Expert Systems with Applications 38 (2011) 2966–2973

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Usefulness of support vector machine to develop an early warning system for financial crisis

Jae Joon Ahn [a], Kyong Joo Oh [a,*], Tae Yoon Kim [b,1], Dong Ha Kim [a]

[a] Department of Information and Industrial Engineering, Yonsei University, 134, Shinchon-Dong, Seodaemun-Gu, Seoul 120-749, South Korea

[b] Department of Statistics, Keimyung University, Daegu 704-701, South Korea

ARTICLE INFO

Keywords:
Traditional financial crisis
EWS classification
Support vector machine

ABSTRACT

Oh, Kim, and Kim (2006a), Oh, Kim, Kim, and Lee (2006b) proposed a classification approach for building an early warning system (EWS) against potential financial crises. This EWS classification approach has been developed mainly for monitoring daily financial market against its abnormal movement and is based on the newly-developed crisis hypothesis that financial crisis is often self-fulfilling because of herding behavior of the investors. This article extends the EWS classification approach to the traditional-type crisis, i.e., the financial crisis is an outcome of the long-term deterioration of the economic fundamentals. It is shown that support vector machine (SVM) is an efficient classifier in such case.

1. Introduction

Recently the financial crises have been sweeping across the major economies and threatened the stability of the international monetary market. To be prepared against such crises, it is quite necessary to have proper early warning system (EWS) that alarms against possible financial crises. In fact, numerous theoretical and empirical works have been done to investigate the economic or financial crisis and build EWS (Goldstein, 1996; Kaminsky & Reinhart, 1999; Khalid & Kawai, 2003). It is a well-accepted view that there are two kinds of explanations regarding financial crises. A classical explanation is that financial crisis is an outcome of long-term deterioration of economic fundamentals. A recent view, however, asserts that financial crisis is self-fulfilling in the sense that crisis occurs by change of expectations of market participants, not entirely by fundamental change of market conditions (see, e.g., Obstefeld, 1986). Indeed, the recent view asserts that stock market crash or frenzied selling mainly driven by irrational herding behavior could lead to a major financial crisis. A typical example of this phenomenon is the financial crisis that many Asian nations had experienced in 1997. Against this type of crisis, it is essential to prepare EWS monitoring market movement on a *daily* basis and Oh, Kim, and Kim (2006a), Oh, Kim, Kim, and Lee (2006b) successfully developed such a system via classification approach, say DFCI (daily financial market condition indicator). Roughly speaking, their EWS (or EWS classification approach) monitors and classifies the financial market on the basis of daily financial market movement. In order to achieve desirable accuracy of their EWS, it is important to find an efficient classifier and Kim, Hwang, and Lee (2004), Kim, Oh, Sohn, and Hwang (2004) actually found that neural network (NN) may monitor daily financial market effectively and Oh and Kim (2007) showed that an improvement over NN is possible by case-based reasoning (CBR).

In this article we extend the EWS classification approach to produce EWS for the traditional-type crisis. Recall that this definition of crisis focuses on long-term and continuous deterioration of economic fundamentals as the major source of trouble. Since there are quite a number of such economic fundamental variables, it is occasionally the main concern of the traditional-type crisis to find "a proper subset of the variables" that describes the crisis efficiently. For instance, Kaminsky, Lizondo, and Reinhart (1998) proposed to use a signal to noise ratio to select a proper subset of given crisis-related variables. This selection procedure might be technically necessary in view of sparsity of data which is originated from the sparsity of crisis itself. In other words, the selection procedure usually comes to support the technically desirable condition that the dimension of data is to be less than the size of data. In this article, our EWS (based on the EWS classification approach) considers all given crisis-related variables available instead of finding a small subset of them and then builds *monthly financial market condition indicator* (MFCI). Note that DFCI uses *daily* financial market data that are plentiful in their character and hence rarely suffers from sparsity of data.

There are two major findings in this article. The first finding is that support vector machine (SVM) is an appropriate EWS classifier for the traditional-type crisis. The second finding is that MFCI based on monthly economic variables behaves similarly to DFCI

* Corresponding author. Tel.: +82 2 2123 5720; fax: +82 2 364 7807.
E-mail addresses: redsamzang@yonsei.ac.kr (J.J. Ahn), johanoh@yonsei.ac.kr (K.J. Oh), tykim@kmu.ac.kr (T.Y. Kim), donghakim@yonsei.ac.kr (D.H. Kim).
[1] Tel.: +82 53 580 5533.

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.08.085

based on daily financial market movements. The latter is interesting from economic point of view since MFCI does not utilize the daily financial market movement used for DFCI but behaves similarly as DFCI. These will be verified empirically by applying the EWS classification to building MFCI for the Korean economy. In addition, it is technically worth mentioning that we introduce a new transformation of crisis-related variables to detect traditional-type crisis more effectively. In fact, the "run" of the crisis-related variable is included as a monitoring variable against continuous deterioration of the economic fundamentals. See, e.g., X2, X9, X13, X14, X17 and X18 of Table 2. The rest of the study consists of as follows: Section 2 discusses the EWS classification approach and SVM. In Section 3, we construct MFCI for the Korean financial market via SVM and compare it with other classifiers. Lastly, the concluding remarks are given in Section 4.

