An Application of Support Vector Machine for Evaluating Credit Risk of Bank

Liu Hongjiu[1], Hu Yanrong[1,2], Wuchong[2]
1 Changshu Institute of Technology, Changshu, P.R.China, 215500
2 School of Management, Harbin Institute of Technology, Harbin, P.R.China, 150001
(E-mail: lionlhj@163.com, rosehyr@hotmail.com, richard_lhj2000@yahoo.com.cn)

Abstract Evaluation of credit risk is an important task before loan of a project. However, effective methods have not developed until Support Vector Machine is brought. In this paper, because SVM is a kind of general forward-feedback network, it is applied to how to evaluate credit risk of commercial loan in banks. Empirical results show that SVM is effective and more advantageous than BP neural network. It has the advantages of easy classification plane, strong generalization, good fitness and strong robust.
Key words Evaluation; Credit risk; Support vector machine

1 Introduction

Banks can provide money, lead asset and adjust supply and demand of society as finance and credit loan centers. However, they also face all kinds of risks during the process of operation, including system and non-system risks. Among non-system risks, credit risk is especially important. Study of World Bank on crisis of global bank industry shows credit risk is the most possible reason to lead banks bankrupt.

Credit risk is the uncertainty of loan safety, which shows the probability of bad debts that because enterprises are not willing or have no ability to return loan and interest. Credit risk management is a basic work of loan risk management in banks. Its aim is to analyze the willing and ability whether borrowers can repay debt.

Research of evaluation methods on credit risk began from 1930s, which became the hot point of study in 1960s. With the development of globalization and financial liberalization, the methods are improved continuously[1]-[5]. Until now, there are three developing stages: proportional analysis, statistics analysis and human intelligence. Discriminate analysis models based on statistics were brought forward after Fisher's heuristic research, such as Discriminate Analysis, Logistic Regression, Major Constituent Analysis and Clustering Analysis etc.[5]-[12]. Among them, the classic methods include Altman's Z-score Model and improved ZETA model. Although Statistics Analysis overcomes the shortcomings of weak comprehensive analysis and quality analysis, there are still some problems: (1) Statistics methods is a kind of graduate theory that samples tends infinite , which demand samples quantity; (2) Statistics models are limited by many supposes, such as Multivariate Discriminate Analysis (MDA) requires data must obey multivariate normal distribution and equal covariance. Whereas, many samples do not accord with the supposes.

After 1980s, artificial intelligence technology is applied to evaluate credit risk, such as Expert System, Neural Network(NN), which overcomes the shortcomings of statistics methods. Especially for Neural Network, it has the characteristics of self-organization, self-adaptation, and self-study. It not only has non-linear mapping and generalization, but also has stronger robust and higher precision. Nevertheless, there are some flaws for NN: (1) network structure is difficult to be determined; (2) easily plunge into local minimum with low training efficiency. Because of no enough samples, the effects are not good for statistics methods and Neural Network. In order to solve the problem, this paper brings forward a general algorithm of Support Vector Machine(SVM) and applies it to evaluate credit risks of banks.

2 Principle and Algorithm of SVM

SVM is a kind of general forward-feedback network, brought forward by Vapnik first. The idea of SUM is to build up a hyperplane as a decision-making surface which make the isolated margin maximum between positive and negative examples. Moreover, SUM is an approximate method of minimizing structure risks. The idea is based on such fact that error rate of testing data is limited in the sum of training error rate and an item depending on Vapnik Chervonenkis dimension. Under the pattern of SUM classifier, the formal value is zero, the latter value is minimized. Thus, although SUM does not need knowledge of question field, it can provide good and particular generalization ability for pattern classification[13].

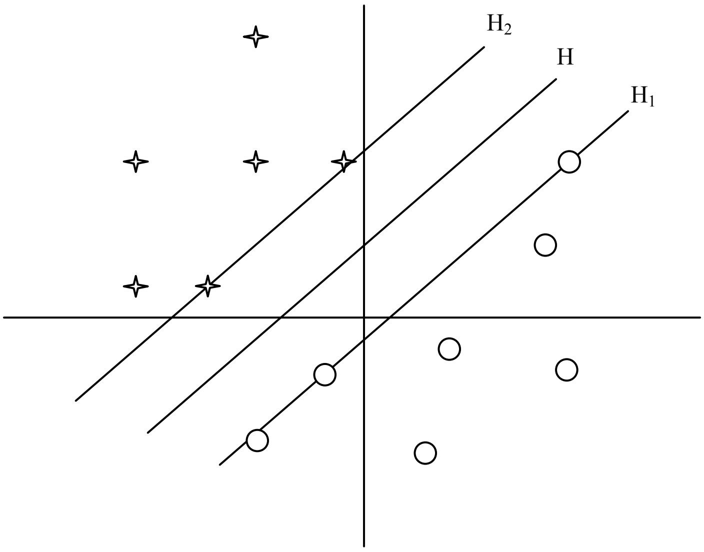

Figure 1 Linear Classification Defined by the Hyperplanes

SUM is developed from optimum classification hyperplane of linear classification. In Figure 1, star points and circular points represent two sorts of samples. H is a classification hyperplane. H_1, H_2 across the samples are the nearest hyperplane to H and parallel H hyperplane. The distance between H_1 and H is equal to that between H_2 and H, which is called classification space[14]. The optimum hyperplane is the biggest space that makes the samples separated correctly.

