ATLANTIS PRESS

Analysis and Optimization of Personal Credit Risk Assessment Model Based on Improved BPNN

Zheyou Guo[1,a*]

[1]Glasgow College, University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu 611731, China

[a]zheyou__guo@126.com

Keywords: Personal credit risk assessment, 5C evaluation method, Time stamp, BPNN

Abstract. In the current situation of "Internet plus Finance", the assessment of personal credit risk is of vital importance to the sustained and steady growth of the whole social economy. In order to more effectively complete the personal credit risk assessment, the innovation of this paper lies in the introduction of time stamp into the 5C evaluation method, and then using the artificial intelligence attribute (crossover factor and mutation factor) of Back Propagation Neural Network(BPNN) to build a personal credit risk assessment analysis. With the dynamic development of thinking, this model constantly adjusts and optimizes the personal credit risk assessment process. Simulation experiment shows that this integrated innovative approach not only reduces the error rate value of personal credit risk assessment, but also improves the efficiency of personal credit assessment.

1. Introduction

With the era of "Internet plus financial", people's consumption concept and mode has been changed. More and more people gradually accept the mode of personal credit consumption, and the proportion of personal credit in bank lending business is increasing [1, 2]. However, due to the lack of personal credit, the risk of personal credit is becoming increasingly prominent, and the bad personal credit business will lead to a sharp increase in the non-performing assets of the bank, which has become an important obstacle to the rapid development of the bank. How to establish a scientific system of personal credit assessment to prevent the risk of personal credit in the business of commercial banks is an urgent problem for commercial banks [3].

In order to effectively assess and prevent personal credit risk, this article is based on the principles of "5C" (Character, Capacity, Capital, Collateral, Condition) in the field of personal credit assessment, introduces the time stamp to assist the neural network artificial intelligence dynamic analysis , and achieves an accurate quantitative analysis of personal credit risk assessment[4,5].

2. Design of Personal Credit Risk Assessment Process Based on Improved BPNN

If a comprehensive analysis of the personal credit risk assessment is conducted, it is necessary to carry out a relevance assessment of the individual's credit risk during all growth stages. Firstly, using time stamps to realize the interconnection of personal credit collection elements in terms of time; secondly, 5C qualitative evaluation methods are used; thirdly, an improved BPNN analysis method is used; finally, personal credit risk analysis reports are implemented.

The design of personal credit risk assessment process based on an improved BPNN is shown in Fig.1.
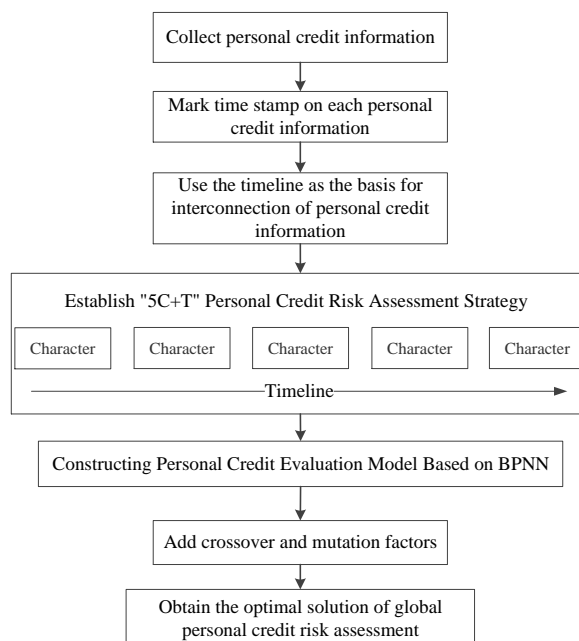
Fig.1 Personal credit risk assessment process based on improved BPNN

2.1 Collection and interconnection of personal credit information

2.2.1 Collection of personal credit information

Based on international standards, domestic and foreign commercial bank experience, and personal credit rating assessment methods, this paper extracts basic customer profiles and customers from three aspects: personal basic information, personal economic status, and personal credit history during the design of a neural network simulation credit evaluation model. There are three indicators of repayment ability and customer credit, and a total of 32 credit evaluation indicators [6].

Specific personal credit collection information includes customer name, ID number, gender, date of birth, household registration, telephone number, mailing address, zip code, education level, job title, work unit, working life, personal insurance, marital status, Health status, whether there is a child, monthly income level, debt ratio, family monthly income, ownership of the house, ownership of automobiles, total value of family assets assessment, loan amount, bank deposits, marketable bonds, credit card processing period, personal loan records Years, credit card credit records, personal judicial records, Internet real-name system behavior records, other reputation records, and cumulative default rates for personal loans.

