China's labor market improved a lot. It not only digest the city economic restructuring of employment pressure, but realize the total employment continued to increase, and the unemployment rate remained at a relatively low level. At the same time, because of the role of market mechanism in the allocation of labor resources participate obvious, the integration degree between cities, urban, rural and regional labor market increased gradually. Therefore, an accurate prediction of population of employment not only can provide the basis for the market, but discover problems in time before the problems take place.

This paper applies the BP neural network of quasi Newton method to predict the situation of employment in China. The BP neural network used in this paper has 3 layers, they are output layer, hidden layer and output layer. Each layer has a number of neurons can be different, there are connections between the first neurons layer and the next neurons layer, and the nonlinear of each neuron transfer is a simple type S function.

Therefore, the most common function form is $\gamma(x)=\frac{1}{1+e^{-\mu x}}$. The parameter α >0 can control the slope. It is going to set the input vector in the *n* cycle training is $a(n)=(a_1(n),a_2(n),...,a_r(n))$, where *r* is the number of the input neurons; the target vector of network output is $b(n)=(b_1(n),b_2(n),...,b_r(n))$, and there are *p* hidden layer in the middle; the output vector of the *q* layer

is $b^{(q)}(n) = (b_1^q(n), b_1(n)^q, ..., b_r^q(n))$; $w_{yz}^{(t+1)}(n)$ is the z neurons weights between y neurons and the t layer, where connected to the hidden layer of $t+1$ layer; the threshold $\omega_y^t(n)$ corresponding to the y neurons in the hidden layer of t layer.

If set $b^{(0)}(n) = a(n)$, $b^{(t+1)}(n) = b(n)$, and the output layer neurons will meet:

$$b_i^{(t+1)}(n) = \gamma(\sum_{i=1}^{r} w_{yz}^{t+1}(n) b_i^t(n) - \omega_i^{(t)}(n))(0 \le t \le p)$$

If set $w_{i0}^{(t+1)}(n) = \omega_i^{(t+1)}(n), b_0^t(n) = -1$, and the result goes to:

$$b_i^{t+1}(n) = \varphi(\sum_{i=0}^{t+1} w_{ti}^{t+1}(n) b_i^t(n))(0 \le t \le p)$$

If set $c_i^{t+1}(n) = \sum_{i=0}^{v} w_{ti}^{t+1}(n) b_i^t(n)$, and the answer is:

$$b_i^{(t+1)}(n) = \varphi(c_i^{t+1}(n))(0 \le t \le p)$$

To define the error and the square errors of m neurons in network output layer is $e_m(n) = d_m(n) - b_m(n), E_m(n) = e_m^2(n)$, and the total square error of the output layer is $E(n) = \sum_{i=1}^{m} E_i(n) = \sum_{i=1}^{m} e_i^2(n)$. Therefore, N samples' total square error of a cycle is $E = \sum_{i=1}^{N} E(n)$ .To identify the objective function is E, which is the square of sample output and the neural network output. Around the iteration point x_y , considering the second approximation $E(x_k)$, so:

$$E(x_y + \eta_y) \approx E(x_y) + \eta_y^T g_y + \frac{1}{2} \eta_y^T H_y \eta_y$$

Where k is the number of iterations, $g = \nabla_X E(x_y)$, H_y is a positive definite matrix. If the stable point of the above right function is $\eta_y = -H_y g_y$, the search direction is $\eta_y = -H_y g_y$. Obviously, using one dimension search algorithm, follow the search direction η_y can be obtained the step size α_y , so $x_{y+1} = x_y + \alpha_y \eta_y, \alpha_y > 0$. In order to determine the H matrix, the BFGS quasi Newton formula goes to:

$$H_{y+1} = H_y - \frac{H_y b_y b_y^T + s_y b_y^T H_y}{b_y^T s_y} + (1 + \frac{b_y^T H_y b_y}{s_y^T b_y}) \frac{s_k s_y^T}{s_k^T b_y}$$

Where $s_y = x_{y+1} - x_y$, $b = \nabla E(x_{y+1}) - \nabla E(x_y)$.