The decision-making curve equation is $w^T x+b=0$ for classification hyperplane. Where, x is input variable, w is adjustable weight vector, b is bias vector, y is corresponding expect (target output). Suppose $y_i=+1$ representing pattern classification is distinguished from $y_i=-1$ representing that, following piecewise function can be acquired:

$$w^T x_i+b\geq 0 \qquad y_i=+1$$
$$w^T x_i+b<0 \qquad y_i=-1.$$

Namely, linear and separable sample set (x_i, y_i) meets

$$y_i[(w^T x_i+b)]-1\geq 0 \tag{1}$$

Where, classification space equals to $2/\|w\|$. Consequently, making space maximized equates to make $\|w\|$ minimized. The optimum hyperplane meets equation (1) and makes $1/2\|w\|^2$ minimized. Support vector is the input vector making equation (1) existence.

The question of optimum classification hyperplance can be turned into dual problem by Lagrange optimum method. The dual problem has same optimum with the old problem. Therefore, the optimum can be acquired by Lagrange operator, namely that is to find the optimal solution of a quadratic function constrained by an inequation. The optimum classifier function is

$$f(x)=\operatorname{sgn}\{(w^T x)+b\}=\operatorname{sgn}\{\sum_{j=1}^{n}\alpha_j^* y_j(x_j.x)+b^*\} \tag{2}$$

Where, α_j^* is the operator of Lagrange for every sample. It can be proved that α_j^* if a part of are not zero, corresponding samples are support vectors. b^* are classification thresholds which can be acquired by any support vector or any pair of support vectors of two samples[15]. If sample set can not be divided, a slack variable can be increased to equation (1), then, equation (1) is changed as follows:

$$y_j[(w^T x)+b]-1+\xi_j\geq 0 \tag{3}$$

The aim is changed that makes $(w,\xi)=1/2\|w\|^2+C[\sum_{i=1}^{n}\xi_i]$ minimized. Namely, constructing a soft interval considering wrong-distributed samples and maximum classifier space eclectically, which is so-called generalized optimum classification plane? Where, $C>0$ is a constant that controls the extent to punishing wrong-distributed samples.

For non-linear problems, they can be converted into linear problems of a certain high dimensions through non-linear conversion. In conversion space, the optimum classification plane can be gotten. If proper inner product function $K(x_j, y_j)$ takes place of inner product of the old space, linear classification can be realized after non-linear conversion[16]. So, classification function is described as follows:

$$f(x) = \mathrm{sgn}\{\sum_{i=1}^{n} \alpha_i^* y_i K(x_I . x) + b^*\} \tag{4}$$

3 Demonstration of Credit Risk Appraisal Based on SVM

3.1 Constructing index frame

Credit risks of banks are related with credit status of loaner, distribution of loan and industrial convergence. They consist of loan enterprise risk, bank risk, macroeconomic risk and others.

According to enterprise performance appraisal index of Statistical Bureau of Ministry of Financial and enterprise credit standing index frame of Industrial and Bank of China Limited and other research of foreign and civil literatures, considering particularity of Chinese credit risks and getatability of data, sixteen variables are used to evaluate risks of banks: sales income/total asset, total asset turnover, current asset turnover, fixed asset turnover, inventory turnover, accounts receivable turnover, liquidity ratio, operation capital/total asset, quick ratio, over quick ratio, return on assets, net return on assets, credit mode.

3.2 Processing data of sample

The data of this paper resource from Harbin branch of Industrial and Bank of China. Because industry, operation environment, business arrange are different, financial and non-financial index of different enterprises are not comparative. Therefore, samples are chosen from call loan of same industry in the model. Before demonstration, samples should be processed by two or three times of standard deviation while abnormal data will be erased. Finally, one hundred and fifty seven samples are acquired. Among them, financial standing is good for eighty enterprises whose loan risks are less and named as “implementation enterprise”. Others’ financial standing is so bad that probability of default is big, which are called as “default enterprise”. Because the quantity of two sorts of samples is close, requirement of SVM is met. We can use SPSS to do factor analysis for financial data of 157 enterprises. According to eigenvalue rule, sixteen index are divided into four explain factors including operation factor, debt service factor, earning power factor and loan mode factor: On the basis of above-mentioned, sample set is divided into train sample and test sample set. In order to show the learn ability of SVM for small samples and generalization ability of the model, thirty five percent (fifty six enterprises) is chosen to construct SVM model as train sample set randomly. Other sixty five percent (one hundred one enterprises) is chosen randomly to test generalization of the model as test sample set.

