Credit Risk Assessment in Commercial Banks Based on Fuzzy Support Vector Machines

Qifeng ZHOU, Chengde LIN
Automation Department
Xiamen University
Xiamen, China 361005
Email: zhouqf@xmu.edu.cn

***Abstract*— Credit risk assessment plays an important role in banks credit risk management. The objective of credit assessment is to decide credit ranks, which denote the capacity of enterprises to meet their financial commitments. Traditional "one-versus-one" approach has been commonly used in the multi-classification method based on Support Vector Machine (SVM). Since SVM for pattern recognition is based on binary classification, there will be unclassifiable regions when extended to multi-classification problems. Focus on this problem, a new credit risk assessment model based on fuzzy SVM is introduced in this paper that can give a reasonable classification for unclassifiable examples. Experiment results show that the fuzzy SVM method provides a better performance in generalization ability and assessment accuracy than conventional one-versus-one multi-classification approach.**

***Keywords*— credit risk assessment, multi-classification, fuzzy support vector machine**

I. INTRODUCTION

Credit risk assessment is essentially a classification problem. It plays an important role in banks credit risk manag [1] ement. An accurate assessment of risk could be translated into a more efficient use of resources or a less loan loss for a bank. In recent years, there has been a lot of finance institutes that employ analysis tools based on intelligent techniques such as Expert System (ES), Artificial Neuron Net (ANN), Genetic Algorithm (GA), etc [1][2]. Now, a new method called as Support Vector Machine (SVM), which is based on Statistical Learning Theory, has being widely used in pattern recognition problems. It is also applied in finance problems such as financial distress prediction of listed firms, enterprises credit risk assessment and personal credit scoring … etc, and has acquired good performances [3][4].

According to the common-rules of credit rating, we usually classify the enterprises into 3 ~ 5 or more ranks. That is, credit risk assessment is a multi-classification problem. SVM is originally a binary classifying technique, and there are some methods to extend it to adapt to the situation, such as one-versus-rest, one-versus-one, directed acyclic graph (in short, 1-v-r, 1-v-1, DAG, respectively) strategies, and … etc. These methods are all based on the combination of the binary classifiers. Among them, '1-v-1' is commonly considered to be more efficient [5]; but in practice, when we use simple majority voting method to decide the example's rank there will be some examples which are non-determinable, so they will decrease the accuracy of the evaluating.

Focus on this problem, Fuzzy Support Vector Machine (Fuzzy-SVM) [6] is applied in this paper. For each credit rank we will adopt an one-dimensional membership function which is orthogonal to the optimal separating hyper-plane. In the classifiable region, the classifying results are the same with 1-v-1 approach; in the non-classifiable region it can also give an appropriate classifying result. The experiment results show that the performance of the model based on Fuzzy-SVM is satisfied and this technique can well be applied to credit risk assessment.

This paper is organized as following: the theory of SVM and multi-classification are briefly introduced in section 2; Fuzzy-SVM is presented in section 3; in section 4, credit risk assessment model based on Fuzzy-SVM is constructed; finally, several modeling approaches are compared and conclusions are drawn in section 5.

II. MULTI-CLASSIFYING METHOD BASED ON SVM

A. Support Vector Machine

Let

$$\mathrm{T} = \{(x_i, y_i)\}_{i=1}^{L} \subset \Omega \times \mathrm{Y} \tag{1}$$

be a training data set following an unknown probability density function. The problem of multi-classification from examples addresses the general problem of finding a decision function $f(x,\omega)$ approximating to an unknown function, defined from an input space $\Omega \subset R^n$ into a set of classes $Y = \{1,\ldots K\}$ having the smaller discrepancy with the real system answer.

