

The Credit Risk Assessment Model of Internet Supply Chain

Finance: Bayesian Networks Model Based on Relative Weight

Baoyu Zhong^{}, Yueliang Su*

College of Business Administration, South China University of Technology, China

^{*}Corresponding author: Baoyu Zhong, 490943753@qq.com

Abstract

In this study, we first design a credit risk evaluation index of Internet Supply Chain Finance from the perspective of electronic commerce enterprises. Then, we determine the relative weights of variables in logistic regression. After that, we construct a credit risk assessment model based on the relative weights and the basic thought of Bayesian Networks. Finally, we perform a comparative analysis of a specific case using logistic regression model and Bayesian Networks based on relative weights. The findings of this study reveal that Bayesian Networks based on relative weights may have a higher accuracy in assessing the credit risk level of small and medium sized enterprises in Internet Supply Chain Finance.

Key words: *credit risk assessment model; relative weights; Bayesian Networks*

1. Introduction

On the background of Internet and the participation of electronic commerce enterprise, supply chain financing model has changed with opportunities and challenges. First, as a network platform for enterprises to trade, the addition of electronic commerce enterprise changes the traditional supply chain finance financing methods. And then, with the rise of the loan companies, online lending becoming one of the main sources of funds for individuals and small businesses. And the comprehensive strength of the electronic commerce enterprise also began to imitate the pattern of network lending. In the Internet supply chain finance based on e-commerce platform, electronic commerce enterprises try to provide financing and loan services legally through the establishment of loan companies and online banks and so on. Therefore, the participation of electronic commerce enterprise in the supply chain finance not only changes the traditional supply chain finance financing methods, but also changes the status quo of traditional supply chain finance that the source of funds is single and the bank occupies the absolute dominance.

As a trading platform, electronic commerce enterprises can get the trading behavior information and logistics information of the two sides of the transaction. With the comprehensive consideration of the financial situation of the financing enterprise and

online transaction data, the electronic commerce enterprise assesses the credit risk level of the financing enterprise in the supply chain.

2. Credit risk evaluation index of Internet Supply Chain Finance

Table 1 The credit risk evaluation index of Internet Supply Chain Finance

First-grade index	Second-grade index	third-grade index	Index description
Internal factors X19	Credit status of the financing enterprise X13	Registration information X1	The basic information of a financing enterprise when it is registered on the platform
		Financial disclosure quality X2	Financial statement audit and information disclosure
		Sales profit margin X3	Sales profit / sales revenue
	Financing project status X14	Purchasing data X4	Project related purchasing data
		Sales data X5	Project related sales data
		Returned merchandise data X6	Whether there is a record of returned merchandise
	Operation status online X15	Online operation level X7	The age of online operation of the financing enterprise, etc.
		Informatization level X8	Online business conditions
External factors X20	Comprehensive ability of Core enterprise X16	Return on net assets of core enterprises X9	Measure the profitability of the core enterprise
	The overall situation of the supply chain X17	Level of Information sharing X10	Information sharing and system compatibility of the supply chain
		Online trading life X11	The age of online financing
Macro environmental factors X18	Policy support X12	Policy support for the applicant's industry	

Table 1 gives the credit risk evaluation index of Internet Supply Chain Finance which comprehensively considered the financial situation of the financing enterprises and online transaction data from the perspective of electronic commerce enterprises.

3. Bayesian Networks model based on relative weights

This study tries to improve Bayesian Networks model based on relative weights in logistic regression^{1, 2}. Logistic Regression first is adopted in data analysis, then the relative weights of the variables in logistic regression model are evaluated. Next, quantitative and qualitative methods are adopted to determine the conditional

probabilities of the Bayesian Networks model. Finally, we build a credit risk assessment of Supply Chain Finance based on Bayesian.³

Given a logistic regression function

$$\hat{y} = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)} \quad (1)$$

the process of determining the relative weights of the variable in this function is followed⁴:

1. Do logit transform to the equation (1), a linear regression function can be obtained:

$$Y = \text{logit}(\hat{y}) = \ln \frac{\hat{y}}{1 - \hat{y}} = b_0 + b_1 x_1 + \dots + b_n x_n \quad (2)$$

2. An estimate of the proportional contribution each x_j makes in predicting $\text{logit}(\hat{y})$ can be

$$\varepsilon_{OLS} = \Lambda^2 \beta^2 \quad (3)$$

where $\Lambda = (Z'Z)^{-1}Z'X$, $\beta = (Z'Z)^{-1}Z'Y$, $Z = PQ'$, $X = P\Delta Q'$, P is the matrix of eigenvectors associated with XX' , Q is the matrix of eigenvectors associated with $X'X$, and Δ is the diagonal matrix containing the singular values of X

3. Menard proposed a fully standardized logistic regression coefficient:

$$\beta_M = (b)(s_x)(R_o)/(S_{\text{logit}(\hat{y})}) \quad (4)$$

where b is the unstandardized logistic regression coefficient, s_x is the standard deviation of the predictor X, R_o is the square root of the OLS coefficient of determination for logistic regression, $(S_{\text{logit}(\hat{y})})$ is the standard deviation of the predicted value of the logit.

