Credit risk assessment based on Improved BP neural network based on L-M method

Yiyang Chen

Dept. of Economic Management, North China Electric Power Univ., Hebei 071003, PR China

Keywords: BP neural network; Levenberg Marquardt; Monte Carlo

Abstract: After the 21st century, China's small and micro enterprises have sprung up like mushrooms. Due to their small size, they often do not have fixed assets to mortgage loans, so they can only apply for credit loans with high risks for banks. To solve this problem, this paper proposes a set of small and micro enterprise credit risk assessment model based on Levenberg Marquardt improved BP neural network and genetic algorithm. Through the analysis of cash flow of enterprises, the paper extracts 7 indexes of three categories, including profit margin, positive profit growth proportion, negative cash flow proportion, number of main suppliers, number of main customers, void rate and return rate. The "excellent" standard of each index is determined according to the specific index value of 425 enterprises. We use the number of indicators that meet the "excellent" standard to measure the credit risk of enterprises in eight grades. The virtual data of 1000 enterprises are simulated by Monte Carlo random method, and the rough grades are used as training samples of BP neural network. LM method is used to improve the training efficiency, and the final fitting result is good.

1. Introduction

China’s steady economic development and unique and long-term vitality are due to its adherence to the economic policy of taking public ownership as the main body and diversified ownership developing together. In recent years, a special kind of enterprise, small and micro enterprises, has sprung up like bamboo shoots after a spring rain. As the name suggests, small and micro enterprises, in contrast to well-known large enterprises, have a small number of employees, low assets and low taxes. However, the stars in the night sky can bloom everywhere. Small and micro enterprises can not only expand employment and improve people’s livelihood, but also play an important role in national taxation and market economy.

However, even though the state has promulgated the support policies for small and micro enterprises, they still face difficulties such as high taxes and lack of funds. In terms of financing scale and financing times, bank loans are the main financing channels. And the proportion of bank loans is as high as 45%.

Bank credit is indeed a low-cost and stable financing method, which is chosen by most enterprises. However, as a bank, due to the lack of collateral, it is necessary to establish a set of credit strategy for small and micro enterprises, evaluate the business status and credit of enterprises, and determine the amount and interest rate of loans (too high interest rate will lead to the loss of a number of potential customers).

2. Problem analysis

In order to quantitatively analyze the credit risk of 123 enterprises, it is necessary to establish a certain index evaluation model to score the credit risk of 123 enterprises, so as to establish the optimal allocation of loan amount and annual interest rate for different enterprises when the total amount of credit is fixed.

Through the literature review of the credit strategy of the bank securities industry, we can understand the evaluation methods of the industry for small and micro enterprises, and extract several indicators that can affect the credit risk of enterprises. Using VFP program to deal with the

business data of 123 enterprises in batch, we can get the specific index value and form a 123 line matrix.

On the other hand, since the risk assessment results of these enterprises are not given in the question stem, in order to quantitatively describe the credit risk of these enterprises, the super average level of each index is taken as the standard of "low-risk and high-quality enterprise". Every enterprise reaches one standard, the higher the score, the lower the credit risk of the enterprise. Therefore, each group of indicators can correspond to an integral value. Even though the integral comes from a large number of statistical data, in order to increase the objectivity of the scoring and reduce the subjective factors of this "artificial assignment method", 1000 samples are simulated by Monte Carlo random simulation method, and the corresponding integral value is used as the training data of BP neural network. The trained BP neural network can score the risk of each group of input indicators, thus 123The credit risk of enterprises is analyzed quantitatively.

3 Preparation work

3.1 Data preprocessing

When dealing with the data of 123 enterprises, it is different from the general excel operation. Because of the huge amount of data, although the processing content is mostly simple operation, if the traditional manual processing method is adopted, it is too tedious and boring, and it is impossible to complete the processing of nearly one million data in a short time. Therefore, we use VFP language, through the preparation of macro program to achieve fast, large quantities of processing.

Compared with the database software, writing macro program directly in Excel runs slowly, but it saves the tedious import and export process. The running results directly generate more than 1000 data tables in Excel, with clear structure and more convenient call.

