Based on Factor Analysis and TOPSIS Weighted Bank Credit Decision Model

Zheng Zihao[1], Liu Ye[1], Zhao Tao[2]

[1]School of Electrical Engineering, Beijing Jiaotong University, Beijing, 100044

[2]School of Computer and Information Technology, Beijing Jiaotong University, Beijing, 100044

Abstract: The bank is a financial institution with profit as its ultimate goal. On the one hand, its credit decision-making process can reduce the bank's lending risk and ensure the bank's income stability. On the other hand, it can urge enterprises to standardize their operation, strengthen their own credit construction, and form a good social competition environment, which has important practical significance. For the enterprises with credit rating, this paper first establishes the credit risk evaluation index system to evaluate the enterprise credit risk from the enterprise strength, reputation, management ability and other aspects. Because of the large number of indicators, the factor analysis method is used to reduce the dimension of the indexes. In order to verify whether the enterprise credit rating can be evaluated according to the TOPSIS score, this paper establishes a RBF neural network model. At the same time, a linear regression model is established to fit the annual interest rate of bank loans and customer churn rate. In order to study the impact of sudden factors on different industries, the data before the impact of sudden factors on the industry is used for function fitting through the cfool toolbox of MATLAB.

1. Introduction

Due to the relatively small scale of SMEs and the lack of mortgage assets, SMEs often need to borrow from banks to make the transaction smooth. Banks usually provide loans to enterprises with strong strength and stable supply-demand relationship according to the credit policy, transaction note information of enterprises and the influence of upstream and downstream enterprises, and can give preferential interest rate to enterprises with high reputation and small credit risk. The bank first evaluates the credit risk of small and medium-sized enterprises according to their strength and reputation, and then determines whether to make loans and credit strategies such as loan amount, interest rate and term according to credit risk and other factors.

2. Two factor analysis to reduce dimension of index two factor analysis to reduce dimension of index

For enterprises with credit records, there are 123 samples and 15 indicators. The following factor model can be established:

$$\begin{cases} z_1 = h_1 + a_{11}fact_1 + a_{12}fact_2 + \cdots a_{1m}fact_m + \sigma_1 \\ z_2 = h_2 + a_{21}fact_1 + a_{22}fact_2 + \cdots a_{2m}fact_m + \sigma_2 \\ \quad\vdots \\ z_{123} = h_{27} + a_{271}fact_1 + a_{272}fact_2 + \cdots a_{27m}fact_m + \sigma_{123} \end{cases} \quad (1)$$

The above equations are written in matrix form as follows:

$$z = h + Afact + \sigma \quad (2)$$

3. Determining weight vector of virtual index by entropy weight method

The index weight determined by AHP is very subjective, so we choose entropy weight method to weight the six virtual weights obtained above, and the positive matrix is as follows:

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{16} \\ f_{21} & f_{22} & \cdots & f_{26} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1231} & f_{1232} & \cdots & f_{1236} \end{bmatrix} \tag{3}$$

3.1 Normalization of positive matrix

Normalize the positive matrix, and mark the standardized matrix as to judge whether there is a negative number in the matrix. If there is a negative number, it needs to be standardized once to ensure that every element in the matrix is non negative. The formula is as follows:

$$\tilde{t}_{ik} = \frac{f_{ik} - \min\{f_{1k}, f_{2k}, \cdots f_{123k}\}}{\max\{f_{1k}, f_{2k}, \cdots f_{123k}\} - \min\{f_{1k}, f_{2k}, \cdots f_{123k}\}} \tag{4}$$

3.2 Calculate the entropy weight of each index

For the kth index, its information utility value is equal to:

$$\tau_k = 1 - \frac{1}{\ln 6} \sum_{i=1}^{n} p_{ik} \ln(p_{ik})(k = 1, 2, \cdots 6) \tag{5}$$

The entropy weight of each index can be obtained by normalizing the information utility value

$$\omega_k = \tau_k \bigg/ \sum_{k=1}^{6} \tau_k \ (k = 1, 2, \cdots, 6) \tag{6}$$

Using MATLAB to calculate the weight vector of the virtual index

$$\vec{\omega} = (0.066, 0.007, 0.132, 0.006, 0.007, 0.782)$$

It is the score matrix of 123 enterprises with credit records calculated by TOPSIS model optimized by entropy weight method, and the amount of bank loans is allocated proportionally according to the scores of each enterprise.