3.3 Construct SVM model

According to above-mentioned analysis, we can construct sample set (x, y), where, dimension of x is 4 and y is sort attribution of sample. For “implementation enterprises”, y=1; For “default enterprises”, y=-1. If inner product kernel function chooses polynomial kernel function, radial-based function or Sigmoid kernel function, SVM can obtain the result of approximate performance, and distribution of support vectors is not big. In this paper, inner product function of SVM model selects radial-based function:

$$K(x, x_i) = \exp\{-\frac{|x - x_i|^2}{\sigma^2}\} \tag{5}$$

Comprehensively considering minimum wrong-distinguished and maximum classification space, soft space is constructed in high dimension interspaced. By cross validation, α^2=225 and $C=10^5$. Then, we can use *Matlab* 7.1 toolbox to test and analyze model.

3.4 Analysis of demonstration result

We compared the result of SVM with that of BP neural network in Table 1. In BP model, number of performance error and hidden layers is attained by cross validation (performance error=0.1, number of hidden layers=12). In Table 1, wrong rate is the value that the sum of judging “default enterprises” as “implementation enterprises” and judging “implementation enterprises” as “default enterprises” is divided by total samples. Seen table 1, the accurate rate of SVM model is 83.78 percent in test sample set. It is superior to BP mode with 77.38 percent accurate rate evidently. In addition, we also compared robust of two models. For train sample set, accurate of BP model is 82.14 percent, and that of SVM is 85.71 percent. In test sample set, accurate rate descends to some extent. BP model descends 3.6 percent and SVM descends 2.9 percent. Apparently, robust of SVM is better and can meets requirements of application.

Table 1 Distinguish Result of SVM and BP Model

Model	Train sample set(56)		Test sample set(101)	
	accurate	wrong	accurate	wrong
BP	82.14%(46)	17.864%(10)	79.2%(80)	20.8%(21)
SVM	85.71%(46)	14.29%(8)	83.17%(80)	16.83%(17)

Note: number in parenthesis is quantity of samples.

4 Conclusions

SVM is a general learn algorithm based on small samples with strict basis of theory, which can solve the problems of non-linear, high-dimension and local minimum that traditional methods can not solve. By applying SVM to evaluate credit risk of banks and comparing with BP neural network model, we find that SVM has the advantages of easy classification plane, strong generalization, good fitness and strong robust. However, However, there are still some problems to be studied such as mapped space of kernel, optimization scale.

References

[1] Shilton Alistair, Lai Daniel T. H.. Iterative fuzzy support vector machine classification. 2007 IEEE International Conference on Fuzzy Systems, 2007: 109
[2] Kostka P.S., Tkacz E.J.. Feature extraction for improving the support vector machine biomedical data classifier performance[J]. Information Technology and Applications in Biomedicine, 2008(12): 32-34
[3] Peltier Scott J. etc. Support vector machine classification of complex fMRI data. 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, 2009: 56-59
[4] Besrour R., Lachiri Z.; Ellouze N. ECG beat classifier using support vector machine. 3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008: 234-239
[5] Tamura Hiroki, Tanno Koichi. Midpoint-validation method for support vector machine classification. IEICE Transactions on Information and Systems, 2008(05): 122-125
[6] Altman E I. Corporate financial distress: a complete guide to predicting, avoiding, and dealing bankruptcy[M]. New York: John Wiley & Sons, 1983: 16-18
[7] Jackson P, Perraudin W. Regulatory implications of credit risk modeling[J]. Journal of Banking and Finance, 2000（24）: 1-14
[8] Ohlson J. Financial ratios and the probabilistic prediction of bankruptcy[J]. Journal of Accounting Research, 1980（2）:109-130
[9] Craig W R. A factor analytic approach to bank condition [J]. Journal of Banking and Finance, 1985（9）:253-266
[10] Lundy M. Cluster analysis in credit scoring. Credit scoring and credit control[M]. New York: Oxford University Press, 1993: 25-36.
[11] Altman E I, Haldeman R G, Narayanan P. Zeta analysis: a new model to identify bankruptcy risk of corporations[J]. Journal of Banking and Finance, 1977,1(1):29-54
[12] Kaastra I, Boyd M. Forecasting futures trading volumes using neural networks. Journal of Futures Markets, 1995,(15):953-970
[13] Boritz J E, Kennedy D B. Effectiveness of neural network types for prediction of business failure[J]. Expert System with Applications, 1995,9(4):503-512
[14] Lee K C, Han I, Kwon Y. Hybrid neural network models for bankruptcy predictions[J]. Decision Support System, 1996,(18):63-72
[15] Burges C J C. A tutorial on Support Vector Machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998,2(2):955-974
[16] Vapnik V N. The nature of statistical learning theory[M]. New York: Spring-Verlag, 1995: 42