Support Vector Machines that learns classification problems –in short SVMC– are specific to binary classification problems. In their basic form, SVM learns linear decision rules $f(x,\omega) = sign(\omega' \cdot x + b)$ described by a weight vector ω and a threshold b. Its basic idea is to map data into a high

This work was supported by the second phase of the "985" project, "Information Technique Platform", Xiamen University, P.R. China.

dimensional space and find a separating hyper-plane with the maximal margin. In the SVMC binary pattern recognition problem, given training vectors $x_i \in R^n \quad i=1,....L$ with outputs $y \in \{\pm 1\}$, computing the separating hyper-plane is equivalent to solving the following optimization problem, (OP1):

$$\begin{aligned} \min_{w,b,\varepsilon} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^{L} \xi_i , \\ s.t \quad & y_i (w^T \cdot x_i + b) \geq 1 - \xi_i , \\ & \xi_i \geq 0, i = 1, ..., L. \end{aligned} \tag{2}$$

where ξ_i is the slack variable. The factor C is a parameter that allows one to trade off training error vs. model complexity. Instead of solving OP1 directly, it is easier to solve its dual form (OP2):

$$\begin{aligned} \max \quad & \sum_{i=1}^{L} \alpha_i - \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j), \\ s.t \quad & \sum_{i=1}^{L} y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, ..., L. \end{aligned} \tag{3}$$

The above SVMC is an essentially linear method, but it can be easily generalized to non-linear decision rules by replacing the inner-products $(x' \cdot x)$ with a kernel function $k(x', x) := < \phi(x'), \phi(x) >$ which accomplishing Mercer's theorem [7], where $\phi : \Omega \to F$ being a non-linear mapping.

B. *Multi-class Support Vector Machines*

The standard method of decomposing multi-classification problem into dichotomies is to place binary classifiers in parallel. In the original method [9], the *i*-th SVMC is trained with positive labels for all the examples in the *i*-th class, and negative labels for all other examples. We refer to SVM trained in this way as 1-v-r SVMC. The training time of the standard method scales linearly with K.

Another common method to construct multi-class classifiers is to build all the possible binary classifiers $K(k-1)/2$ hyper-planes decision functions from training set of classes, each classifier being trained on only two out of classes. We refer to the SVM trained with this method as 1-v-1 SVMC. The size of the 1-v-1 classifiers may grow superlinearly with K. The combination of these binary classifiers to determine the label assigned to each new input can be made by different schemes, such as voting. The 1-v-1 approach is preferable to the 1-v-r [5] in general. As the number of class in credit risk assessment is not very large, so usually, we choose 1-v-1 as the multi-classifying method.

III. FUZZY SUPPORT VECTOR MACHINE

A. *Unclassifiable Pattern in Conventional One-versus-one Method*

1-v-1 method constructs all the K·(K-1)/2 possible binary machines from a K-class training set, each SVMC being trained on only two out of all K classes. Since the extension to nonlinear decision functions is straightforward, to simplify discussions, we consider linear decision functions. Let the decision function for class i against class j, with the maximum margin is

$$h_{ij}(x, \omega) = < \omega_{ij} \cdot x > + b_{ij} \tag{4}$$

where ω_{ij} is the m-dimensional vector, b_{ij} is a scalar, and $h_{ij}(x, \omega) = -h_{ij}(x, \omega)$. For the input vector x we calculate

$$h_i(x) = \sum_{j=1, j \neq i}^{n} sign(h_{ij}(x, \omega)), \tag{5}$$

and classify x into the class

$$\arg \max_{i=1,...,n} h_i(x). \tag{6}$$

If (6) is satisfied for multiple i, the sample x is unclassifiable. In Fig.1, the shaded region depicts the unclassifiable zone for 1-v-1 method on a three-class problem. For illustration purposes, the original space has been considered as the feature space, i.e. a linear kernel.