4. Evaluation of reputation grade by RBF neural network model

RBF (radial basis function) neural network model is a three-layer forward network, which is composed of input layer, hidden layer and output layer. The neural cells in the hidden layer are based on the nonlinear radial function, and the vector input is not connected according to the weight, but directly transmitted to the corresponding space. The output of the hidden layer neural cells is weighted sum to obtain the linear output of the neural network. The essence of RBF neural network is to transform the input data from one space to another. Through spatial transformation, the input data is transformed from non-linear to linear, and the complex problems become easy to solve .

For the layer (hidden layer), the output value of the node is:

$$o_\gamma = g(score_i) \tag{7}$$

Where G is the activation function of hidden layer node. Here we select soft max function, that is:

$$g(score_i) = \frac{\exp(score_i)}{\sum_{i=1}^{123} \exp(score_i)} \tag{8}$$

For the layer (output layer), the output value of the node is:

$$o_\rho = h(net_\theta) \tag{9}$$

H is the activation function of the output layer node.

The error between the predicted credit rating and the actual credit rating calculated by SPSS is shown in Table 1: the percentage of incorrect prediction is 0, that is, the predicted result is basically consistent with the actual result. Therefore, the TOPSIS score model constructed above can be used for corporate reputation rating.

Table 1 Surface Prediction error of RBF neural network

train	Sum of squares error	1.194E-27
	Percentage of incorrect predictions	0%
	Bayesian information criterion (BIC)	.a
	training time	0:00:00.31

5. Relationship between annual interest rate of bank loans and customer churn rate

Both commercial banks and enterprises are profit-making institutions. The determination of the annual interest rate of bank loans is one of the key factors for banks to reduce credit risk and obtain satisfactory returns. Establishing the functional relationship between the annual interest rate of bank loans and the rate of customer churn is of great reference significance to the formulation of Bank credit strategy. According to the statistical data of the relationship between customer churn rate and annual interest rate of bank loans in 2019, the scatter diagram is shown in Figure 1
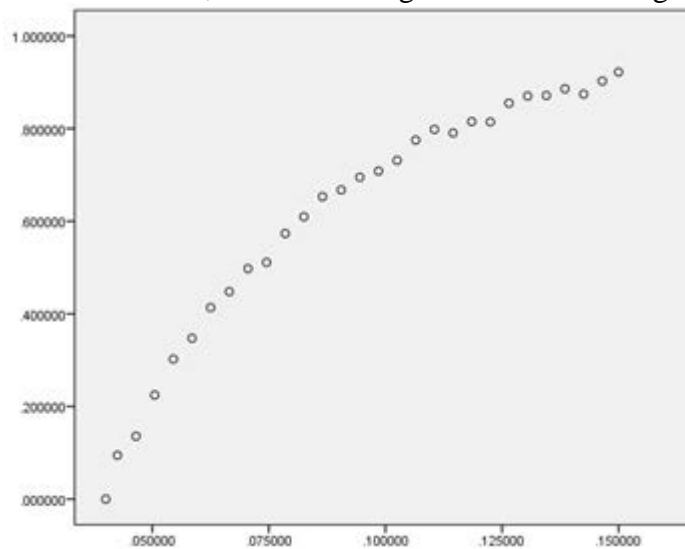

Figure 1 Scatter chart of the relationship between customer churn rate and annual interest rate of bank loans

It can be seen from the scatter chart that there is an obvious linear relationship between the customer churn rate and the annual interest rate of bank loans, so we establish a linear regression model to analyze it.

The annual interest rate of the bank loan of the third enterprise with credit record is α_i .The corresponding customer churn rate is f_i .The relationship between the customer churn rate and the annual interest rate of the loan meets the following requirements:

$$f_i = \lambda + \mu\alpha_i + \varepsilon_i \tag{10}$$

When

$$\sum_{i=1}^{123}(f_i - \hat{f}_i)^2 = \sum_{i=1}^{123}(f_i - \hat{\lambda} - \hat{\mu}\alpha_i)^2 = \min \sum_{i=1}^{123}(f_i - \lambda - \mu\alpha_i)^2 \qquad (11)$$

The residual error is the minimum.

