

# Study of Personal Credit Evaluation Method Based on PSO-RBF Neural Network Model\*

Shuai Li, Yuanmei Zhu, Chao Xu, Zongfang Zhou<sup>#</sup>

School of Management and Economics, University of Electronic Science and Technology of China, Chengdu, China.  
Email: <sup>#</sup>zhouzf@uestc.edu.cn

Received April 1<sup>st</sup>, 2013; revised May 10<sup>th</sup>, 2013; accepted June 10<sup>th</sup>, 2013

Copyright © 2013 Shuai Li *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

Personal credit evaluation is the basic method for the commercial banks to avoid the consumer credit risk. On one hand, the credit behavior of individuals is complex; on the other hand the personal credit assessment system in our country is not sound, assessment methods are mostly objective, therefore, more and better scientific methods for credit risk assessment need to be introduced. This paper proposed a method for personal credit evaluation based on PSO-RBF neural network, which used PSO algorithm to optimize the parameters of RBF neural network, then applied the optimized RBF neural network in the personal credit evaluation. This method combined the global searching ability of PSO algorithm and the high effectiveness of local optimize of RBF together, overcame the unstability of PSO algorithm and the drawback of RBF which easily leads to local minimum. The result shows that the personal credit assessment method based on PSO-RBF neural network is highly accurate in classification and prediction, and is suitable in personal credit assessment and prediction.

**Keywords:** Personal Credit Evaluation; PSO Algorithm; RBF Network; PSO-RBF Model

## 1. Introduction

With the rapid development of Chinese economy, the personal credit is being taken more and more attention, which strongly challenges the credit evaluation system in our country. At present, the personal credit problem in our country has seriously affected the development of Chinese economy, therefore, to build scientific personal credit evaluation system, and to introduce in scientific personal credit evaluation is in urgent need. Regards to the commercial bank, personal credit assessment is to make comprehensive judgment and evaluate the possibility of the borrower to return the money on time, according to the observation of the indexes which reflect the personal credit ability, and so as to make a decision on whether lend money [1].

In western countries, the theories and methods on credit risk evaluation are maturing, and many interdisciplinary theories and methods are applied to credit risk evaluation [2]. Personal credit risk evaluation methods

are separated into two types, statistic and non-statistic method, and the statistic method includes decision theory method, logic regression, linear and non-linear regression, *K*-nearest-neighbor estimation etc, non-statistics method includes mathematical programming, neural network, genetic algorithm, expert system and classification tree method etc. [3]. Particle swarm optimization (PSO) [4] is a recently developed numerical method for optimization, which is simple, easy to apply and has a strong smart background, and it has been used in many fields such as function optimization, and pattern recognition. In 2007, Fu Peizhong, Ying Yixin [5] pioneered in apply PSO algorithm in determining the parameter of RBF neural network (connection weight, centers and widths of the hidden units), and checked its validity using simulation. RBF neural network is a three-layered BP network, which can approximate the continuous function with arbitrary precision. It is characterized by single best approximation, no local minimum, less calculation and fast learning, and is widely applied to pattern classification, system identification and functional approximation. At present, PSO-RBF neural network model [5] is applied to many industries, such as water quality assessment [6], short time traffic flow prediction [7], and boiler super-

\*Funding Project: Natural Science Foundation of China (71271043); The Special Research Foundation of Ph.D. Program of China (2011 0185110021) and Sichuan Province Science and Technology Support Project (2012SZ0001).

<sup>#</sup>Corresponding author.

heat steam temperature control systems identification [8] etc.

This paper built personal credit assessment model by using RBF neural network, and chose PSO algorithm to train and optimize the network, then built personal credit assessment model based on PSO-RBF neural network, and made a further comparative analysis of it with single RBF neural network, the second chapter of this paper provided the modeling solution; the third chapter processed the sample data; the fourth chapter analyzed the result of simulation.

## 2. Basic Theory and Modeling Solution

### 2.1. RBF Neural Network Model

RBF neural network [9] is firstly proposed by J. Moody and C. Darken, in the late 1980s, it is a three-layered BP network with a single-hidden Layer. The first layer is the input layer, constructed with signal source; the second layer is the hidden layer, its action function is Gaussian function, the hidden layer proceed space mapping and transformation to the input information; the third layer is output layer, its action function is linear function, the output layer linear weighted the information from the neuron of the hidden layer, then output the result of the entire network. Its structure is as shown in **Figure 1**.