4. Finally, we can compute the relative weights in logistic regression:

$$\varepsilon_{log} = \Lambda^2 \beta_M^2 \quad (5)$$

4. Case analysis

In this paper, the data were collected by enterprises and staff on electronic commerce platform. Qualitative indexes took a 5 point scoring system while quantitative indexes were the historical data from the e-commerce platform. Finally, the number of valid samples was 83, all of them were used to determine the relevance and the association relationship of the logistic regression model and 13 of them were tested the accuracy of the model as validation samples.

4.1 Logistic regression

Given $X = (X_1, X_2, \dots, X_{12})$, twelve dimensional random variable representing twelve the third-level indexes in the Table 1, and Y_1, Y_2, \dots, Y_{83} , the dependent variable representing repayment of the 83 financing enterprises. The random variable Y_i can be 1 or 0, $Y_i=1$ means the financing enterprise i has a high credit risk level that it won't repay on time while $Y_i=0$ means the financing enterprise i has a low credit risk level. Logistic regression model can be obtained by using SPSS software

$$\ln \frac{\hat{y}}{1 - \hat{y}} = 169.732 - 27.956x_1 + 3.096x_2 + \dots - 17.409x_{12} \quad (6)$$

So the probability of small and medium financing enterprises repay on time can be obtained as follow

$$P = \frac{\exp(169.732 - 27.956x_1 + 3.096x_2 + \dots - 17.409x_{12})}{1 + \exp(169.732 - 27.956x_1 + 3.096x_2 + \dots - 17.409x_{12})} \quad (7)$$

From equation (3), we can obtain all probability of 83 financing enterprises repay on time. Let 0.5 be a boundary value, then those the probability is higher than 0.5 will be regarded as the high-risk enterprises. On the other hand, the financing enterprises whose probability is lower than 0.5 will be regarded as the low-risk enterprises. The following table is the result of the logistic regression model to verify the prediction of the samples.

Table 2 The observed and predicted value of 13 samples by logistic regression model

Validation sample	1	2	3	4	5	6	7	8	9	10	11	12	13
Observed value	0	0	1	0	0	0	0	1	0	0	0	0	0
predicted value	0	0	1	0	0	0	0	0	0	0	0	0	0

4.2 Bayes network model based on relative weights in logistic regression model

The structure learning process of Bayes network model is to select the variables that affect the credit risk level of small and medium enterprises in the supply chain finance and then analyze the correlation between variables. In this paper, we select all variables $(X_1, X_2, \dots, X_{19}, X_{20})$ in the credit evaluation index and the repayment (X_{21}) . We can construct a Bayesian Networks structure as Fig 1 shows,

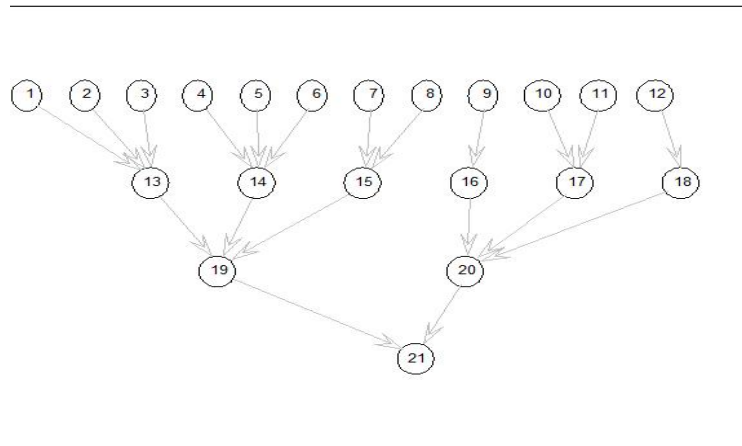


Fig 1 Bayesian Networks structure

By the section 3, we can obtain the relative weights of variables in logistic.