B. *The Membership Functions*

Here, we use a membership function to resolve unclassifiable region while realizing the same classification results with that of the conventional 1-v-1 classification. To do this, for the optimal separating hype-plane $h_{ij}(x, \omega) = 0$ we define one-dimensional membership functions $\mathrm{m}_{ij}(\mathrm{x})$ on the direction orthogonal to $h_{ij}(x, \omega) = 0$ as follows:

$$\mathrm{m}_{ij}(\mathrm{x}) = \begin{cases} 1 & \text{for } \mathrm{h}_{ij}(\mathrm{x}) \geq 1 \\ \mathrm{h}_{ij}(x) & otherwise \end{cases}. \tag{7}$$

Using $\mathrm{m}_{ij}(\mathrm{x}) \; (\mathrm{j} \neq \mathrm{i}, \mathrm{j} = 1, \ldots, \mathrm{n}),$ we define the class i membership function of x using the minimum operator:

$$\mathrm{m}_{\mathrm{i}}(\mathrm{x}) = \min_{\mathrm{j}=1,\ldots\mathrm{n}} \mathrm{m}_{\mathrm{ij}}(\mathrm{x}). \tag{8}$$

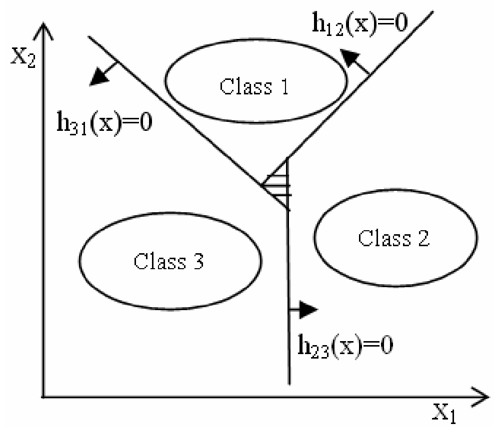

Figure 1. Unclassified regions by the 1-v-1 method

Equation (8) is equivalent to

$$m_i(x) = \min\left(1, \min_{j=1,\dots n, j\neq i} h_{ij}(x)\right). \tag{9}$$

Since $m_i(x) = 1$ holds for only one class, (9) reduces to

$$m_i(x) = \min_{j=1,\dots n, j\neq i} h_{ij}(x). \tag{10}$$

Now an unknown datum x is classified into the class

$$\arg\max_{i=1,\dots n} m_i(x). \tag{11}$$

Thus, the unclassified region shown in Fig. 1 is resolved as shown in Fig. 2.

IV. CREDIT RISK ASSESSMENT MODEL BASED ON FUZZY SUPPORT VECTOR MACHINE

Credit risk assessment is a synthetical evaluation to a company, finance organization etc according to theirs credit-ability. Usually we use the companies finance states to denote their credit-ability. The methods can be described simply as studying the history sample, finding the rule from the sample, structuring a classification model, and applying it to predict new example's credit rank. It is essentially a multi-classification problem [10].

In a standard classification problem the input vector are associated with one of the *K* classes, thus $y_i \in Y = \{1, \dots, K\}$ denoting the class membership. In credit risk assessment problem, Y is a finite set such as {AAA, AA, A, B, C} which denotes the "excellent", "very good", "good", "general" and "bad" credit degree respectively.

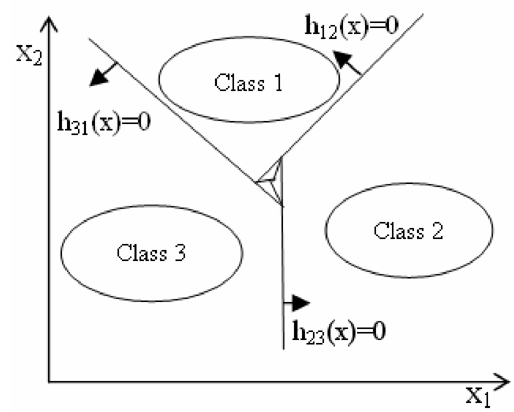

Figure 2. Fuzzy-SVM classification method

Usually we evaluate the companies' credit-ability based on their financial status. Since different industries pattern have different financial status. So we divide the companies into several patterns according to their industry.

