ISSN:2790-1688 DOI: 10.56028/aetr.2.1.221

Credit Risk Evaluation of Supply Chain Finance Based on K-Means-SVM Model

Li Zhu[1, a], Wenhao Liu[1, b] , Rongdi Zhang[1, c] , Bingjie Dong[2, d]

[1] China University of Mining and Technology, Xuzhou 221000, China;

[2] Northwestern Polytechnical University, Xi'an 710018, China.

[a]Liz_yzcs4226@163.com, [b] lwh08180324@163.com, [c] 13203710@cumt.edu.cn,

[d] 2019302987dbj@mail.nwpu.edu.cn

Abstract. With the rapid development of supply chain finance, it is important to evaluate its credit risk effectively. The Support Vector Machine (SVM) is designed to construct the credit risk measurement model of supply chain finance. Considering the characteristics of SVM model, we select the clustering center based on K-Means clustering algorithm and the edge points far from the clustering center as training samples to train the SVM model. Experimental results show that compared with single SVM model, the overall classification accuracy of K-means-SVM model is increased by 7.2%, and the first type error rate is reduced by 5.0%, which verifies the superiority and effectiveness of k-means-SVM model applied to enterprise credit risk assessment under supply chain finance mode.

Keywords: Supply chain finance; Credit risk; K-means clustering algorithm; Support vector machine

1. Introduction

Small and medium-sized enterprises are an important pillar of China's economic development. However, due to the common problems of this type of enterprise, such as weak credit, less collateral and difficult capital turnover, it is hard for them to obtain financing from commercial banks. The emergence of supply chain finance provides a brand-new way to overcome the financing difficulties. Supply chain finance takes the whole supply chain as the object of investigation, and changes the traditional risk management mode, and turns the risk management for a single enterprise into the risk management for the whole supply chain. Take the related enterprises in the supply chain as a whole, conduct comprehensive credit granting for them, and provide comprehensive financing and loan services for the whole supply chain. By effectively injecting credit capital into the supply chain and strengthening the cooperative relationship between the core enterprises and the upstream and downstream enterprises of the supply chain, the win-win situation of all participants can be realized.

Due to information asymmetry, commercial banks cannot fully grasp the operation and profitability of smes, and there is great uncertainty in the financing process, which is easy to cause credit risks. At the same time, in addition to the conductivity of credit risk in the supply chain, the credit risk of a single enterprise is easy to infect other enterprises in the supply chain, so that the risk harm is multiplied and the stable operation of the supply chain is impacted. On July 9, 2019, THE CBRC issued the Guidelines of the General Office of the CBRC on Promoting Supply Chain Finance to Serve the Real Economy to major banks and insurance companies, insisting on the availability of transaction information to ensure direct access to first-hand original transaction information and data. We will continue to comprehensively manage risks. We will not only pay attention to the risks of core enterprises, but also test the risks of enterprises in the upstream and downstream chains. Therefore, how to effectively obtain transaction information and data, improve the level of credit risk evaluation, for the healthy development of supply chain finance is of great significance.

In order to improve the supply chain financial capacity, mode of small and medium-sized enterprise credit risk assessment in SVM model to construct financial supply chain mode of small and medium-sized enterprise credit risk assessment model, on the basis of factor analysis, using the

ISSN:2790-1688 DOI: 10.56028/aetr.2.1.221
K - means clustering algorithm to select clustering center and clustering center far distance of sample points as the training sample in order to train the SVM model. By comparing the evaluation effect with the single SVM model, the accuracy and effectiveness of the K-means-SVM model are verified, providing reference for promoting the healthy development of supply chain finance.

2. Literature review

There are few foreign researches on credit risk of supply chain finance, which mainly focus on risk management. Aberdeen (2006) [1] found that optimization of supply chain finance in technology, financing and visualization can improve the financial status of financing enterprises and better control credit risks in the financing process. Manuj (2008) [2] made an in-depth study of the accounts receivable mode of supply chain finance and believed that in the process of risk management, it is necessary to reduce the dependence on risk avoidance mechanism to achieve better results in the risk control of supply chain finance. Lustrato (2014) [3] constructed a model combining supply chain finance with physical supply chain, and found that supply chain integration is an important method to improve the risk management ability of supply finance. It can realize information visualization and reduce information asymmetry, which is beneficial to all participants. Sum (2016) [4] believes that the lack of appropriate risk analysis models in banks is the main reason for the risks in the supply chain finance model, and applying reasonable models to quantitatively analyze risks is an effective way to control supply chain finance risks. Finch (2016) [5], from the perspective of commercial banks, believes that it is necessary to optimize the information management system, improve the information management ability in the process of supply chain finance business, and thus improve the risk management level. Song et al. (2018) [6] believe that supply chain finance can reduce the information asymmetry of all participants and reduce the default risk of financing enterprises through the use of business closed loop, receivables transfer and relationship embedding.