2.2.2 Interconnection of personal credit information

At the same time, each personal credit data is bound to a time stamp in order to strengthen the internal correlation of personal credit information and predict future development trends. Through effective timestamp technology, it provides referenceable time-retrospective attributes for subsequent iterative and back-propagation feedback processes in BP neural network modeling, which enhances the interconnection of personal information data and provides an important reference quantitative analysis of personal information risk assessment [7].

2.3 Establishing "5C+T" Personal Credit Evaluation Strategy

On the basis of the collected personal information data, the personal credit evaluation strategy of the "5C" assessment method and the timeline is formulated in two dimensions. This strategy not only provides a criterion for evaluating indicators, but also provides a time marking concept for evaluation results.

Specific "5C+T" personal credit assessment strategy design steps are as follows:

a. Character+T: In terms of character, time stamp is used to obtain comprehensive character performance values mainly from individual's comprehensive performance in family, community, unit, society, and internet;

b. Capacity+T: In terms of Capacity, time stamp is used to obtain comprehensive Capacity

performance values, mainly from the individual's overall ability to earn a living, learn, create, realize, and repay;

c. Capital+T: In terms of Capital, time stamp is used to obtain the comprehensive Capital performance values, mainly from the individual's comprehensive capital performance on the income of work, opyright, reputation, lease and so on;

d. Collateral+T: In terms of Collateral, time stamp is used to obtain the comprehensive Collateral performance values, mainly from the comprehensive mortgage performance of the individual on fixed assets, negotiable securities, valuables and other mortgages;

e. Condition+T: In terms of the Condition, through the time stamp is used to obtain the comprehensive Condition performance value , mainly from the individual in the human resources, promotion channels, job hopping space, network appeal and other conditions of the comprehensive conditions;

f. The above steps are intertwined and continue to process the input personal credit parameters continuously to obtain quantifiable personal information risk assessment indicators.

2.4 Design of Personal Credit Evaluation Model Based on BP Neural Network

The learning rules of BP neural network are supervised learning. In the process of personal credit risk assessment, the expected output values of the BP neural network are calculated first, and then the error between the actual and expected output of the unknown individual credit sample is calculated in the training process. The size and direction of the error are used to modify the weight and adjust it repeatedly until the error reaches the expected accuracy. The overall learning process can be divided into the following steps.

a. Determine the number of nodes in the input layer (individual credit information), hidden layer (individual credit information interconnection relationships) and output layer (credit risk assessment results) in the BPNN, and select the appropriate connection structure between the layers.

b. Preset neural network key parameters, such as network weight, learning rate, error allowable range, etc. The default value should be a small random number to avoid over learning.

c. Select the moderate number of learning samples and normalize them, then input the normalized samples to the network input layer and calculate the output of the network.

d. The error between the output result and the expected value in the neural network is obtained. The error value is then transmitted back to the input layer through the output layer, and the weight of each node is adjusted. When the error is lower than the allowable range of the error, the learning is stopped.

e. The new test indicator vector is input to the neural network after learning and training, and the predicted output vector result is obtained.

3. Improved BPNN Mathematical Model Design

According to the BPNN architecture, the input layer variables, the hidden layer correlation content and the output layer evaluation results are determined in turn. At the same time, the crossover factor and mutation factor are introduced to improve the convergence speed which is not ideal and the accuracy of the judgment which is not high [8, 9].

The design flow of the improved BPNN mathematical model is shown in Fig.2 [10].

3.1 Determining the parameters of the input node l

The input node parameters are selected from the content of the personal credit collected, a total of 27 credit evaluation indicators, $Item_{Input}(I_{x1}, I_{x2}, \cdots I_{xn})$. I_{xi} indicates the content of a credit in the input layer.

3.2 Determining the content of the hidden layer

The three evaluation indexes of personal basic information, personal economic situation and personal credit record are used as hidden layer nodes, that is, the number of hidden layer nodes is 3.
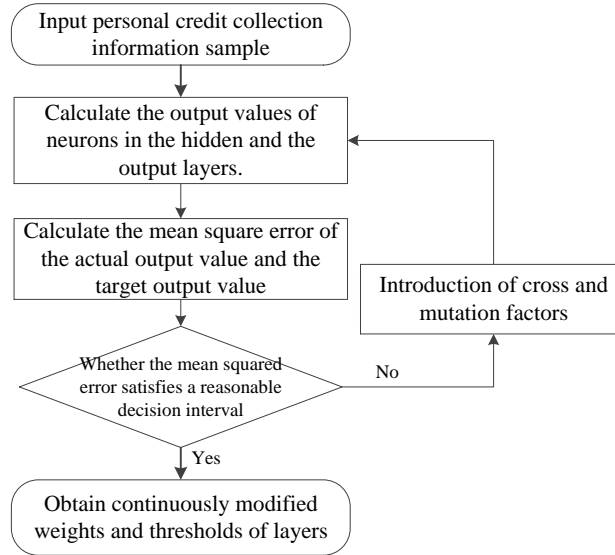
Fig.2 Improved BPNN mathematical model design process

3.3 Determining the decision result of the output node

The results of the output node are five levels: no risk, low risk, medium risk, high risk and maximum risk.