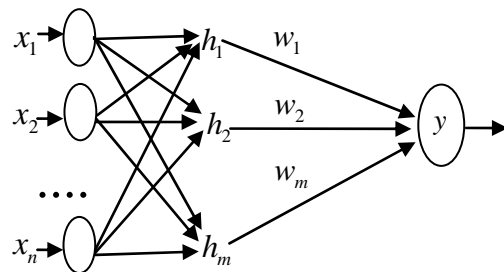
The figure above is the FBF neural network, specially, in this paper, 17 indicators are selected by the inputs layer in the article. Gaussian basis functions are used in Hidden layer. The hidden layer output results are weighted to get the final output, which means that the results of the assessment of the personal credit. As is known, generally, the RBF Neural Network has  $n$  inputs,  $m$  hidden node, 1 output; the action function of the radial base layer (hidden layer) is Gaussian function, the output of this layer neuron  $j$  is:

$$h_i = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right), i = 1, 2, \dots, m \quad (1)$$

Among them,  $x$  is an  $n$  sphere network input vector,  $c_i$  is a center vector with sphere of  $x$ , is the basic width. The output of RBF neural network is:  $y = \sum_{i=1}^m w_i h_i$ , among them,  $w_i$  is the network weight vector between the  $i$ th radius basic layer and output layer. When applying RBF neural network, three parameters need to be determined: the center vector of the Gaussian function  $c_i$ , the basic vector  $\sigma_i$  and the network weight  $w_i$ .

### 2.2. Weighted PSO Algorithm

Particle Swarm Optimization (PSO) [10-12], proposed by Kennedy and Eberhart in 1995, is a Cluster optimization calculation method, which can be used as neural network



**Figure 1. The structure of RBF neural network based on gaussian function.**

training method. Its basic idea is to update the local optimal  $p_{best} = (p_1, p_2, \dots, p_d)$  and the global optimal  $g_{best} = (g_1, g_2, \dots, g_d)$  of the particles with the fitness (the fitness function normally is determined by the optimized function in practical issues) of each particle.

The particle updates their speed and location, by using the following formula to dynamic trace the optimal of individuals and the global optimal.

$$v_{ij}(t+1) = wv_{ij}(t) + c_1r_1[p_j(t) - x_{ij}(t)] + c_2r_2[g_j(t) - x_{ij}(t)] \quad (2)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (3)$$

Among them,  $j = 1, 2, \dots, d$ ;  $t$  is the number of iteration;  $r_1, r_2$ , are the random number uniform distributed in  $(0,1)$ ;  $c_1, c_2$  are the accelerated factor, generally  $c_1 = c_2 = 2$ ;  $w$  indicates the influence of the previous speed to the present.

For better searching results, this paper adopts the improved weighted PSO algorithm. The weigh  $w$  maintains the balance of global and local searching capacity, high  $w$  inclines to global search, otherwise inclines to local searching [13]. This paper chose the following mean of square error function as the fitness function of the PSO algorithm.

$$MSE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (y_{j,i}^d - y_{j,i})^2 \quad (4)$$

Among them,  $n$  indicates the number of the input variables,  $m$  indicate node number of the hidden layer,  $y_{j,i}^d$  indicates the idea output value,  $y_{j,i}$  indicates the actual value.

### 2.3. RBF Neural Network Optimization Process Based on PSO Algorithm

As is mentioned above, there are three important parameters in RBF neural network, namely the center of the Gaussian function  $c_j$ , the basic vector  $\sigma_i$  and the network weight  $w_i$ . However, Currently, it is difficult to obtain the best values of the parameters of the network structure in theory, and the methods of training RBF

neural network parameters have a major impact on the performance of the network, so we must select the appropriate parameters to improve the prediction performance of the RBF neural network. This paper chooses particle swarm optimization (PSO) to determine the parameters of RBF neural network, so that they are no longer random. The optimized RBF neural network, compared to the independent PSO algorithm and RBF neural network algorithm, has a better evaluation effect on the assessment of the personal credit with higher prediction accuracy. Specific training steps [14] are as follows:

