Risk Quantification of Small and Medium-Sized Enterprises and Bank Optimal Credit Strategy Model

Hu Hanqing, Du Jian, Bi Gaoang

School of Economics, Central University of Finance and Economics, Beijing, 102206

Keywords: bank credit decision model, risk quantification, BP neural network, nonlinear programming, game theory

Abstract: Based on principal component analysis and BP neural network model, this paper constructs a quantitative model of enterprise loan risk, and solves the optimal loan strategy of different types of enterprises according to the idea of game theory and nonlinear programming model. Given the credit rating and default data of an enterprise, the number of upstream enterprises, the total input amount, the total input tax, the number of downstream enterprises, the total output amount, the total output tax, and the output void rate are obtained through data cleaning and mining. Combined with the data of rating and default, we get the average score of each enterprise by principal component analysis. In the case of no given enterprise credit rating and default data, according to the possible internal relationship between enterprise indicators and credit rating, a BP neural network model is established. Considering the impact of emergencies on enterprise risk and bank credit strategy, this paper uses the method of text analysis to classify enterprises into corresponding industries. The framework of each industry risk model is established by using Delphi method, and the industry risk value of corresponding enterprises is modified, and the optimal loan amount and the optimal interest rate under the influence of industry risk are obtained.

1. Introduction

In today's economic operation, the development of the real economy often needs the financial system and other virtual economy as the hub to reasonably allocate the resources such as funds and preferential policies to the enterprises in need, so that the social resources can be allocated with maximum efficiency. In the market, many industry giants will borrow money from banks to make up for their cash flow. The debt ratio of some large companies is even more than 50%. This shows that the loan funds provided by banks are the key factor for enterprises to continue to develop. Compared with the large enterprises with strong strength and rich assets, many small and medium-sized enterprises also need the financial support. However, due to their own lack of strength, they can not make loans based on the value of collateral provided by small and medium-sized enterprises. They need to pay attention to the strong strength and stable supply-demand relationship according to the credit policy, the transaction note information of enterprises and the influence of upstream and downstream enterprises. Enterprises with high reputation and small credit risk can be given preferential interest rate policies.

2. Bank Credit Decision

Principal component analysis is an excellent dimensionality reduction algorithm based on data. It uses linear transformation to convert multiple evaluation indicators into a small number of "principal component" indicators, and can calculate the score matrix by the contribution rate of each principal component (i.e., the characteristic value of the correlation coefficient matrix), so as to obtain a unified multivariable measurement index.

According to the seven indicators extracted from data mining, the rating and default indicators of each enterprise, principal component analysis is carried out on each index value of each quarter of each enterprise. The corresponding eigenvalue sequence and load matrix are shown in the table

Table 1: Sequence of eigenvalues

component	characteristic value	Total contribution rate	Is it a principal component
1	2.17144719	24.127%	yes
2	1.92009274	45.462%	yes
3	1.70660488	64.42%	yes
4	1.15506542	77.26%	yes
5	0.99670328	88.33%	no
6	0.5275005	94.19%	no
7	0.29345736	97.45%	no
8	0.1472835	99.09%	no
9	0.08184512	100%	no

Table 2: Load matrix

index	Ingredient 1	Ingredient 2	Ingredient 3	Ingredient 4
1	0.4919	-0.1257	-0.4118	0.5446
2	0.1217	0.9697	0.0029	0.0614
3	0.1374	0.9664	-0.0005	0.0765
4	0.3713	-0.1096	-0.5038	0.5852
5	0.6300	-0.0940	0.7122	0.0987
6	0.6011	-0.0846	0.7255	0.1598
7	0.0497	-0.0152	0.0188	0.1418
8	0.7112	-0.0379	-0.3865	-0.4379
9	-0.7010	0.0148	0.3159	0.5092

Figures 1 and 2 show the distribution of the first and second principal component scores among the observed values