In terms of credit risk assessment methods of supply chain finance, Xiong and Ma Jia et al. (2009) used principal component analysis and regression model to measure credit risk of enterprises through the evaluation method of subject plus debt rating, overcoming the limitation of relying on expert scoring method. Huang Sijing et al. (2014) combined the influencing factors of the supply chain and the macro environment of the enterprise, and concluded that the systemic risk of the whole supply chain should be attached importance through the analytic hierarchy process [8]. Liu Yanchun et al. (2017), through the establishment of structural equation model and grey correlation model, concluded that the situation of enterprises, industry risks and supply chain operation capacity are positively correlated with the credit risk of enterprises under the supply chain finance model, and believed that the construction of sme database should be strengthened. Xu Hongfeng et al. (2018) analyzed the credit risks of small and medium-sized enterprises by using analytic hierarchy Process and multi-level grey comprehensive evaluation method based on the analysis of the three modes of supply chain finance, and the results showed that the supply chain finance system is conducive to solving the financing difficulties of small and medium-sized enterprises. Yu Desheng et al. (2019) selected 14 credit risk evaluation indicators, screened out 4 important variables for risk prediction by Logistic model stepwise backward method, and proposed to improve the credit risk evaluation system and pay attention to risk warning signals.

With the continuous development of computer technology, many scholars apply machine learning algorithms to supply chain finance credit risk measurement. Hu Haiqing et al. (2011) combined the credit status and supply chain relationship of core enterprises and found that the SVM-based credit risk assessment model is more advantageous in supply chain financial risk assessment. Wu Ping et al. (2015) established a credit risk assessment system based on BP neural network and verified its effectiveness by summarizing the characteristics of online supply chain financial risk factors. Li Jian et al. (2019) took the automobile supply chain as a sample, used the random forest model and blind number theory to screen variables, and compared the prediction

ISSN:2790-1688 DOI: 10.56028/aetr.2.1.221
effects of various evaluation models, and found that the PSO-SVM model had high prediction accuracy in the evaluation of supply chain financial credit risk . Pan Yongming et al. (2020) used SVM model to build a classification and prediction model of supply chain financial credit risk, and introduced Information gain (IG) to screen characteristic variables, proving that IG-SVM model has higher prediction accuracy than single SVM model, and further improving the classification efficiency of SVM model.

The quality of training set samples has a direct impact on model classification results. In the past, when machine learning algorithm is applied to supply chain finance credit risk evaluation, the random method is used to select training samples, which failed to maximize the classification efficiency of machine learning algorithm. Zhao Zhigang et al. (2011) selected training samples based on clustering, which effectively improved the classification efficiency of SVM model. In view of this, by selecting a high-dimensional space computing power and strong generalization ability of SVM model to build enough supply chain financial prediction model of credit risk classification, using clustering effect can explain better and stronger K - means clustering algorithm training SVM model selection of the most effective training set samples, in order to improve the credit risk prediction ability of the SVM model.

3. Evaluation index

Taking financial data as the starting point, we conclude and summarize the evaluation indicators used in previous literatures. Based on the characteristics and principles of scientificity, rationality and applicability of supply chain finance credit risk, our study screened out 16 variables from 5 aspects, including profitability index, debt paying ability index, growth ability index and operation ability index.

Since 16 initial variables are selected and the variable dimension is high, both classical econometric algorithm and machine learning algorithm have high correlation and high dimension of model indicators, resulting in excessive model fitting and invalid parameter estimation. Therefore, factor analysis of variables is firstly carried out to extract variables with major analytical ability, and then empirical analysis is carried out using the obtained variables.

Table 1 Credit risk evaluation system

Level indicators	Secondary indicators
Profitability	ROE (X_1)
	Net interest rate on assets (X_2)
	Net margin on sales (X_3)
	Cost expense margin (X_4)
Debt-paying ability	Asset-liability ratio (X_5)
	Current ratio (X_6)
	Quick ratio (X_7)
	Cash flow liability ratio (X_8)
Growth ability	Earnings per Share Growth rate (X_9)
	Growth rate of Operating Income (X_{10})
	Net profit Growth rate (X_{11})
	Growth rate of Total Assets (X_{12})
Supply chain finance	Inventory turnover (X_{13})
	Accounts receivable turnover (X_{14})
Operation ability	Current Asset turnover (X_{15})
	Total Asset turnover (X_{16})

ISSN:2790-1688 DOI: 10.56028/aetr.2.1.221

4. Empirical analysis

4.1 Sample data selection and descriptive statistics

We select 210 listed small and medium-sized manufacturing enterprises in China, and the relevant financial data of these enterprises in 2018 are collected through RESSET database. Among them, 60 ST and *ST enterprises are considered as risky enterprises; There are 150 non-ST and non-* ST enterprises, which are regarded as non-default risk enterprises. SPSS23.0 is used to make descriptive statistics of relevant data.