- 1) To set the related parameters of the PSO algorithm, initialize the particles, including the value assignment for hidden center vector, base width vector, weight of the network;
- 2) Calculate the fitness value of each particle according to Formula (4), and make the current position of the particle as the individual maximum  $p_{best}$ , find out the particle of the minimum fitness value, and make it the initial  $g_{best}$ ;
- 3) Compare fitness value of the current particle with  $p_{best}$ , if the fitness value of the current one is smaller, then update  $p_{best}$  with the current fitness value;
- 4) For each particle, compare  $p_{best}$  with  $g_{best}$ , if  $p_{best}$  is better,

- 5) Update the speed and location of the particle according to formulas (2) and (3);
- 6) Repeat steps 4)-6), until the terminal condition is met, which is either meet the maximum iterations or the error accuracy requirement;
- 7) Set  $g_{best}$  as the parameter of the RBF neural network. Specific processes are shown in **Figure 2**.

### 3. Sample Data and Preprocess

#### 3.1. Sample Data and the Selection of Variables

Because of the difficulty in collecting domestic credit data, this paper uses the personal credit data of a Germany bank, a total of 1000 data. Considering the specific circumstance of Germany, removing too much index [15] will influence the effect of the model, therefore this paper removed two indexes, “the purpose of loan” and “nationality”, which have little influence to the model. Specific indexes and quantification method are shown in **Table 1** [16].

First, this paper using stratified sampling, divided the 1000 sample clients into “good credit” and “bad credit”, with 700 and 300 respectively, and for each sample contains 17 attribute indexes. Then, 350 “good credit” and

