Available online at www.sciencedirect.com

SCIENCE DIRECT®

Computers & Operations Research 32 (2005) 2513–2522

computers & operations research

www.elsevier.com/locate/dsw

Forecasting stock market movement direction with support vector machine

Wei Huang[a,b], Yoshiteru Nakamori[a], Shou-Yang Wang[b,*,1]

[a]School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

[b]Institute of Systems Science, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100080, China

Accepted 29 March 2004

Abstract

Support vector machine (SVM) is a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution. In this paper, we investigate the predictability of financial movement direction with SVM by forecasting the weekly movement direction of NIKKEI 225 index. To evaluate the forecasting ability of SVM, we compare its performance with those of Linear Discriminant Analysis, Quadratic Discriminant Analysis and Elman Backpropagation Neural Networks. The experiment results show that SVM outperforms the other classification methods. Further, we propose a combining model by integrating SVM with the other classification methods. The combining model performs best among all the forecasting methods.

Keywords: Support vector machine; Forecasting; Multivariate classification

1. Introduction

The financial market is a complex, evolutionary, and non-linear dynamical system [1]. The field of financial forecasting is characterized by data intensity, noise, non-stationary, unstructured nature, high degree of uncertainty, and hidden relationships [2]. Many factors interact in finance including political events, general economic conditions, and traders' expectations. Therefore, predicting finance market price movements is quite difficult. Increasingly, according to academic investigations, movements

* Corresponding author. Tel.: +86-10-62651381; fax: +86-10-62568364.
E-mail address: swang@mail.iss.ac.cn (S.-Y. Wang).

[1] This author is also with University of Tsukuba in Japan and Hunan University in China.

doi:10.1016/j.cor.2004.03.016

in market prices are not random. Rather, they behave in a highly non-linear, dynamic manner. The standard random walk assumption of futures prices may merely be a veil of randomness that shrouds a noisy non-linear process [3–5].

Support vector machine (SVM) is a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution [6–8]. Established on the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error, SVM is shown to be very resistant to the over-fitting problem, eventually achieving a high generalization performance. Another key property of SVM is that training SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of SVM is always unique and globally optimal, unlike neural networks training which requires nonlinear optimization with the danger of getting stuck at local minima.

Some applications of SVM to financial forecasting problems have been reported recently [9–13]. In most cases, the degree of accuracy and the acceptability of certain forecasts are measured by the estimates' deviations from the observed values. For the practitioners in financial market, forecasting methods based on minimizing forecast error may not be adequate to meet their objectives. In other words, trading driven by a certain forecast with a small forecast error may not be as profitable as trading guided by an accurate prediction of the direction of movement.

The main goal of this study is to explore the predictability of financial market movement direction with SVM. The remainder of this paper is organized as follows. In Section 2, we introduce the basic theory of SVM. Section 3 gives the experiment scheme. The experiment results are shown in Section 4. Some conclusions are drawn in Section 5.

2. Theory of SVM in classification

In this section, we present a basic theory of the support vector machine model. For a detailed introduction to the subject, please refer to [14,15]. Let D be the smallest radius of the sphere that contains the data (example vectors). The points on either side of the separating hyperplane have distances to the hyperplane. The smallest distance is called the margin of separation. The hyperplane is called optimal separating hyperplane (OSH), if the margin is maximized. Let q be the margin of the optimal hyperplane. The points that are distance q away from the OSH are called the support vectors.

Consider the problem of separating the set of training vector belonging to two separate classes, $G = \{(x_i, y_i), i = 1, 2, \dots, N\}$ with a hyperplane $w^{\mathrm{T}}\varphi(x) + b = 0$ ($x_i \in R^n$ is the ith input vector, $y_i \in \{-1, 1\}$ is known binary target), the original SVM classifier satisfies the following conditions:

$$w^{\mathrm{T}}\varphi(x_i) + b \geqslant 1 \qquad \text{if} \quad y_i = 1, \tag{1}$$

$$w^{\mathrm{T}}\varphi(x_i) + b \leqslant -1 \quad \text{if} \quad y_i = -1, \tag{2}$$

or equivalently,

$$y_i[w^{\mathrm{T}}\varphi(x_i) + b] \geqslant 1 \quad i = 1, 2, \dots, N, \tag{3}$$

where $\varphi : R^n \to R^m$ is the feature map mapping the input space to a usually high dimensional feature space where the data points become linearly separable.