3.4 Construction of BP neural network computing model

Initialize the BPNN. According to the number of indicators and the number of output predicted by the system, the number of input nodes NoIn, the number of output layer nodes NoOut, the number of hidden layer nodes NoIm of the BPNN are determined. Initialize the connection weights between the input layer, the hidden layer and the output layer. The weights of the connections are ω_{ij} and ω_{jk}. Initialize hidden layer threshold value I, output layer threshold value O. Set learning rate and neuron activation function.

Hidden layer output calculations: according to the input variable Ix, the connection weight value ω_{ij} between the input layer and the hidden layer and the implicit layer threshold value Imv, the output value ImOutput of the hidden layer can be calculated.

$$\mathrm{Im}\,Output_j = f(\sum_{i=1}^{n}\omega_{ij}I_{xi} - \mathrm{Im}_{vi}) \qquad j = 1, 2, \cdots, No\,\mathrm{Im}$$

(1)

In the formula (1), f is the incentive function of the hidden layer.

Output layer output calculation: according to the output value ImOutput of the hidden layer, the weight value ω_{jk} and the threshold value O are connected, and the predicted output value Output can be calculated.

$$Output_k = \sum_{j=1}^{No\,\mathrm{Im}} \mathrm{Im}\,Output_j \omega_{jk} - O_k \qquad k = 1, 2, \cdots, NoIn$$

(2)

Adaptive adjustment calculation: in order to ensure the convergence of BPNN and the precision of output data, the cross factor Pc and mutation factor Pm are introduced to realize the adaptive adjustment function of the BPNN algorithm. When the output result of personal credit assessment falls into the trend of local optimal solution, Pc and Pm are correspondingly increased. When the output result of personal credit assessment diverges in the solution space, Pc and Pm are reduced.

$$P_c = \begin{cases} \dfrac{c_1(f_{Fit\max} - f_{Fit0})}{f_{Fit\max} - f_{Fitavg}} & f_{Fit0} \geq f_{Fitavg} \\ c_2 & f_{Fit0} < f_{Fitavg} \end{cases}$$

(3)

$$P_m = \begin{cases} \dfrac{c_3(f_{Fit\max} - f_{Fit})}{f_{Fit\max} - f_{Fitavg}} & f_{Fit} \geq f_{Fitavg} \\ c_4 & f_{Fit} < f_{Fitavg} \end{cases}$$

(4)

In the formula (3) and (4), fFitmax represents the maximum fitness value in the evaluation results of the hidden layer; fFitavg represents the average fitness value of the evaluation results for each hidden layer; fFit0 represents the higher fitness value of the hidden layer data in the two credit indicators to be crossed, and the fFit represents the fitness value of the mutation individual. c_1, c_2, c_3 and c_4 take the value of the (0,1) intervals to be adjusted adaptively.

4. Simulation and Explanation

4.1 Experimental Environment

The experimental parameters of this personal credit risk assessment model are shown in Table 1.

Table. 1 Statement of Experimental Environment

No.	Item	Details
1	The number of input items in the model input layer	27
2	The number of the model hidden layers	3
3	The number of the candidate form in the model output layer	5
4	The shortest validity period of time stamp	10 Years
5	The number of original samples for personal information risk assessment	1000

5. Simulation and Explanation

In this mathematical model, the objective information collected by personal credit is used as the input of credit risk assessment, and the calculation process of the whole simulation model is associated with the time stamp, and the evaluation time and the accuracy of the evaluation results are considered as the assessment items.

Based on the two criteria of error rate and time-consuming evaluation, the following 5C+T-IBPNN and 5C-BPNN are used for comparative analysis.