It is obvious that in the iterative learning process of neural networks, the convergence speed and performance can be improved by certain rules of the iterative algorithm. For example, using an adjusted decline rule, set $E(x_{y+1}) < E(x_y) + \delta g_y^T \eta_y$, if it can match it, it is going to use quasi-Newton formula to exam it. Otherwise, it is advice to skip the update in *H* array. The specific algorithm is as follows: Firstly, give the output samples, permissible error δ; secondly, when k=0, take a random number as the initial value in α_0, then compute the gradient vector g; thirdly, set H_y is the unit matrix; fourthly, if $\|g\| \le \delta$, stop going forward; fifthly, use first search algorithm for $a_y > 0$, determine the search direction η_y, then work out $x_{y+1} = x_y - \alpha_y H_y \nabla E(x_y)$; finally, if $E(x_{y+1}) < E(x_y) + \delta g_y^T \mu_y$, then H_y will goes to H_{y+1}; if not, set $H_{y+1} = H_y$, *y*=*y*+1, then turns to step four.

3. The empirical analysis

This paper applies three layers BP neural network; builds up a nonlinear mapping model on employment. The inputs in this model are the national employment between 2005 and 2009 in China. And the target vector is the national employment in 2010(All the data come from Zhonghong database). Throughout a repeated neural network training process, and finally confirm a hidden layer nodes, which is 3. The result of neural network topology diagram as shown in figure 1.

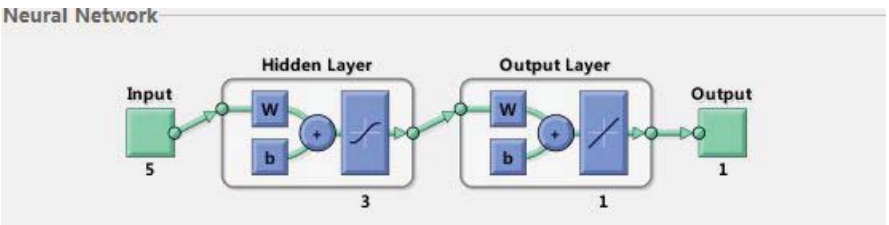

Figure 1: Neural network topology diagram

In order to train the network, neural network uses the data of 2010 as the target vector, and then use the trained network to predict its futures, by using the data from 2005 to 2009.

Table 1: The national employment from 2005 to 2010

Provinces	2005	2006	2007	2008	2009	2010
Beijing	920	920	1111	1174	1255.08	1317.66
Tianjin	427	427	433	503	507.26	520.78
Hebei	3467	3467	3567	3652	3899.73	3790.19
Shanxi	1476	1476	1550	1583	1599.65	1665.08
Mongolia	1041	1041	1082	1103	1142.47	1184.68
Liaoning	1979	1979	2071	2098	2189.96	2238.09
Jilin	1099	1099	1096	1144	1184.71	1248.67
Heilongjiang	1626	1626	1660	1670	1687.47	1743.39
Shanghai	856	856	877	896	929.24	924.72
Jiangsu	3878	3878	4193	4384	4536.13	4731.73
Zhejiang	3203	3203	3615	3692	3825.18	3989.18
Anhui	3485	3485	3598	3595	3689.75	3846.76
Fujian	1868	1868	1999	2080	2168.85	2181.32
Jiangxi	2107	2107	2196	2223	2244.15	2306.09
Shandong	5111	5111	5262	5352	5449.77	5654.67
Henan	5662	5662	5773	5835	5948.78	6041.56
Hubei	2676	2676	2763	2876	3024.48	3116.52
Hunan	3658	3658	3749	3811	3907.7	4007.75
Guangdong	4702	4702	5293	5478	5643.34	5776.93
Guangxi	2703	2703	2760	2807	2862.63	2945.34
Hainan	378	378	415	412	431.45	445.72
Chongqing	1721	1721	1790	1837	1878.48	1912.13
Sichuan	4604	4604	4779	4874	4945.23	4997.61
Guizhou	2216	2216	2283	2302	2341.11	2402.17
Yunnan	2461	2461	2601	2679	2730.2	2814.11
Tibet	140	140	154	160	169.07	175.03
Shanxi	1883	1883	1922	1947	1919.48	1952.03
Gansu	1348	1348	1374	1389	1406.62	1431.86
Qinghai	268	268	276	277	285.54	294.1
Ningxia	300	300	309	304	328.51	325.98
Xinjiang	764	764	801	814	829.17	852.59