4.3 The Probabilistic reasoning of Bayes network model

Probability reasoning of credit risk assessment of financing enterprises based on Bayesian Networks is actually that given all nodes, calculate the probability of the repayment of one enterprise. Combined with the Bayesian network structure and the node conditional probabilities, the probability of $X_{21} = 1$ (the enterprise is regarded as a high-risk enterprise)

$$\begin{aligned}
 P(X_{21} = 1) &= P(X_{21} = 1|X_{19} = 1, X_{20} = 1)P(X_{19} = 1)P(X_{20} = 1) \\
 &+ P(X_{21} = 1|X_{19} = 1, X_{20} = 0)P(X_{19} = 1)P(X_{20} = 0) \\
 &+ P(X_{21} = 1|X_{19} = 0, X_{20} = 1)P(X_{19} = 0)P(X_{20} = 1) \\
 &+ P(X_{21} = 1|X_{19} = 0, X_{20} = 0)P(X_{19} = 0)P(X_{20} = 0) \quad (8)
 \end{aligned}$$

From equation (8), we can also obtain all probability of 83 financing enterprises repay on time. Let 0.2、0.4、0.6、0.8 be boundary values, then those the probability is smaller than 0.2 will be regarded as the low-risk enterprises with credit risk level 1. The enterprises whose probability is between 0.2 and 0.4 will be regarded as the low-risk enterprises with credit risk level 2. By that analogy, the enterprises whose probability is higher than 0.8 will be regarded as the high-risk enterprises with credit risk level 5. And the more the probability is close to 1, the higher the credit risk level of the financing enterprise will be. The following table is the result of the Bayesian Networks model to verify the prediction of the samples based on relative weights

Table 3 The observed and predicted value of 13 samples by Bayesian Networks model

Validation sample	1	2	3	4	5	6	7	8	9	10	11	12	13
Observed value	1	3	4	1	2	2	3	4	2	3	3	2	3
predicted value	1	3	4	1	2	2	3	3	2	3	3	2	3

Table 2 and Table 3 illustrates the result of the logistic regression model and Bayesian network model based on a special case. After comparison, some conclusions are obtained

1. From the result of the logistic regression model and Bayesian network model, the prediction accuracy of the two models is as high as 92.3%, which have a high accuracy.
2. In logistic regression model, the dependent variable was transformed into binary variables. Actually, the dependent variable was credit risk level of financing enterprises which was 5 point scoring system. As a result, under the premise of the same accuracy of the predictive value, the actual accuracy of the Bias Network is higher than that of the logistic regression model.
3. In the 13 validation samples, the predicted value of sample 8 was contrary to observed one according to logistic regression model. while it was different from observed one according to Bayesian Network model. In detail, the credit risk level predicted of sample 8 was 3 while the credit risk level observed was 4. The observed value differs with the predicted values in a risk level, which was in acceptable error range.

5. Conclusion

This study analyzed the credit risk level assessment model of Internet Supply Chain Finance from the perspective of electronic commerce enterprises. we construct a credit risk assessment model based on the relative weights and the basic thought of Bayesian Networks. Finally, we perform a comparative analysis of a specific case using logistic regression model and Bayesian Networks based on relative weights.

Through our analysis, a credit risk assessment model with higher accuracy is established. It is proved that the credit risk assessment model can be improved to a certain extent if the prior probabilities and conditional probabilities of the Bayesian Networks model are determining used the relative weights of variables.

Finally, there are some possible directions that can be studied deeply. First, we can study the credit risk assessment model based on all data electronic commerce enterprises owned. Then the case can be simulated by traditional Bayesian Networks model and compared the Bayesian Networks model on relative weights of variables.

Acknowledgements

This research was supported by “the Fundamental Research Funds for the Central Universities” (2015ZKYJZX02) and Guangzhou Financial Services Innovation and Risk Management Research Base.

References

1. *J.W. Johnson*, A heuristic method for estimating the relative weight of predictor variables in multiple regression, *J. Multivariate Behavioral Research*, 35(2000) 1- 19.
2. *S. Tonidandel, J.M. LeBreton*, Determining the relative importance of predictors in logistic regression:An extension of relative weights analysis, *J. Organizational Research Methods*, 13(2000)767- 781.
3. *J. H. Bierman, W.H. Hausman*, The Credit Granting Decision, *J. Mangement Science*,16(1970)519-532.
4. *R. Azen, D. V. Budescu*, Comparing predictors in multivariate regression models: An extension of dominance analysis, *J. Journal of Educational and Behavioral Statistics*, 31(2003)157-180.