Here are: $\hat{\lambda}, \hat{\mu} = \arg\min_{\lambda,\mu}(\sum_{i=1}^{123}(f_i - \hat{f}_i)^2) = \arg\min_{\lambda,\mu}(\sum_{i=1}^{123}(f_i - \hat{\lambda} - \hat{\mu}\alpha_{ii})^2)$

$$\hat{\lambda}, \hat{\mu} = \arg\min_{\lambda,\mu}(\sum_{i=1}^{123}(\hat{\mu}_i)^2) \qquad (12)$$

The data were analyzed by SPSS and the following conclusions were obtained

$$f = \begin{cases} f_A = -0.098 + 7.524\alpha \\ f_B = -0.118 + 7.351\alpha \\ f_C = -0.138 + 7.468\alpha \\ f_D = 1 \end{cases} \qquad (13)$$

among $f_k(\alpha)$ It is the functional relationship between the customer churn rate of enterprises with credit rating (k = a, B, c) and the annual interest rate of bank loans.

6. Optimization model for solving annual loan interest rate

6.1 Establishment of optimization model

The annual interest rate of bank loan is closely related to the bank & apos; s income this year. From the perspective of the bank, it is necessary to ensure that the bank can obtain as much revenue as possible when determining the annual interest rate of bank loan. Therefore, the problem of solving the annual interest rate of loan can be transformed into the optimization problem of maximizing the bank & apos; The decision variable is the annual interest rate of the third enterprise with credit record. If the annual income function of the bank is, then the objective function is: IQ

$$\max Q = [1 - f(\alpha)] \cdot score' \cdot \alpha \cdot N \qquad (14)$$

Where is the function of customer churn rate on the loan annual interest rate, is the normalized enterprise score, and is the total credit. Because it is a constant, it does not affect the solution of the objective function, so that = 1 can simplify the objective function. The constraints to be satisfied are the range of annual interest rate of bank loans and the normalization of scores.

$$\begin{cases} 4\% \le \alpha_i \le 15\% \\ \sum_{i=1}^{123} score' = 1 \end{cases} \qquad (15)$$

6.2 Creative transformation of solution

Considering that different regression equation coefficients are needed when the credit rating of enterprises is different, which hinders lingo & apos; solution, we transform the piecewise solving problem into a matrix operation solving problem.

If a coefficient matrix of 4×1 is constructed, then $M^T = [7.524, 7.531, 7.468, 0]$

$$f = L \cdot M \qquad (16)$$

Matlab is used to calculate the matrix, and lingo is used to optimize the solution to get the annual loan interest rate of each enterprise when the bank & apos; annual income is maximized. The enterprise with the annual interest rate of 15% does not lend by default.

7. Conclusion

Less than zero in industry, retail, catering, logistics and transportation, medical and health industry, and humanities industry, indicating that the new epidemic had a negative impact on these industries. During the epidemic period, in order to effectively control the spread of the virus, most enterprises shut down production except for medical devices, pharmaceuticals, mask production and necessary logistics transportation. The frequency of people dining out suddenly dropped and the catering industry suffered huge losses. Most of the materials are transported by means of transportation and transportation routes designated by the state. The expenses for medical treatment and isolation of patients with fever are subsidized by the state, and the implementation of home isolation measures reduces the spread of other influenza. Therefore, the logistics and transportation industry, medical and health industry, etc., also have a certain "negative growth". However, the promotion of cloud office, online doctor consultation and online courses has promoted the development of high-tech industry to a certain extent, and the total industrial output value has a greater growth rate.

References

[1] Wei Xiao. Research on credit decision of commercial banks for SMEs [D]. Shandong University of Finance and economics, 2016

[2] MBA think tank encyclopedia, value added tax (VAT), https://wiki.mbalib.com/wiki/%E5%A2%9E%E5%80%BC%E7%A8%8E,2020.

[3] Di Zhang ying. Research and application of SME bank loan decision model [D]. Shanghai Jiao tong University, 2014