The data processing results are as follows: (only the top 10 enterprises are shown)

Table 1 Excel processing results

	profit margin	Proportion of profit margin increase	Capital flow early warning duration ratio	Number of major suppliers	Number of major customers	Void rate	Chargeback rate
E1	-0.29216	0.69402784	0.938661877	1	6	0.037336	0.024172
E2	2.854413	0.265939675	0	1	0	0.039297	0.011502
E3	11.19952	0.688183924	3.56443E-05	3	3	0.020221	0.158099
E4	7.430778	0.794301327	0.00039032	1	3	0.066308	0.007168
E5	0.025896	0.326642927	0.368080285	5	6	0.043357	0.008671
E6	0.234469	0.082544555	0.00102328	1	2	0.050206	0.008314
E7	8.39025	0.318355502	0.001596131	4	3	0.026853	0.110986
E8	1.23416	0.271767985	0.000199097	3	0	0.057101	0.0102
E9	13.02555	0.568631389	0.001405764	6	2	0.023532	0.014609
E10	63.83278	0.124771481	0	1	7	0.061951	0.007717
...							

3.2 Determination of "excellent" standard

In the subsequent modeling, due to the lack of quantitative evaluation value of risk, it is necessary to set a "excellent" standard for each index. The maximum, minimum, average and variance of each index are calculated as follows:

Table 2 Excel processing results

	profit margin	Proportion of profit margin increase	Capital flow early warning duration ratio	Number of major suppliers	Number of major customers	Void rate	Chargeback rate
Min	-0.99537	0.013261	0	0	0	0	0
Max	8341.582	0.99424	0.99973	9	22	0.754933	0.729572
Ave	128.2259	0.472108	0.151686	3.701176	3.183529	0.072893	0.015743
Var	470167.4	0.071009	0.089267	3.644822	5.20867	0.005034	0.001892
standard	>200	>0.6	<0.1	>5	>5	<0.05	<0.01

Generally, the standard is slightly better than the average value. For items with large variance, the standard is far away from the average value; for the item with small variance, it is closer to the mean value.

4. Modeling

The evaluation model is mainly composed of a set of BP neural network improved by elastic gradient descent method. BP neural network is a kind of multi input single output multi-layer negative feedback linear system, which can imitate the output of nonlinear function. The typical neural network is divided into input layer, hidden layer and output layer (as shown in the figure).When the input exceeds the threshold after a linear transformation, the next layer can be triggered for output. Through continuous training, the neural network can adjust and optimize its own weights and thresholds, so that it can be closer to the training samples.

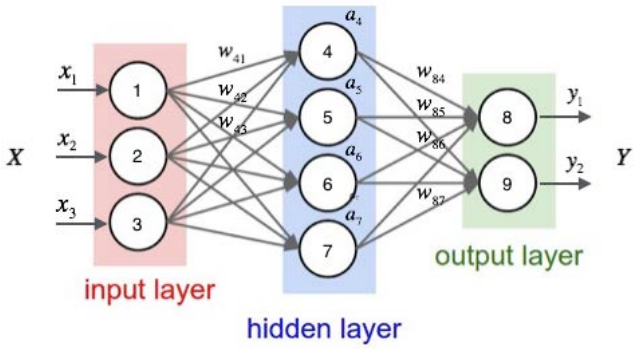

Figure 1 BP network

Because the number of enterprises given is small, and the credit risk of each enterprise is not given clearly, there is not enough training samples and necessary output conditions, so it does not meet the requirements of neural network training. Therefore, we first use Monte Carlo simulation method to simulate a batch of 7-dimensional index data, and then artificially assign risk score as the output condition according to these index data. On the other hand, the trained neural network can also weaken the "stiff sense" of artificial assignment, and get more natural and diverse risk assessment results.

