Expert Systems with Applications 30 (2006) 427–435

Expert Systems with Applications

www.elsevier.com/locate/eswa

A study of Taiwan's issuer credit rating systems using support vector machines

Wun-Hwa Chen [a,1], Jen-Ying Shih [b,1,*]

[a] Graduate Institute of Business Administration, National Taiwan University, Taiwan, ROC
[b] Securities and Futures Institute, 9F, 3, Nan-Hai Rd., Taipei 100, Taiwan, ROC

Abstract

By providing credit risk information, credit rating systems benefit most participants in financial markets, including issuers, investors, market regulators and intermediaries. In this paper, we propose an automatic classification model for issuer credit ratings, a type of fundamental credit rating information, by applying the support vector machine (SVM) method. This is a novel classification algorithm that is famous for dealing with high dimension classifications. We also use three new variables: stock market information, financial support by the government, and financial support by major shareholders to enhance the effectiveness of the classification. Previous research has seldom considered these variables. The data period of the input variables used in this study covers three years, while most previous research has only considered one year. We compare our SVM model with the back propagation neural network (BP), a well-known credit rating classification method. Our experiment results show that the SVM classification model performs better than the BP model. The accuracy rate (84.62%) is also higher than previous research.

Keywords: Credit ratings; Support vector machines; Taiwan's banking industry

1. Introduction

An issuer credit rating is a current assessment of an issuer's overall capacity to meet its financial obligations based on its ability and willingness to meet its financial commitments on a timely basis. Such information is a useful and important reference for raising corporate capital, the granting of credit by banks, and providing credit risk information for investment decisions. It therefore benefits most participants in financial markets, including issuers, investors, market regulators and intermediaries.

Although rating agencies claim that both financial and non-financial information is considered in the rating decision process, their rating criteria are not explicit. Thus, many researchers have tried to construct automatic classification systems by using data mining methods, such as statistical and artificial intelligence (AI) techniques. The former include linear regression (Horrigan, 1966), linear multivariate discriminant analysis (MDA) (Belkaoi, 1980), quadratic MDA (Pinches & Mingo, 1975), probit regression (Ederington, 1985), logit analysis (Ederington, 1985) and multidimensional scaling (Mar, Apellaniz & Cinca, 1996). The latter consist of back propagation neural networks (BP) (Dutta & Shekhar, 1988) and case base reasoning (CBR) (Shin & Han, 2001).

A novel classification technique, the support vector machine (SVM) proposed by Vapnik, (1999), has been successfully applied and evaluated in many areas, including financial time series forecasting (Tay & Cao, 2001), credit scoring (Baesens et al., 2003), the prediction of protein structural classes (Cai & Lin, 2002), drug design (Burbidge et al., 2001) and the identification of organisms (Morris & Autret, 2001). However, very few researchers have studied the suitability of SVMs for credit rating systems. In this paper, we apply the SVM method to the automatic classification of issuer credit ratings. Our goal is to test the SVM's capacity to perform the classification of credit ratings, not to examine the speed of a new algorithm.

The classification categories of previous credit rating research were based on issue credit ratings, including bond ratings (Ederington, 1985) and commercial paper ratings (Shin & Han, 2001), instead of issuer credit ratings. The main reason is that the classification accuracy rates of issue credit ratings can be improved by some issuing variables, such as subordination, the scale, and the period of issue (Horrigan, 1966; Pinches & Mingo, 1975; West, 1970). However,

* Corresponding author. Address: Graduate Institute of Business Administration, National Taiwan University, 9F, 3, Nan-Hai Rd, 100 Taipei, Taiwan. Tel.: +886 2 23574357; fax: +886 2 23224302.
E-mail address: d88741008@ntu.edu (J.-Y. Shih).
[1] These authors contributed equally to this work.

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.10.003

we believe that issuer credit ratings provide a more fundamental type of credit risk information because issue credit ratings are generally determined on the basis of issuer credit ratings. Generally, junior debt may be rated below the issuer credit rating, while well-secured debt can be rated above. We therefore propose an issuer credit rating model to explore this research area.

Most previous research has focused on using financial variables as input variables. According to (Atiya, 2001), trading information, such as the volatility of stock prices, can be taken into consideration when measuring the possibility of bankruptcy. Thus, we included a market information variable in the classification model, namely: each share's average daily closing price in the previous year. Also, according to the rating criteria published by Standard and Poor's and the Taiwan ratings corporation (TRC), financial support is very important in determining ratings. Therefore, two financial support variables regarding government and major shareholders are included in the classification model. In addition, it should be noted that the data period of the input variables used in our model covers the three years prior to the rating year, whereas previous researchers have only used data for the year immediately before the rating year. In our opinion, it is more appropriate to use multi-year data because rating agencies use it as the benchmark for the rating decisions.