Unit：Ten Thousand People

Acquire matlab software as the programming tool. Then the setting parameters are: the maximum number of training is 1000, the learning rate is 0.05, the iterative training process is 2000, and longitude training requirements is 0.0001. Other values are used by default matlab neural network toolbox value. The perdition results show in table 2.

Table 2: The comparison between the prediction results and the actual values and the errors

Provinces	The actual value	the prediction value	Errors	Error rate
Beijing	1317.66	1179.2394	138.420597	11.74%
Tianjin	520.78	538.983731	-18.203731	-3.38%
Hebei	3790.19	3988.91516	-198.72516	-4.98%
Shanxi	1665.08	1617.31356	47.7664383	2.95%
Mongolia	1184.68	1129.04428	55.6357228	4.93%
Liaoning	2238.09	2285.65418	-47.564184	-2.08%
Jilin	1248.67	1176.90484	71.7651551	6.10%
Heilongjiang	1743.39	1739.88579	3.50420885	0.20%
Shanghai	924.72	922.748114	1.97188571	0.21%
Jiangsu	4731.73	4563.88056	167.84944	3.68%
Zhejiang	3989.18	3884.70168	104.478315	2.69%
Anhui	3846.76	3845.72747	1.03253326	0.03%
Fujian	2181.32	2228.12837	-46.808367	-2.10%
Jiangxi	2306.09	2362.96515	-56.87515	-2.41%
Shandong	5654.67	5601.5827	53.0872952	0.95%
Henan	6041.56	6118.28069	-76.72069	-1.25%
Hubei	3116.52	3172.90413	-56.38413	-1.78%
Hunan	4007.75	4035.1949	-27.444904	-0.68%
Guangdong	5776.93	5750.4611	26.4689046	0.46%
Guangxi	2945.34	3049.86453	-104.52453	-3.43%
Hainan	445.72	488.866719	-43.1467	-8.83%

			19	
Chongqing	1912.13	1931.26966	-19.139662	-0.99%
Sichuan	4997.61	5046.00901	-48.399007	-0.96%
Guizhou	2402.17	2481.8445	-79.674495	-3.21%
Yunnan	2814.11	2873.01633	-58.906332	-2.05%
Tibet	175.03	298.743839	-123.71384	-41.41%
Shanxi	1952.03	2010.61491	-58.584913	-2.91%
Gansu	1431.86	1422.80057	9.05943239	0.64%
Qinghai	294.1	384.385495	-90.285495	-23.49%
Ningxia	325.98	415.756604	-89.776604	-21.59%
Xinjiang	852.59	827.664327	24.9256727	3.01%

Unit： Ten Thousand People

Table 2 shows that the predictions in most area are exceptionally accurate, only the error of Tibet, Qinghai, and Ningxia is relatively large. This is fully indicates that the employment market in Tibet, Qinghai and Ningxia is extremely immature and goes to an uncertain trend. Because of the advantage of BP neural network based on the quasi Newton method is its nonlinear approximation, therefore, the employment market in Tibet, Qinghai and Ningxia is not match to its nonlinear approximation, and there is no rules of the employment in these three provinces can be followed at all. On the other hand, the employment in other provinces can be well predicted by using the BP neural network.

4. System performance analysis

As shown in Figure 2, the optimal solution of neural network iterative is obtained after 27 times. When iterative to the twenty-first step, it obtained the best check value, where as shown in the circle. When the system iteration to the 27th step, the training value, testing value and calibration value converge to the optimal value, it is prove that the BP neural network can predict the employment efficiently.