According to the common-rules and expert's suggestion, we assess a company's financial status from three aspects: solvent ability, payoff ability and management ability [11]. In order to compare the performance of each multi-classification approaches in credit ranking problem, we select a subset of the database obtained from some business bank's client information as modeling data. After pre-processing, about 1000 light industry enterprises are selected, including 48.3% 'AAA' companies, 42.3% 'AA' companies, 8.49% 'A' companies, and 0.91% 'B' companies. Since the rank 'B' companies are few, we combined them with rank 'A' companies. Thus the total ranks are K=3, or Y= {1, 2, 3}, where 1, 2, 3 corresponding to 'AAA' rank, 'AA' rank, and 'ABC' rank respectively. In each rank, we randomly select 75% samples as training set and the rest 25% as testing set, and adopt 'Libsvm' as the tool to construct meta-classifier. Similarly, the data sets of business and heavy industry are selected respectively. In conventional 1-v-1 algorithm, in order to give a certain class label for each example, when the maximum votes are equal for more than one rank, the example was classified into the maximum votes' rank which appeared firstly, thus the error increased in unclassifiable region. Fuzzy-SVM can resolve this problem preferably.

TABLE I. PERFORMANCE COMPARE BETWEEN 1-V-1 AND FUZZY-SVM APPROACH

Industry class	**#test ptr (total ptr)**	**#atr**	**#class**	**#average test accuracy (%)**	
				1-v-1	***Fuzzy-SVM***
light industry	167(837)	12	3	80.24.	81.43
Business	110(552)	14	3	79.09	79.68
heavy industry	42(209)	13	3	77.58	78.16

In our experiments, each pattern's data set is randomly partitioned with a fifth of the data reserved to validate the model's performance. This process is repeated 100 times, the average accuracy showed in Table 1.

From Table 1 we can see that the performance of the evaluating models based on Fuzzy-SVM is superior to the others; its evaluating accuracy is higher than the conventional 1-v-1 method.

V. CONCLUSION

In this paper, we have proposed a credit risk assessment approach based on Fuzzy-SVM, and experimented on the data of several practical industry patterns. Using a membership function, Fuzzy-SVM can give a reasonable classification for unclassifiable examples in conventional 1-v-1 classification.

The experiment results show that the fuzzy support vector machines approach provided a better performance in generalization abilities and assessment accuracy than conventional 1-v-1 multi-classification approach.

REFERENCES

[1] Lee K C, Han I, Kwon Y. "Hybrid neural network models for bankruptcy predictions [J]". Decision Support System, (18), pp.63-72, 1996.

[2] Boritz JE, Kennedy DB,"Effectiveness of neural net work types for prediction of business failure [J]." Expert System with Applications, Vo.l.9 (4), pp.503-512, 1995.

[3] LIU Yuntao, WUChong,"The model of credit risk assessment in commercial banks based on Support Vector Machines." Forecasting, Vol.24 (1), pp.52-55, 2005

[4] Jianping Li, Jingli Liu, Weixuan Xu, "Support Vector Machines approach to credit assessment", ICCS 2004, LNCS 3039, pp.892-899,2004.

[5] M. Moreira, E. Mayoraz, "Improved pairwise coupling classification with correcting classifiers", in: C.NZedellec, C. Rouveirol (Eds.), Proceedings of the ECML-98, Chemnitz, Germany, pp. 160–171. 1998

[6] Shigeo Abe and Takuya Inoue, ESANN'2002 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium), 24-26 April 2002, d-side public., pp. 113-118.

[7] V. Vapnik, "The nature of statistical learning theory", Springer, New York, 1995.

[8] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining Knowledge Discover. 2 (1998) 1–47.

[9] C. Angulo, "Learning with kernel machines into multiclassification frameworks", Automatic Control Department, Technical University of Catalonia, Barcelona, Spain, 2001.

[10] W.X.Wei, "The comprehensive evaluation method of the credit rating of enterprise and its application," System Engineering—Theory & Practice, Vol.5 (2), pp26-31, 1998.

[11] Y.F.He; D.L.Shi, "Banks client credit assessment," Business Publishing House, Beijing, China 2002