In this research, we choose the banking industry to test the applicability of SVM models for the issuer's credit rating. Banks play a pivotal role in an economy, and, as loans are one of their major sources of revenue, it's important that the level of non-performing loans (NPL's) is kept to a reasonable minimum. In Taiwan, however, the ratio of NPL's has increased rapidly in the last decade because of a steep decline in the local stock market and over-competition between banks. The Economist (Nov. 11, 2000) reported that bad loans in Taiwan's domestic banks had reached new highs, while Business Week (Dec. 11, 2000) pointed out that the NPL ratio among listed banks in Taiwan amounted to more than six percent, However, because of the narrow definition of official NPL statistics, the ratio could, in reality, be as high as 10–15 percent. As rising NPLs may increase the banks' risks, it's important to be able to measure those risks. In addition, in most case, the banking industry is highly regulated, information about its operations is usually open to the public and relatively easy to obtain.

The remainder of this paper is organized as follows. In Section 2, we briefly introduce the SVM algorithm, which is the research method we applied in this paper. In Section 3, we describe the test bed, i.e. Taiwan's banking industry. The research design is described in Section 4, while our experiment results are presented and discussed in Section 5. Finally, in Section 6, we present our conclusions and suggestions for further research.

2. SVM methodology

The SVM model is a type of learning machine that is based on statistical learning theory. The basic procedure for applying SVMs to a classification model can be stated briefly as follows. First, map the input vectors into a feature space, which is possible with a higher dimension. The mapping is either linear or non-linear, depending on the kernel function selected. Then, within the feature space, seek an optimized division, i.e. construct a hyper-plane that separates two or more classes. Using the structural risk minimization rule, the training of SVMs always seeks a globally optimized solution and avoids over-fitting. It, therefore, has the ability to deal with a large number of features. The decision function (or hyper-plane) determined by a SVM is composed of a set of support vectors, which are selected from the training samples. A complete description of the theory of SVMs can be found in Vapnik, (1999).

2.1. Support vector classification (SVC)

In the following, we briefly describe how the support vector classification (SVC), which is a type of SVM algorithm for classification purposes, can determine an optimal separating hyper-plane from the division composed of some of the existing samples. The basic principle is to find a maximum margin, i.e. maximize the distance between the hyper-plane and the nearest data point of each class.

2.1.1. Linear separable case

Consider the problem of separating a set of training vectors belonging to two separate classes, $(x_1, y_1), ..., (x_l, y_l)$, $x \in R^n$, $y \in \{+1, -1\}$, where x are input vectors, and y are output vectors. The data set can be separated by a hyper-plane as two classes, +1 and –1. The set of vectors is said to be optimally separated by the hyper-plane if it is separated without error and the distance between the closest vector and the hyper-plane is maximal. In such a case, the classifier is called the largest margin classifier. The hyper-plane in canonical form must satisfy the following constraints,

$$y_i[\langle w, x\rangle + b] \geq 1, \quad i = 1, ..., l \tag{1}$$

w is the weight vector and b is the bias. The optimal hyper-plane is obtained by maximizing the margin ρ, subject to the constraints of Eq. (1). The margin is given by:

$$\rho(w,b) = \min_{x_i, y_i=-1} d(w,b;x_i) + \min_{x_i, y_i=1} d(w,b;x_i) = \min_{x_i:y_i=-1} \frac{|\langle w, x_i\rangle + b|}{||w||} + \min_{x_i:y_i=1} \frac{|\langle w, x_i\rangle + b|}{||w||} = \frac{2}{||w||} \tag{2}$$

where $d(\,)$ is the distance function. Hence, the hyper-plane that optimally separates the data is the one that minimizes $\Phi(w) = 1/2||w||^2$, subject to constraint (1). By applying the Lagrange relaxation method, the problem can be formulated as follows:

$$\min_{w,b} \Phi(w,b,\alpha) = \frac{1}{2}||w||^2 - \sum_{i=1}^{l} \alpha_i (y_i[\langle w, x_i\rangle + b] - 1), \tag{3}$$

In essence, the SVM training procedure is the same as solving a convex quadratic programming problem. The solution is