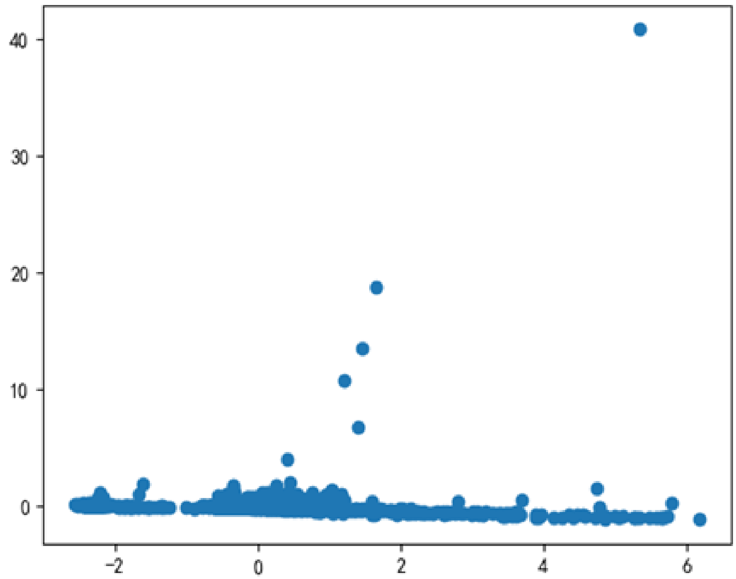

Figure 1: Scatter plot of principal component one and two

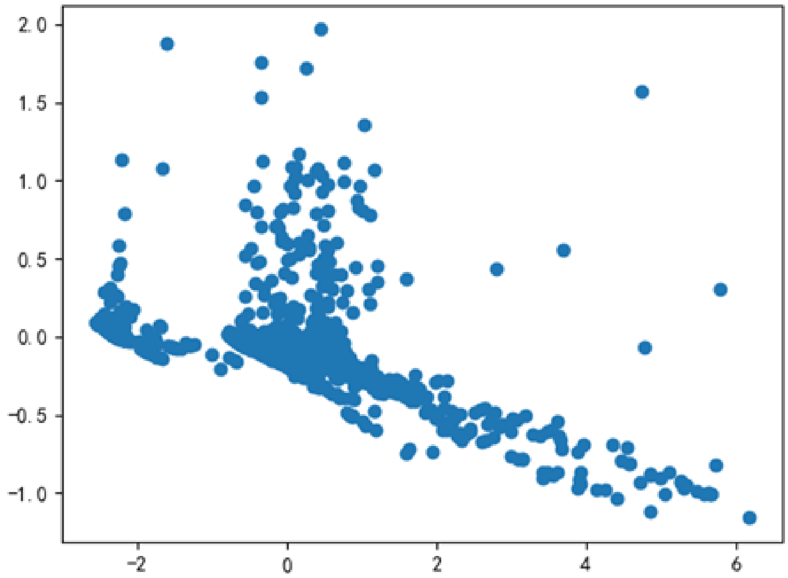

Figure 2: Scatter distribution with outliers removed

In the above score chart, we observe that the removal of outliers has a clearer understanding of the rest of our data distribution display. Therefore, in the following classified loans, we temporarily remove outliers and only classify them in the remaining and evenly distributed enterprises, so as to achieve the goal of evenly distributing scores and risk probability.

3. BP Neural Network Model Prediction

BP neural network is mainly divided into configuration stage, training stage and prediction stage. In the configuration stage, we need to select the appropriate index input and select the appropriate number of hidden layer nodes. In the training stage, we need to constantly debug to minimize the error function; in the prediction stage, we will use the trained neural network to predict the enterprises. After data cleaning, we construct the number of upstream enterprises and other indicators as the input vector of the network. The basic information is as follows:

Table 3: Descriptive statistics of input vector indicators

variable	Number of samples	mean value	standard deviation	minimum value	Maximum
Number of upstream enterprises	1313.00	44.00	93.33	0.00	878.00
Total input amount	1313.00	14550.57	120000.00	0.00	3100000.00
Total input tax	1313.00	1878.34	16307.66	0.00	530000.00
Number of downstream enterprises	1313.00	34.19	87.15	0.00	875.00
Total amount of sales	1313.00	64732.14	160000.00	0.00	1000000.00
Total output tax	1313.00	6139.75	19107.37	0.00	170000.00
Cancellation rate	1313.00	0.09	0.14	0.00	1.0

The neural network is trained based on the structure of BP neural network. Specifically, using MATLAB neural network toolbox, based on Levenberg Marquardt method, after 13 times of training, we finally completed our training goal.

The performance chart and training state of the training process are shown in the figure. We can see that the error between the prediction result and the actual result of the neural network is gradually reduced, and the set goal is achieved.