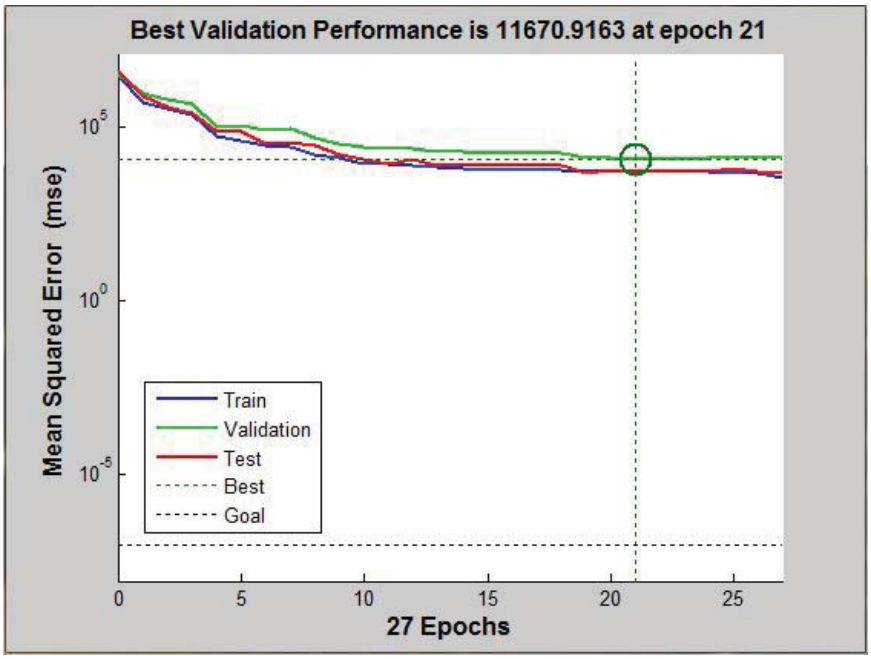

Figure 2: neural network iteration map

As shown in Figure 3, the system after six times check, the convergence of the gradient evident goes to the optimal value. Figure 3 also shows that the BP neural network has good convergence of gradient prediction for the employment in china.

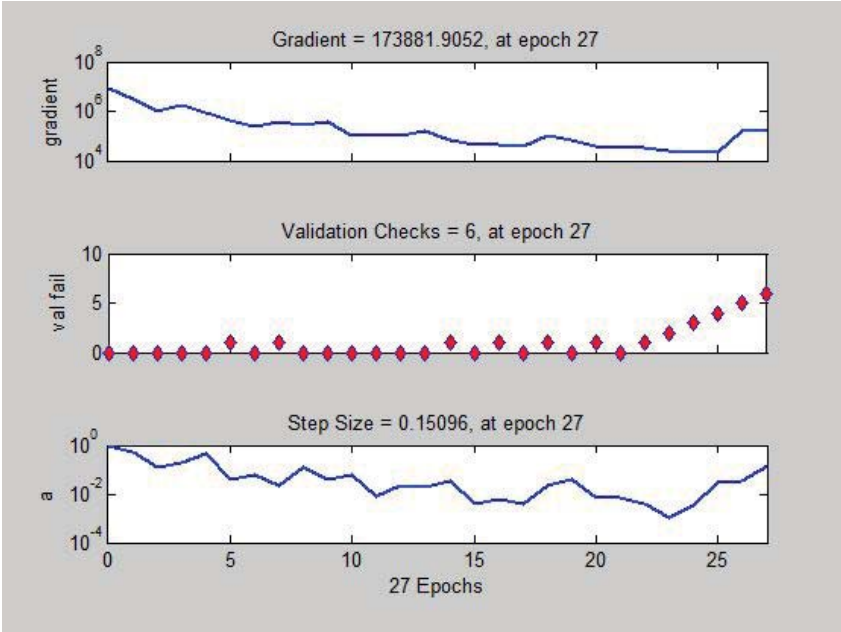

Figure 3: network stability analysis map

As shown in Figure 4, the fitting values of the neural network can fit the data well, where allow the data distributed in the fitting line on both sides evenly. At the same time, the check value and test value of fitting degree is exceedingly high. It is clearly that in table 4, the neural network can show the effect of good nonlinear approximation, the data can be distributed evenly in the fitting line on both sides.