For the risk score, we have already mentioned the process of screening and calculating the index parameters, and obtained three categories, seven indicators and corresponding "excellent" standard values. Through these indicators, we can judge whether the enterprise meets the standard of excellence. If so, add one, if not, you will get zero. After 7 judgments, a score of 0-7 points can be obtained. The probability of the final bankruptcy of small and micro enterprises is about 0.3, that is, for enterprises with high risk, 30% of them may not be able to repay the loan. Based on this, we give the corresponding risk scores of different scores as follows

Table 3

score	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Risk score	0.030	0.027	0.024	0.021	0.018	0.015	0.012	0.009

In this way, the standard output corresponding to a batch of simulated data can be determined. As a sample, the neural network can be trained well.The 425 groups of data obtained are put into the neural network to get the risk score of each enterprise.

5. Model solving

Through the literature review of the credit strategy of banking and securities industry, we know the evaluation methods of small and medium-sized enterprises in the industry, and extract seven indicators that can affect the credit risk of enterprises, namely profit margin$P$The time ratio of rising profitsσThe time ratio of capital flow early warningδ, number of major suppliersQ_s, number of major customersQ_c, invoice void rateη, return rateβ.

(1) Establishment of evaluation matrix

After getting seven different indexes, the VFP program is used to process the data of 123 enterprises in batch, and the specific index values are obtained to form a 123 × 7 evaluation matrix.

(2) Sample simulation and training

In order to quantify the credit risk of enterprises, we select the "excellent" criteria of seven indicators as follows.

Table 4 Standard values of indicators

index	Standard value
P	>200
σ	>0.6
δ	<0.1
Q_s	>5
Q_c	>5
η	<0.05
β	<0.01

After that, we compare the value of each index of 123 enterprises. At the same time, it is simulated by Monte Carlo random simulation method 1000 The risk score corresponding to the integral value is taken asBPTraining samples of neural network.use $Levenberg-Marquardt$As a training function, it replaces the traditional gradient descent method to obtain better optimization efficiency.

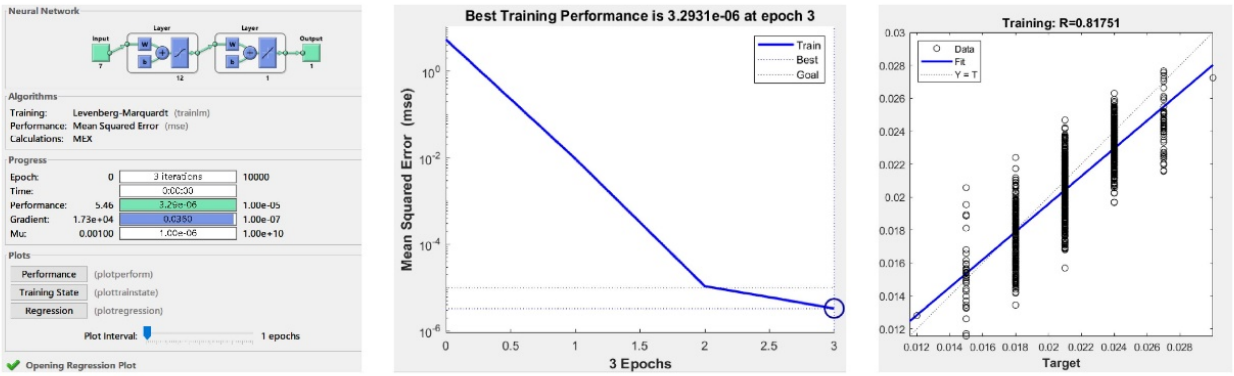

Figure 2 Training results of neural network

6. Conclusion

The $Levenberg-Marquardt$ As a training function instead of the traditional gradient descent method, the BP neural network is improved and the optimization efficiency is improved to a certain extent. The Monte Carlo stochastic process is used to simulate a large number of data, which improves the training effect of neural network. Therefore, we can conclude that the model can provide a set of feasible and scientific credit evaluation for banks.

References

[1] Arat Kerim T.,Hagen Cornelis W.. Model sensitivity analysis of Monte-Carlo based SEM simulations [J]. Results in Physics, 2020, 19.

[2] Haiyan Wang,Jinyan Fan. Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound [J]. Optimization Methods and Software, 2020, 35(4).

[3] Fang Yunlong. Research on bank credit risk management under the new situation [J]. China collective economy, 2020 (30): 100-101