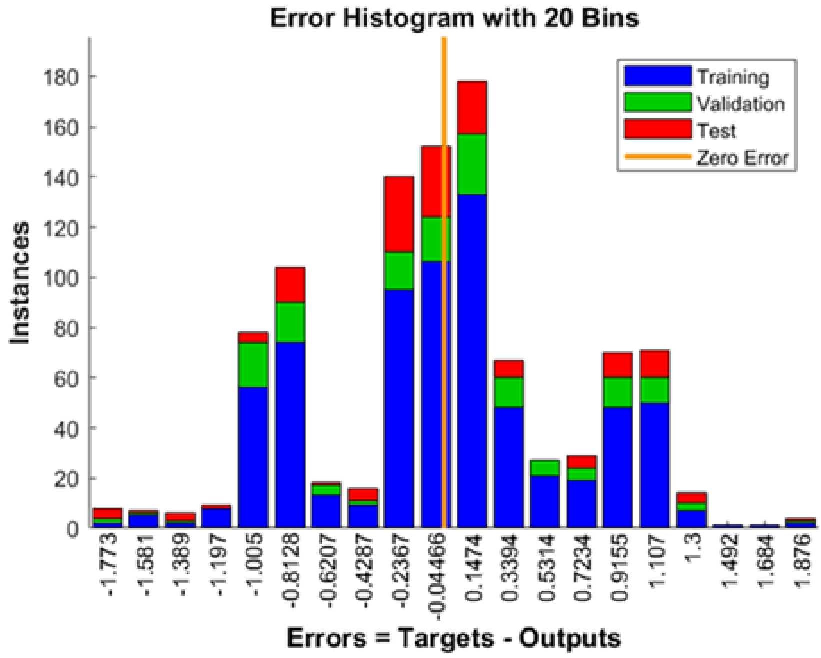

Figure 3: Estimation error histogram

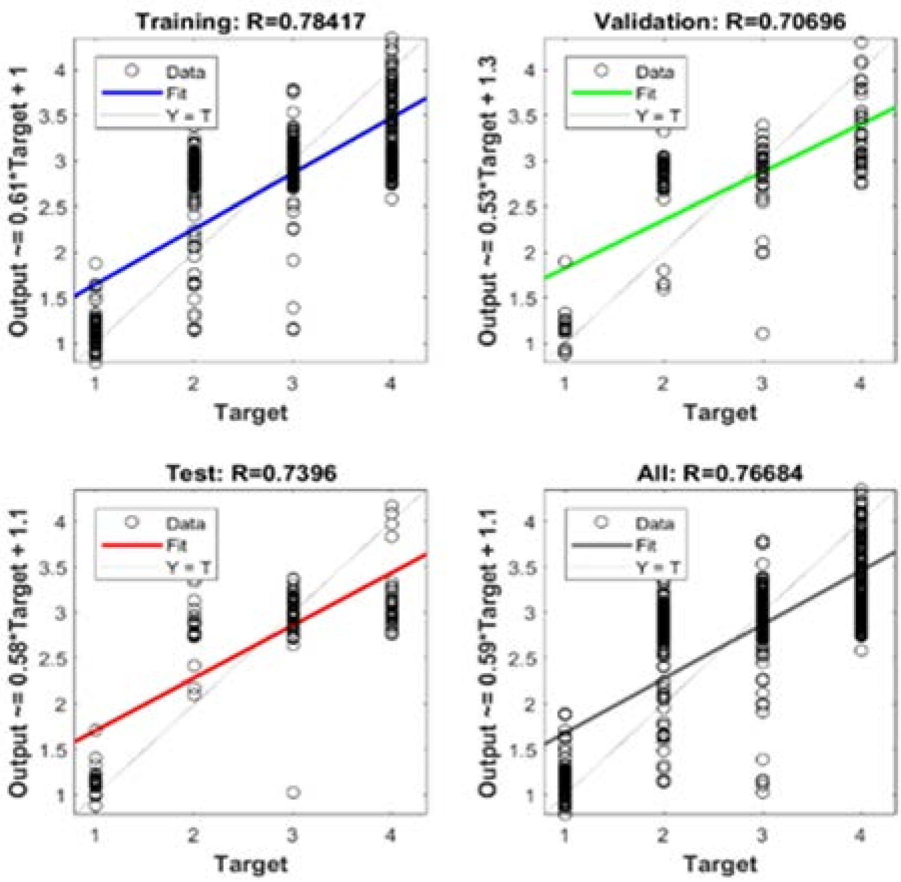

Figure 4: Degree of fit

Finally, we use the training set output results compared with the actual results, the final matching degree is 79.572%, which is basically able to meet the requirements.

After cleaning the data, we use the trained network structure to forecast and analyze the enterprise.

The basic situation of the prediction results is shown in Table 4

Table 4: neural network prediction score table

variable	sample	mean value	standard deviation	minimum value	median	Maximum
Neural network prediction score	3257	-0.14	1.53	-7.58	-0.37	11.39

4. Promotion and Application

Aiming at how to quantify the loan risk of small and medium-sized enterprises and how to formulate the optimal loan strategy, the established model first solves the problem that enterprises do not have credit rating, and then takes into account the exogenous factors that impact the market. If we continue to expand the amount of data, we can calculate the model parameters more accurately and make the evaluation results closer to reality. We can also subdivide the industries and categories to which each enterprise belongs, and formulate different credit strategies for different industries and categories. We can also include the probability of business success after borrowing, and take into account the optimal decision-making of enterprises, so as to make the model more complete, and finally form a complete and comprehensive risk assessment model for small and medium-sized enterprises and bank credit decision-making model.

References

[1] Jun Li, Xibo Xu, Jianshu Lv, Quan yuanWu, Mengyan Ren,Jianfei Cao,Peiyuan Liu. Source apportionment and health risk quantification for heavy metal sources in soils near aluminum-plastic manufacturing facilities in northeast China[J]. Human and Ecological Risk Assessment: An International Journal,2020,26(8).

[2] Samuel Stevens, Lesley Gibson, David Rush. Conceptualising a GIS-based risk quantification framework for fire spread in informal settlements: A Cape Town case study[J]. International Journal of Disaster Risk Reduction,2020,50.

[3] Esther B. Del Brio,Andrés Mora-Valencia,Javier Perote. Risk quantification for commodity ETFs: Backtesting value-at-risk and expected shortfall[J]. International Review of Financial Analysis, 2020, 70.