2. Technical issues

2.1. Support vector machine

SVM is a learning machine that can perform pattern recognition tasks based on the statistical learning theory presented by Vapnik (1998). The basic concept of SVM considers a typical two-class classification problem (see Kecman, 2001; Schölkopf and Smola, 2000; Cristianini and Shawe-Taylor, 2000). SVM is known as the algorithm that finds a special kind of linear model, the maximum margin hyperplane. Consider the problem of separating the set of training vector belonging to two separate classes, $G = \{(xi, y_i), i = 1, 2, \ldots, N\}$ with a hyperplane $\mathbf{w}T\mathbf{x} + b = 0$ ($\mathbf{w} \in R^d$ is a normal vector, $\mathbf{x}_i \in R^d$ is the ith input vector and $y_i \in \{-1, 1\}$ is a known binary target). The original SVM classifier satisfies the following conditions:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geqslant 1, \quad i = 1, 2, \ldots, N. \tag{1}$$

The maximum distance between two hyperplanes, $\mathbf{w}^T\mathbf{x}_i + b = 1$ and $\mathbf{w}^T\mathbf{x}_i + b = -1$, is $2/\|\mathbf{w}\|$. Hence, we can find the optimal hyperplane by solving the optimization problem:

$$\min_{b,\mathbf{w}} \frac{1}{2}\|\mathbf{w}\|^2 \tag{2}$$

under the constraints of Eq. (1). The solution to the above problem is equivalent to determining the saddle point of the Lagrange function, i.e.,

$$\min_{\mathbf{w},b} L = \min_{\mathbf{w},b}\left\{\frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{N}\alpha_i[y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1]\right\} \tag{3}$$

subject to $\alpha_i \geqslant 0$ for $i = 1, 2, \ldots, N$. At the optimal point, we have the following saddle point equations:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \quad \text{and} \quad \frac{\partial L}{\partial b} = 0 \tag{4}$$

which implies

$$\mathbf{w} = \sum_{i=1}^{N}\alpha_i y_i \mathbf{x}_i \quad \text{and} \quad \sum_{i=1}^{N}\alpha_i y_i = 0. \tag{5}$$

Substituting Eq. (5) into Eq. (3), we obtain the dual quadratic optimization:

$$\max L_D = \max\left[\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \mathbf{x}_i^T\mathbf{x}_j\right] \tag{6}$$

subject to $\sum_{i=1}^{N}\alpha_i y_i = 0$ and $\alpha_i \geqslant 0$ for $i = 1, \ldots, N$. The Karush–Kuhn–Tucker (KKT) optimality conditions play an important role in determining the optimal value of the b and $\mathbf{w}$, respectively, i.e.,

$$b = y_i - \mathbf{w}^T\mathbf{x}_i \quad \text{for } i = 1, \ldots, N, \tag{7}$$

$$\mathbf{w} = \sum_{i=1}^{N}\alpha_i y_i \mathbf{x}_i. \tag{8}$$

Note that the classification task is only a function of the support vectors, the training data that lie on the margin.

For a non-separable case, we introduce a slack variables ξ_i and penalty C to formulate soft margin as follows:

$$\min \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i \tag{9}$$

subject to $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geqslant 1 - \xi_i$, $i = 1, 2, \ldots, N$ and $\xi_i \geqslant 0$ for all i. By applying the Lagrangian technique to Eq. (9), we will have a similar dual quadratic optimization problem like Eq. (6),

$$\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \mathbf{x}_i^T\mathbf{x}_j \tag{10}$$

subject to

$$\sum_{i=1}^{N}\alpha_i y_i = 0, \quad 0 \leqslant \alpha_i \leqslant C, \ \ i = 1, \ldots, N. \tag{11}$$

Since real-life pattern recognition problems cannot be solved using linear machine learning algorithms, a non-linear decision function must be applied. Hence, the linear classification problem must be converted into a non-linear classification problem by the mapping function Φ. By applying the kernel function $(\Phi^T(\mathbf{x}_i), \Phi(\mathbf{x}_j)) = K(\mathbf{x}_i, \mathbf{x}_j)$ to Eq. (11), we will obtain a following equation:

$$\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \tag{12}$$

subject to Eq. (11). Followed by the steps described in the liner generalized case, we obtain decision function of the following form:

$$f(x) = \text{sign}\left(\sum_{i=1}^{N}\sum_{j}^{N}\alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b\right). \tag{13}$$

Any function satisfying the Mercer's condition (Vapnik, 1995) can be used as the kernel function. There are different kernel functions used in SVM, such as linear, polynomial, and radial basis function (RBF). In this study, RBF kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2)$ is used as a kernel function of SVM since the RBF kernel tends to give good performance under general smoothness assumptions. Note that SVM has reported excellent performances in financial applications. For instance, time series prediction such as stock price index (Cao & Tay, 2001; Tay & Cao, 2002), classification such as credit rating (Huang, Chen, Hsu, Chen, & Wu, 2004), and bankruptcy prediction (Fan & Palaniswami, 2000) are the main applicable areas of SVM.

2.2. EWS classification approach

Kim, Hwang, et al. (2004), Kim, Oh, et al. (2004) first proposed the classification approach for establishing EWS for the economic crisis and Oh et al. (2006a, 2006b) has used it to develop *daily* financial condition indicator (DFCI). DFCI proved to register a good performance in judging the given financial market condition in the sense that it reflects the real financial market situation fairly accurately. In the EWS classification approach, the financial market conditions are classified into three phases: (i) stable period (SP), (ii) unstable period (UP), and (iii) crisis period. It is UP that gives the unique feature to the classification approach, i.e., the EWS based on the classification approach is designed to activate (or issue a warning) whenever the financial market enters the UP. As the UP usually occurs just prior to a crisis, it can be interpreted as a