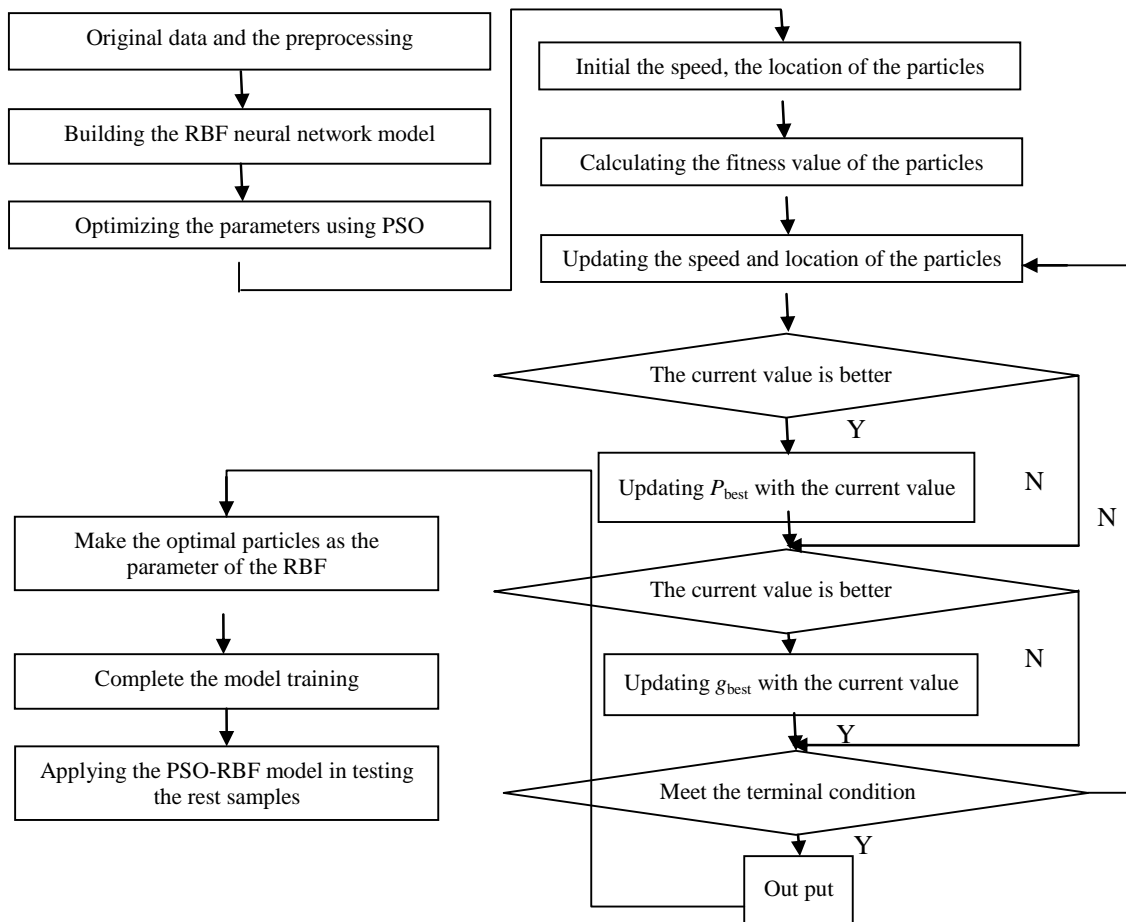


Figure 2. The flow chart of PSO-RBF neural model.

**Table 1. Variable and its quantification.**

Indexes	Variable	Definition
Age	$X_1$	Actual value
Marriage	$X_2$	1 = single; 2 = married
Supporting family members	$X_3$	Actual value
Occupation	$X_4$	1 = unemployed/manual workers, non-resident; 2 = non-proficient worker, resident; 3 = proficient worker/officer; 4 = manager/independent entrepreneurs
Year of working	$X_5$	1 = unemployed; 2 = less than 1 year; 3 = 1 - 4 years; 4 = 4 - 7 years; 5 = more than 7 year
Housing condition	$X_6$	1 = rent; 2 = owned; 3 = free housing
Year of live in current house	$X_7$	Actual value
Installment to deposable disposable income rate	$X_8$	Actual value
assets	$X_9$	1 = real estate; 2 = if not 1: agreement of public construction savings/life insurance; 3 = if not 1 or 2: Automobile or other; 4 = vain
Current payment account status	$X_{10}$	1 = less than 0 mark; 2 = 0 - 200 dollar; 3 = more than 200 dollar or Salary contract has been signed for at least a year; 4 = no payment account
The rest plan for the installment	$X_{11}$	1 = bank; 2 = stock; 3 = no
Debt amount	$X_{12}$	Actual value
Saving account / bonds	$X_{13}$	1 = less than 100 mark; 2 = 100 - 200 dollars; 3 = 500 - 1000 dollars; 4 = more than 1000 dollars; 5 = no saving account/bonds
Loan Period	$X_{14}$	Actual value
Credit record	$X_{15}$	0 = no bad credit record; 1 = has overdue payment record/other bad credit record; 2 = overdue payment; 3 = has late payment record; 4 = no credit record/credit record is no in this bank
Existing loan project number in this bank	$X_{16}$	Actual value
other note debtor/guarantor	$X_{17}$	1 = no; 2 = joint applicants; 3 = secured

150 “bad credit” from the sample are selected as the training sample; the rest 500 samples are used in testing the model, to exam the classification effect of the model.

**3.2. Normalizing the Data**

To accelerate the convergence of the model and reduce the effect to the network classification caused by the unbalance of the data, first normalize the training sample and the test sample. Divide the adopted 17 explanatory variables into discrete variable and continuous variable.

For the discrete variable, use the Minimum and maximum normalization method to process.

$$X = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

Among them,  $X \in [0,1]$  is the value of the variable after normalizing,  $x_{\min}$ ,  $x_{\max}$  are the maximum and minimum value of the variables.

For the continuous variables, through observing the distribution of the variables , we found that the amount of the loan  $x_{12}$ ,and the time of the loan  $x_{14}$  both approximately obey the law of normal distribution, that is

$x \sim N(u, \sigma^2)$  , and convert them into [0,1].

$$X' = \Phi\left(\frac{x-u}{\sigma}\right) \tag{6}$$

Among them,  $\phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$  is the standard normal distribution, through converting, we can get the corresponding probability.

