Expert Systems with Applications 30 (2006) 535–542

Expert Systems with Applications

www.elsevier.com/locate/eswa

Kernel methods for short-term portfolio management

Huseyin Ince [a], Theodore B. Trafalis [b,*]

[a] *School of Business Administration Studies, Gebze Institute of Technology, Çayirova Fab., Yolu, No: 101 P.K:141 41400, Gebze/Kocaeli, Turkey*

[b] *School of Industrial Engineering, The University of Oklahoma, 202 West Boyd, Ste 124, Norman, OK 73019, USA*

Abstract

Portfolio optimization problem has been studied extensively. In this paper, we look at this problem from a different perspective. Several researchers argue that the USA equity market is efficient. Some of the studies show that the stock market is not efficient around the earning season. Based on these findings, we formulate the problem as a classification problem by using state of the art machine learning techniques such as minimax probability machine (MPM) and support vector machines (SVM). The MPM method finds a bound on the misclassification probabilities. On the other hand, SVM finds a hyperplane that maximizes the distance between two classes. Both methods prove similar results for short-term portfolio management.

Keywords: Support vector machines; Minimax probability machine; Kernel methods; Portfolio management; Earning announcements

1. Introduction

The objective of this study is to develop a model for stock selection by using state of the art machine learning techniques, such as minimax probability machine (MPM) and support vector machines (SVMs). Stock selections decisions are derived from earning announcements and volatility of stocks around the earning time. The majority of the studies on stock markets and earning announcements focus on the information content and timeliness- measured by changes in the characteristics of stock return distribution (e.g. mean, variance, and serial correlation) and trading volume—when corporate earnings are announced to the market (Eilifsen, Knivsflae, & Sættem, 2001). Numerous studies, including Ball & Brown (1968), Chari, Jagannathan, and Ofer (1988), Easton & Zmijewski (1989), Gennotte & Truemann (1996), and Kross & Schroeder (1984), find that stock prices respond positively to announcements of increase in earnings and negatively to announcements of decrease in earnings for the USA firms. Beaver (1968) argued that earnings announcements possess information content if stock price volatility and/or trading volume increase around the time of the announcement (see also Atiase, 1985; Bamber, 1987; Bamber, Barron, & Stober, 1997; Barron, 1995; Holthausen & Verrecchia, 1990). According to this view, stock prices reflect the market's aggregated (or average) interpretation of the information, while trading volume measures investor activity, and thus reflects differential beliefs among investors.

Three previous studies have examined short-term price movements around earnings announcements—Chari et al. (1988), covering the 1976–1984 period, Ball & Kothari (1991), covering the years 1980–1988, and Trueman et al. (2003) covering January 1998 to August 2000. There are two primary differences between the findings of the first two studies and the findings of Trueman et al. (2003). First, while the stocks in the first two studies' samples do increase in price prior to earnings announcements, there is no consistent price movement (either positive or negative) afterwards. Second, the magnitude of the pre-announcement returns they found is very small in comparison to what the third study reports (less than one tenth in size). Furthermore, those returns are significant only for the day before and the day of the earnings announcement.

Quarterly earnings announcement is probably one of the most highly anticipated events and receives significant media and investor attention. Research consistently shows that the market assimilates more and more of the information in earnings as the announcement date approaches (see Kothari, 2001; for a survey of this research starting with Ball & Brown, 1968). Hence, the extent to which an earnings announcement will provide 'useful' information to market participants should be a function not only of the nature of information released but

* Corresponding author. Tel.: +1 405 325 3721; fax: +1 405 325 7555.
E-mail addresses: h.ince@gyte.edu.tr (H. Ince), ttrafalis@ou.edu (T.B. Trafalis).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.10.008

also when it is released. Firms tend to release good news earnings reports earlier than those containing bad news (Chambers & Penman, 1984; Kross & Schroeder, 1984). If this is the case, then the longer a firm goes without reporting its earnings, the less positive (or, alternatively, more negative) will investors expect the ultimate earnings news to be. This implies that a firm should experience negative stock returns from the end of its fiscal quarter up until the time of its earnings announcement. When the earnings are finally released by the manager, the stock price should rise.

Based on these studies, we propose a classification model with two classes; 'buy a certain stock whose earning/eps is higher than some threshold' and 'do not buy a certain stock if its earning is less than some certain threshold'. The threshold value is subjective and depends on the investor. As far as we know, there have been no studies that propose a classification model by using the earning announcements and anomalies. Data mining techniques such as SVMs, MPM and multilayer perceptron (MLP) would be good candidates addressing this problem instead of using classical statistical techniques.

Data mining is the process of discovering and analyzing hidden patterns in data sets. It has wide application areas from engineering to commerce. Most widely used data mining techniques in finance are artificial neural networks (ANNs), SVMs and their variations. Bankruptcy prediction, portfolio management (choosing individual stocks for portfolio), option pricing, and forecasting indices or individual stock prices are typical application areas in finance (Chen, Leung, & Daouk, 2003; Dourra & Siy, 2002; Galindo, 1998; Hutchinson, Lo, & Poggio, 1994; Trafalis & Ince, 2000).