From the analysis of Fig.3, we can see that by using the 5C+T-BPNN method, when the individual credit collection sample is iterated about 300 times, the evaluation error rate is basically maintained within 2%. By using the 5C-BPNN method, when the individual credit collection sample is iterated about 450 times, the evaluation error rate is basically maintained within 16%. This shows that 5C+T-BPNN can obtain a good sample evaluation result fitting degree in a smaller sample iteration number process.
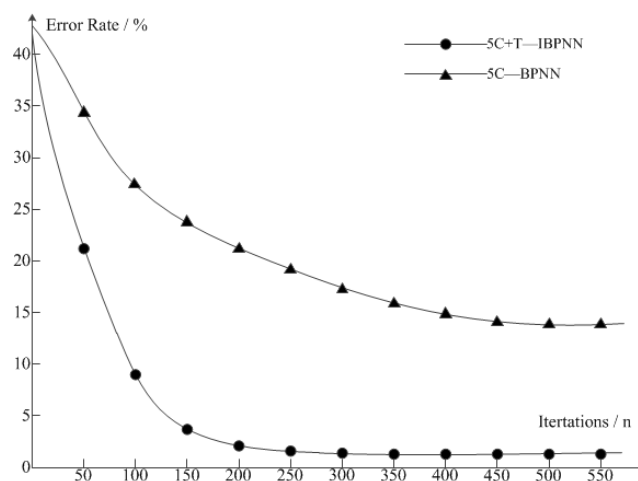

Fig.3 Comparison of error rates between 5C+T - IBPNN and 5C - BPNN in relative real results

From the analysis of Fig.4, it can be found that the 5C+T-BPNN method takes a certain amount of time in the early stage of training and the degree of normalization of the evaluation results is not high, but once the mathematical model is stable, it shows its advantages both in time consuming and the normalization of evaluation results.
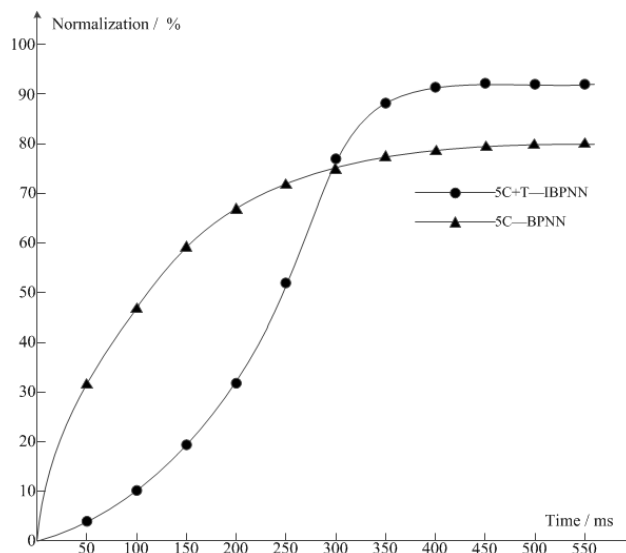
Fig.4 Comparison of normalization between 5C+T - IBPNN and 5C - BPNN

6. Conclusion

On the basis of the original collection of personal credit risk assessment samples, this paper innovatively introduces the concept of timestamp and integrates it into the 5C evaluation method, and then with the help of BPNN, the artificial intelligence analysis of personal credit risk assessment is realized, and the universality and accuracy of the algorithm are strengthened.

References

[1] Wang D. Coordination Mechanism of Seller's Consumption Credit with Service Level Affecting Credit Consumption Demand[J]. Value Engineering, 2017.

[2] Ouanphilalay S. The Impact of Household Credit on Consumption in Laos[J]. Journal of Southeast Asian Economies, 2017, 34.

[3] XIE Linquan, ZHAO Nan, XU Hao, et al. Analysis on the Development of Personal Credit Risk Assessment Model[J]. Henan Science and Technology, 2018.

[4] Fitria N, Sari R L. ANALISIS KEBIJAKAN PEMBERIAN KREDIT DAN PENGARUH NON PERFORMING LOAN TERHADAP LOAN TO DEPOSIT RATIO PADA PT. BANK RAKYAT INDONESIA (PERSERO), TBK CABANG RANTAU, ACEH TAMIANG. (PERIODE 2007-2011)[J]. Ekonomi Dan Keuangan, 2012.

[5] Li M, Gvozdenovic S, David R, et al. A real-time implementation of precise timestamp-free network synchronization[C]// Signals, Systems and Computers, 2015, Asilomar Conference on. IEEE, 2016:1214-1218.

[6] Zhou Z, Chen L, Niu X. Research on "Three Dimensional Credit Theory"-Based Credit Evaluation Indicator System of Small and Micro Enterprises[J]. Credit Reference, 2017.

[7] Osoro J. Credit information sharing -the potential impact of credit information sharing on the level of personal unsecured lending among commercial banks in Kenya[J]. Pediatric Pulmonology, 2011, 27(11):1193-1196.

[8] Wang X, Liu S, Tian W. Improved backtracking search optimization algorithm with new effective mutation scale factor and greedy crossover strategy[J]. Journal of Computer Applications, 2014.

[9] Aalto J, Lampinen J. A mutation and crossover adaptation mechanism for differential evolution algorithm[C]// Evolutionary Computation. IEEE, 2014:451-458.

[10] Liang Y, University Y N. Design of association mining model based on improved BP neural network[J]. Modern Electronics Technique, 2018.