**4. Simulation and the Result Analysis**

In the process of designing RBF neural network, the key is the determining of the number of the hidden neurons, because there are many parameters in this paper, when we set the number of the hidden neurons as 3, the result meet the setting condition, and the input parameters has 17 variables, then the structure or RBF network is 17-3-1. To obtain better training result, this paper chose the improvement PSO algorithm with the inertia weight, the inertia weight  $w = 0.1$ , the learning factor  $c_1$   $c_2$  are both 2, and, the maximum iterations  $t_{\max}$  is 1500, which is the termination condition of the algorithm. To keep the con-

sistency of the data, we chose the 500 sample data in the testing sample as the fundament of the assessment, when evaluating the credit of the creditor, use 0.5 as the critical value, the one whose predicted value is higher than 0.5 is viewed as “good credit” client, otherwise is “bad credit” client, the result of the comparison is shown in **Table 2**.

In practice, there are two types of errors: type one and type two, the former is to misjudge the good credit client as bad credit client and rejects their loan application; the latter is to misjudge the bad credit client as good credit client and accept their loan application. Therefore type two errors cause more damage to the bank, and the false rate of it should be reduced.

From the table above, it can be seen that PSO-RBF neural network when compared with the pure RBF neural network model, both the type one error, and type two

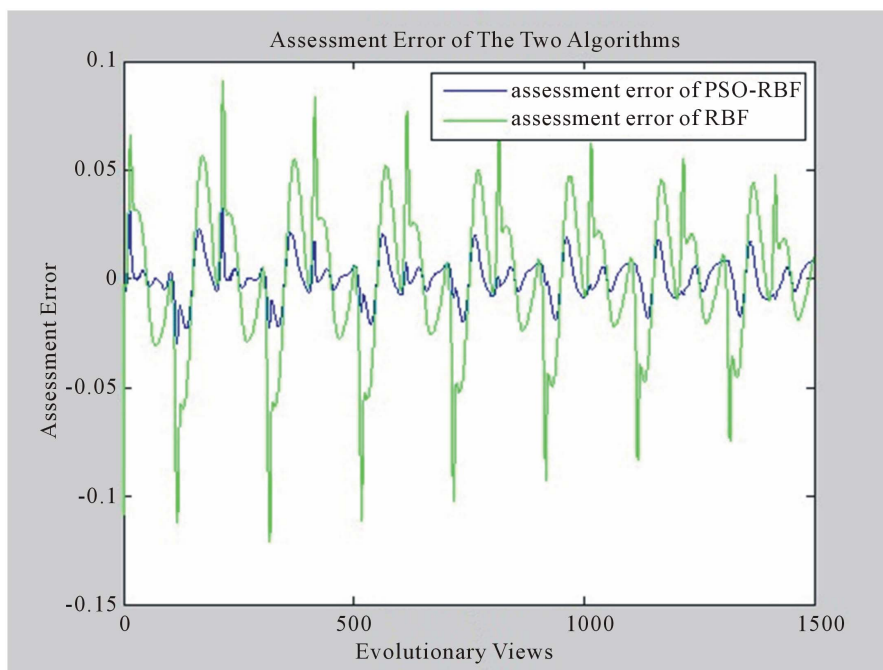
error rate are reduced, therefore the total false rate is reduced, what’s more, the type two error rate is reduced from 8.67% to 4%, namely, the probability of misjudging “bad credit” client as “good credit” client is reduced, which has more practical value for the commercial bank to avoid the personal credit risk.

To test the perdition accuracy of the model, after 1500 iterations, we get the following evaluation error chart

As the below **Figure 3** shown, evaluation error rate of PSO-RBF model is apparently lower than RBF model, which means PSO-RBF model is effective and of better accuracy. If adopting the RBF model directly, because the initial value of the network is random, the evaluation effect is poor. In conclusion, first adopt PSO algorithm to adjust and optimize the parameter of the RBF model, and reduce the inner error of the model, and then apply the

**Table 2. Comparison between RBF neural network and PSO-RBF neural network model prediction accuracy.**

Model	Actual value	Predicted value		Sample number	Accuracy (%)	False rate (%)
		1.00	0.00			
RBF	1.00	317	33	350	90.57	9.43
	0.00	13	137	150	91.33	8.67
	Total	330	170	500	90.80	9.20
PSO-RBF	1.00	326	24	350	93.14	6.86
	0.00	6	144	150	96.00	4.00
	Total	332	168	500	94.00	6.00



**Figure 3. Evaluation error curve of PSO-RBF model and RBF model.**

optimized model in the personal credit evaluation, the evaluation result is more accurate than the single RBF model.

