Forecasting Direction of the S&P500 Movement Using Wavelet Transform and Support Vector Machines

Salim Lahmiri, Department of Computer Science, University of Quebec at Montreal, Montreal, QC, Canada, & ESCA School of Management, Casablanca, Morocco

ABSTRACT

Using the wavelet analysis for low-frequency time series extraction, we conduct out-of-sample predictions of the S&P500 price index future trend (up and down). The support vector machines (SVMs) with different kernels and parameters are used as the baseline forecasting model. The simulation results reveal that the SVMs with wavelet analysis approach outperform the SVMs with macroeconomic variables or technical indicators as predictive variables. As a result, we conclude that the wavelet transform is appropriate to capture the S&P500 trend dynamics.

Keywords: Classification, Financial Time Series, Forecasting, Stock Market Trend, Support Vector Machines, Wavelet Transform

INTRODUCTION

Recently, a large attention has been given to the problem of financial data modeling and decision making (Xidonas & Psarras, 2008; Davalos et al., 2009; Sun, 2010; Joseph & Mazouz, 2010; Hammami & Boujelbene, 2012; Lai & Joseph, 2012; Strang, 2012). One of the most attractive fields of study is the stock returns prediction. Indeed, stock market forecasting is a difficult and challenging task since stock data are noisy non-stationary, and chaotic. In particular, several factors affect the behaviour of the stock market including macroeconomic conditions, political events, psychology of investors, and traders’ expectations. Nevertheless, there have been many studies in the field of stock market forecasting using soft computing techniques in the last decade (refer to the survey by Atsalakis and Valavanis (2009)). More recent studies include bacterial foraging optimization, adaptive bacterial foraging optimization, genetic algorithms and particle swarm optimization (Majhi et al., 2009), bacterial chemotaxis optimization and artificial neural networks (Zhang and Wu, 2009), probabilistic neural network, rough set and C4.5 (Cheng et al., 2010b), Markov chain and fuzzy logic (Wang et al., 2010), rough sets theory and genetic algorithms (Cheng et al., 2010a), artificial neural networks (Wang

DOI: 10.4018/jsds.2013010105
et al., 2011), support vector machines (Yeh et al., 2011), fuzzy logic and artificial neural networks (Chakravarty & Dash, 2012; Kumar, 2012), and partially connected networks (Chang et al., 2012). The advantage of using soft computing techniques is that they can handle the uncertain, chaotic, and nonlinear structure of the stock markets (Majhi et al., 2009; Chakravarty & Dash, 2012). Furthermore, they are not based on the assumption of linearity of the underlying model and normality distribution of the variables. Indeed, these assumptions may not be satisfied in modeling the stock price movements (Wang et al., 2011). Support vector machines are the most successfully used soft computing technique for modeling and forecasting financial time series due to the outstanding performance in classification and regression problems (Wen et al., 2010; Ismael et al., 2010; Kara et al., 2011; Ni et al., 2011). Indeed, SVM which is introduced by Vapnik (1995) is based on the structural risk minimization principle that considers both the training error and the capacity of optimal generalization, has a global optimum, and is effective and powerful as a discriminant function (Vapnik, 1995; Cristianini & Shawe-Taylor, 2000; Kecman, 2001; Sun, 2010). To predict stock market movements, most of works used as predictive inputs to soft computing systems either macroeconomic variables, technical analysis indicators or simply historical values (Atsalakis & Valavanis, 2009). However, recent studies have given attention to the wavelet analysis (Daubechies, 1992; Chui, 2002) to transform stock market data to a time-frequency feature space suitable for financial modeling and forecasting using soft computing systems (Huang & Wu, 2010; Wang et al., 2011; Huang, 2011, Hsieh et al., 2011). The wavelet characterizes the coarse structure of data to identify the long-run trend. Because of these attractive features, it was concluded that the wavelet analysis is effective in forecasting stock market future price index (Huang & Wu, 2010; Wang et al., 2011; Huang, 2011, Hsieh et al., 2011).

The purpose of this study is to apply the wavelet analysis in the problem of forecasting future stock market trends (ups and downs). The wavelet transform is used to de-noise the stock market time series to extract low frequency components to be fed to support vector machines. The SVMs with two different kernels; namely polynomial and radial basis function; are used to perform classification task; for example forecasting price index ups and downs. The second purpose is to compare the performance of the SVM depending on the type of the inputs: wavelet components, macroeconomic variables or technical indicators. The goal is to check the effectiveness of the wavelet analysis approach against conventional approaches. Because of the high dimensionality and possible redundancy of the input data, the most efficient predictive inputs (macroeconomic variables and technical indicators) are identified based on entropy measure. Then, the selected inputs will be fed to the classifier (SVM) to perform final forecasting/classification.

The reminder of the paper is as follows. First, our methodology is described. Then, the data and simulations results are provided. Finally, we conclude the paper.

**METHODOLOGY**

In this study, three experiments will be conducted to forecast stock market future moves $y$ (up = +1 for ups and -1 for downs). In the first one, support vector machines use low frequency components obtained from the wavelet transform of the stock market price index times as predictive inputs. In the second experiment, entropy-based selected macroeconomic variables are fed to SVM. In the third experiment, SVM uses entropy-based selected technical indicators as predictive variables. The following
Fuzzy Adaptive Controller for Uncertain Multivariable Nonlinear Systems with Both Sector Nonlinearities and Dead-Zones
www.igi-global.com/chapter/fuzzy-adaptive-controller-for-uncertain-multivariable-nonlinear-systems-with-both-sector-nonlinearities-and-dead-zones/176797?camid=4v1a

A Predictive Analytic Model for Maternal Morbidity
www.igi-global.com/chapter/a-predictive-analytic-model-for-maternal-morbidity/138642?camid=4v1a