The classification problem is to learn a discriminating hyperplane or hypersurface, which separates two classes from examples by minimizing the generalization error (Vapnik, 1995). MPM attempts to control misclassification probabilities for two-class classification problems. Specifically it minimizes the worst case (maximum) probability of misclassification of future data points under all possible choices of class densities with a given mean and covariance matrix, which is positive definite. MPM minimizes directly an upper bound on the generalization errors (Lanckriet, El Ghaoui, Bhattacharyya, & Jordan, 2002). In contrast to MPM, SVMs achieve low generalization error by maximizing an associate quantity termed margin that describes the distance between two classes.

In this study, we will provide a framework for short-term portfolio selection problem by using MPM and SVMs. This can be done by analyzing the behavior of the stocks around their earning announcement. We believe that some variables such as volume, difference between expected eps and actual eps, and expectation have big effect on return. Several studies have shown that the stock market is inefficient during the earning announcement in terms of individual stocks.

The paper is organized as follows: in Sections 2 and 3 we present the basics of the MPM and SVM techniques, respectively; in Section 4, a short-term portfolio classification model is explained; computational results are given in Section 5; and finally, Section 6 concludes the paper.

2. Minimax probability machine

The MPM attempts to control misclassification probabilities for two-class (binary) classification problems. It minimizes the worst case (maximum) probability of misclassification of future data points under all possible choices of class densities with a given mean and covariance matrix (Lanckriet et al., 2002). MPM uses the following theorem, by Popescu & Bertsimas (2001):

$$\sup_{\mathbf{y}\sim(\bar{\mathbf{y}},\sum y)} \Pr\{\mathbf{y}\in S\} = \frac{1}{1+d^2}, \text{ and, } d^2 = \inf_{\mathbf{y}\in S}(\mathbf{y}-\bar{\mathbf{y}})^{\mathrm{T}}\sum_{\mathbf{y}}^{-1}(\mathbf{y}-\bar{\mathbf{y}}), \tag{1}$$

where $\mathbf{y}$ is a random vector, S is a given convex set, and supremum is taken over all distributions for $\mathbf{y}$ having mean $\bar{\mathbf{y}}$ and covariance matrix Σ_y, assumed to be positive definite.

Let $\mathbf{x}$ and $\mathbf{y}$ denote random vectors in a binary classification problem with means and covariance matrices given by $(\bar{\mathbf{x}},\sum_x)$ and$(\bar{\mathbf{y}},\sum_y)$. Note that $\mathbf{x}$ and $\mathbf{y}$ also denote the two classes. MPM seeks a hyperplane $H(\mathbf{w},b)=\{\mathbf{z}|\mathbf{w}^T\mathbf{z}=b\}$ where $w\in\Re^n\backslash\{0\}$ and $b\in\Re$, which separates the two classes of points with maximal probability with respect to all distributions having these means and covariance matrices. It minimizes the generalization error by finding the hyperplane for which the misclassification probabilities $\Pr(\mathbf{w}^T\mathbf{x}\le b)$ and $\Pr(\mathbf{w}^T\mathbf{y}\le b)$ are minimized. The optimization problem becomes

$$\max_{\alpha,\mathbf{w}\neq 0,b} \alpha \tag{2}$$

s.t.

$$\inf_{\mathbf{x}\sim(\bar{\mathbf{x}},\sum_x)} \Pr(\mathbf{w}^{\mathrm{T}}\mathbf{x}\ge b)\ge\alpha,$$

$$\inf_{\mathbf{y}\sim(\bar{\mathbf{y}},\sum_y)} \Pr\left(\mathbf{w}^{\mathrm{T}}\mathbf{y}\le b\right)\ge\alpha$$

The quantity $(1-\alpha)$ can be interpreted as an upper bound on the generalization error. Furthermore, this quantity is the worst-case (maximum) misclassification probability. Infimum in both constraints are computed via a theorem stated in Popescu & Bertsimas (2001) given in (1). Problem (2) can be restated as

$$\max_{\kappa,\mathbf{w}\neq 0} \kappa \tag{3}$$

s.t.

$$\mathbf{w}^T(\bar{\mathbf{x}}-\bar{\mathbf{y}})\ge\kappa\left(\sqrt{\mathbf{w}^{\mathrm{T}}\sum_{\mathbf{x}}\mathbf{w}}+\sqrt{\mathbf{w}^{\mathrm{T}}\sum_{\mathbf{y}}\mathbf{w}}\right)$$

The problem (3) can be simplified as (Lanckriet et al., 2002):

$$\min_{\mathbf{w}} \|\textstyle\sum_{\mathbf{x}}^{1/2}\mathbf{w}\|_2+\|\sum_{\mathbf{y}}^{1/2}\mathbf{w}\|_2 \tag{4}$$

s.t.

$$\mathbf{w}^{\mathrm{T}}(\bar{\mathbf{x}}-\bar{\mathbf{y}})=1$$