Table 2 Descriptive Statistics

	N	Min	Max	Mean	Standard deviation
X_1	210	-709.801	410.010	-11.52318	80.248496
X_2	210	-114.628	29.790	-1.39232	15.486374
X_3	210	-5432.607	73.635	-45.85692	391.244934
X_4	210	-369.121	103.679	-7.37002	54.378930
X_5	210	9.417	175.835	48.27571	22.554532
X_6	210	0.103	19.624	1.83994	1.852268
X_7	210	0.102	17.592	1.44519	1.651549
X_8	210	-2.393	2.352	0.13246	0.389737
X_9	210	-7906.452	2500.000	-298.45983	1173.969443
X_{10}	210	-98.823	408.654	9.74252	43.326221
X_{11}	210	-9015.560	2549.921	-288.70384	1184.010872
X_{12}	210	-89.626	44.894	.07116	20.591038
X_{13}	210	0.016	1432.970	12.25263	98.646772
X_{14}	210	0.030	2390.679	22.75262	166.842129
X_{15}	210	0.006	9.987	1.29032	1.054248
X_{16}	210	0.006	5.797	0.67470	0.641891

4.2 Factor analysis

In statistical analysis, multicollinearity of independent variables should be avoided as much as possible so as to avoid the influence of strong multicollinearity on model reliability. SPSS23.0 is used to perform factor analysis on the variables, and the variables with major analytical ability are extracted.

Before factor analysis, it is necessary to test the correlation degree between variables and judge whether variables are suitable for factor analysis. Using KMO and Bartlett tests, the KMO value is 0.649 (KMO>0.6), and on the basis of the cumulative variance contribution rate of 75.5%, the first 6 variables are selected as the main components.

4.3 Empirical analysis based on K-means-SVM model

Based on the data of 105 enterprises in the training set, six main factors are used as input vectors for training, radial basis kernel function (RBF) is used to establish SVM prediction model, and the rBF-SVM model after training is used to test the samples in the test set. At the same time, in order to show the efficiency of the K-means-SVM model more intuitively, it is planned to compare the classification prediction results of the K-means-SVM model with the evaluation results of the single SVM model that randomly selects the training set samples and test set samples. The evaluation results are represented by the classification accuracy of training set and test set, the overall classification accuracy and the first category classification error rate. The specific results are shown in Table 3.

ISSN:2790-1688 DOI: 10.56028/aetr.2.1.221

Table 3 Comparison of classification efficiency

Classification model	Correct rate of training set classification	Test set classification accuracy rate	Overall classification accuracy rate	The first category error rate
K-means-SVM	94.3%	91.4%	92.9%	16.7%
SVM	86.4%	84.1%	85.7%	21.7%

It can be seen from Table 3 that the K-Means-SVM model is higher than the single SVM model in terms of the classification accuracy of training set, test set and overall classification accuracy. Among them, the overall classification accuracy of the K-Means-SVM model is 92.9%, 7.2% higher than that of the single SVM model. The first classification error rate represents the model's ability to identify defaulting enterprises and is often used to test the evaluation effect of credit risk models in statistics. The K-Means-SVM model has a category 1 classification error rate of 16.7%, which is 5.0% lower than the single SVM model, proving that the K-means-SVM model can identify defaulting enterprises more accurately and has higher efficiency in credit risk assessment.

5. Conclusions

The credit risk assessment model of smes under the supply chain finance model was built by SVM model, and variables with major explanatory ability are extracted by factor analysis. According to the characteristics of SVM model, K-Means clustering algorithm is used to select the clustering center and the edge points of samples far from the clustering center as training samples to train THE SVM model. The results show that compared with the single SVM model, the overall assessment accuracy of K-Means-SVM model is increased by 7.2%, and the first type error rate is reduced by 5.0%, which verifies that the K-means-SVM model under the supply chain finance model has higher identification ability of credit insurance, and better meets the requirements of commercial banks for credit risk assessment of financing enterprises.

References

[1] Aberdeen Group.Technology Platforms for Supply Chain Finance[J].Aberdeen Group,2007,11(9):3-8.

[2] Manuj I,Mentzer J.Global Supply Chain Risk Management Strategies[J].International Journal of Physical Distribution and Logistics Management,2008(3):192-223.

[3] Blome C, Schoenherr T. Supply Chain Risk Management in Financial Crises—— A Multiple Case-Study Approach[J]. International Journal of Production economics, 2011,134(1):43-57.

[4] Silvestro R, Lustrato P. Integrating Financial and Physical Supply Chains: The Role of Banks in Enabling Supply Chain Integration[J]. International Journal of Operations & Production Management,2014,34(3):298-324.

[5] Sum K. A Review of Idividual and Systemic Risk Measures in Terms of Applicability for Banking Regulations[J].Social Science Electronic Publishing,2016,10(1):71-82.

[6] Finch P. Supply Chain Risk Management Revisited [J]. Supply Chain Management, 2016,19 (3):142-156.