## 5. Conclusions

This paper adopting PSO algorithm in the training of RBF neural network, proposed PSO-RBF neural network method, used PSO algorithm to proceed global dynamic searching, and used RBF neural network to proceed local optimizing, finally got the PSO-RBF neural network. This paper first adopted PSO-RBF neural network in personal credit risk evaluation, and through simulation and the empirical study of the 1000 personal credit sample data from a Germany commercial bank, indicated that this method when compared with the pure RBF model, showed higher classification and evaluation accuracy, reduced the type two error, and effectively reduced the personal credit risk that the commercial bank facing. Therefore PSO-RBF neural network is applicable in personal credit risk evaluation. There is potential improvement in this paper, mainly in three aspects:

1) According to reducing the type two error, how to choose the fitness function of the particles in the PSO algorithm;

2) The number of the hidden neurons in the RBF model can be found through other optimizing algorithm, so as to improve the performance of the model;

3) Because there are too many input variables, the convergence of the model is comparatively slow, the time performance is poor.

## REFERENCES

- [1] J. Z. Dai, "Exploration on the building of Personal Credit Risk System," *Financial Practice*, Vol. 15, No. 3, 2002, p. 11.
- [2] Z. F. Zhou, *et al.*, "The Study of the Evolution Mechanism and Evaluation Methods of Emerging Technology Corporate Credit Risk," Science Press, Beijing, 2010.
- [3] W. E. Henley, "Statistical Aspects of Credit Scoring," The Open University, Milton Keynes, 1995.
- [4] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization," *Proceedings of International Conference on Neural Networks*, Orlando, 1997, pp. 1942-1948.
- [5] P. Z. Fu and Y. X. Ying, "The Study of Neural Network Based on PSO-RBF and Its Application in Credit," *Chinese Journal of Scientific Instrument*, Vol. 8, No. 1 2007, pp. 464-466.
- [6] D. S. Wang, S. H. Li and X. P. Zhou, "The Method and Application of Raw Water Quality Evaluation Based on PSO-RBF Neural Network Model," *Journal of Southeast University* (Natural Science Edition), Vol. 41, No. 5, 2011, pp. 1019-1023.
- [7] M. F. Feng and J. C. Lu, "Use Particle Swarm to Optimize RBF Neural Network in Short-Term Traffic Flow Prediction," *Computer Simulation*, Vol. 27, No. 12, 2010, pp. 323-326.
- [8] B. X. Xiao, X. W. Wang and Y. F. Liu, "Simulation Research of Boiler Superheat System Identification Using MPSO-RBF Optimization Strategy," *Journal Of System Simulation*, Vol. 19, No. 6, 2007, pp. 1382-1386.
- [9] L. L. Li, "RBF Neural Network Integration Research and the Application in the personal Credit Assessment," Shandong Normal University, Jinan, 2008.
- [10] S. Yin and M. H. Jiang, "Personal Credit Assessment Combination Forecast Model Based on PSO," *Economic Research Guide*, Vol. 3, No. 14, 2208, pp. 83-86.
- [11] M. H. Jiang and X. C. Yuan, "The Construction and Application of PSO-SVM Model in Personal Credit Assessment," *Journal of Management*, Vol. 5, No. 4, 2008, pp. 511-516.
- [12] L. Y. Yu, "Evaluation Prediction Model Research on Commercial Bank Credit Risk," *Journal of Management*, Vol. 6, No. 5, 2003, pp. 46-52.
- [13] M. H. Jiang, "Research on Combination Forecast Method for Commercial Bank Personal Credit," Harbin Institute of Technology, Harbin, 2006.
- [14] H. Y. Tang and Z. G. Xu, "Neural Network Based on PSO-RBF and Its Application. Metallurgical Automation," 2006, pp. 464-466.
- [15] J. Wang, "The Establishment of Personal Credit Evaluation Index in Consumer Credit," *Rural Finance Research*, Vol. 21, No. 1, 2001, pp. 48-50.
- [16] H. Z. Huang, Z. F. Zhou and J. K. Yu, "ILMBP Neural Network Model and Its Application in the Personal Credit Assessment," *Managerialist*, Vol. 3, No. 10, 2010, pp. 1-3.