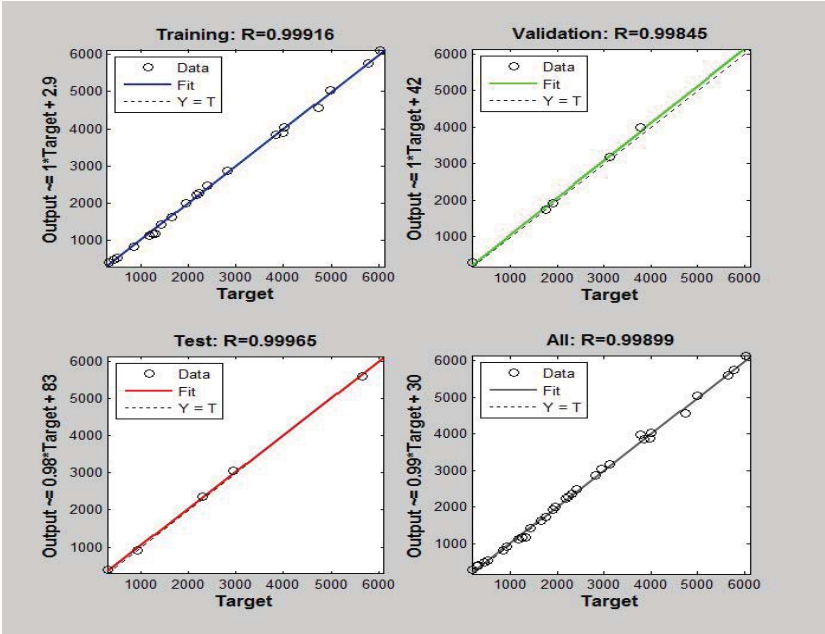

Figure 4: Network training map

5. Conclusion

The prediction model of BP neural network based on BFGS uses the self-learning scheme; modify the weight value during the training process, so that the actual output vector of network is close to the expected output value. Through an example analysis of weight matrix in neural network, revised the weights and threshold with the target vector in the whole network. Finally get the predictive value, the error of the predicted value and the target vector is in an acceptable range. Therefore, it is going to draw a conclusion that is the prediction model of employment by using BP neural network can make an accurate evaluation to the employment industry. It also can be inferred that BP neural network based on BFGS has a great potential development to the future employment in China.

Acknowledgement

This paper is financially supported by the youth project of 2016 Guangzhou Philosophy and Social Sciences and the Project (GD15XYJ09) of Guangdong Philosophy and Social Sciences

Reference

[1] Yuan Yaxiang. Nonlinear programming algorithm. [M] Shanghai: Shanghai science and Technology Press, 1992
[2] Chen S, Billing S A. Neural network for nonlinear dynamic system modeling and identification [J]. Journal of Control, 1992, 56(2):319- 346.
[3] Malik M. The early restart algorithm [J]. Neural Computation, MIT, 2000(12) : 1303- 1312.
[4] Chang K S, Abelg. A universal neural net with guaranteed convergence to zero system error [J]. IEEE Trans. Signal Processing, 1992, 40(12): 3022- 3031.
[5] He Shudong. On the research of multilayer feed-forward neural network structure [J]. Control theory and applications, 1998 (4): 313- 318
[6] Li Yibo, Huang Xiaoyuan, Wu Zhihong. The modified Gauss model neural network chromatographic resolution of overlapped peaks [J]. Computers and applied chemistry, 2001 (9): 484- 488
[7] Xu Chunhui, Xu Xiangdong. A new learning algorithm for feed-forward neural networks [J]. Journal of Tsinghua University: Natural Science Edition, 1999 (3): 1- 3
[8] Shi Hongbao. Neural network and its application. [M] Xi'an. Xi'an Jiao Tong University press, 1993