Chapter 10
Practical Machine Learning in Financial Market Trend Prediction

Salim Lahmiri
University of Quebec at Montreal, Canada & ESCA School of Management, Morocco

ABSTRACT
Using the wavelet analysis for low-frequency time series extraction, the authors in this chapter conduct out-of-sample predictions of the S&P500 price index future trend (up and down) following two trading strategies. In particular, the goal is to separately predict an increase or decrease of stock market by 0.5%. Indeed, predicting market increases by 0.5% is suitable to active portfolio managers, whilst predicting its decreases by 0.5% is suitable to risk-averse portfolio managers to limit losses. The Support Vector Machine (SVM) with polynomial kernel is used as the baseline forecasting model. Its performance is respectively compared to that of the Probabilistic Neural Networks (PNN) and the well known k-Nearest Neighbour (k-NN) algorithm, which is a statistical classifier. The simulation results reveal that the predictive system based on the SVM with wavelet analysis coefficients as inputs outperforms all the other systems. The achieved accuracy is 98.13%. As a result, it is concluded that the wavelet transform and SVM as an integrated system are appropriate to capture the S&P500 price changes by more or less than 0.5%.

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 macro-economic 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

DOI: 10.4018/978-1-4666-5958-2.ch010
Practical Machine Learning in Financial Market Trend Prediction

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 & 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 et al., 2011), probabilistic neural networks, 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 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 & Shave-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 analysis which is widely used in pure science and engineering problems is a powerful tool to decompose a time series into series of different timescales. In particular, it decomposes a given data into high and low frequency components. At high frequency (shorter time intervals), the wavelets are able to capture discontinuities, ruptures and singularities in the original data. At low frequency (longer time intervals), 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 three different machine learning classifiers; namely the support vector machine (SVM), the probabilistic neural networks (PNN), and the k-nearest neighbour algorithm (k-NN). Introduced by Vapnik (1995), the SVM was developed based on the theory of structural risk minimization for binary classification problems. The SVM seeks to implement an optimal marginal classifier that minimizes the structural risk. Its main advantage is its ability to reach global optimum. The PNN (Specht, 1990) provides a general solution to pattern classification problems based on Bayesian theory. It is able to classify a new sample with the maximum probability of success given a large training set using prior knowledge. Finally, the main advantage of $k$-NN (Cover & Hart, 1967) is to use the data directly for classification without the need of an explicit model.

The classifiers are used to perform classification task; for example forecasting price index ups and downs. Their respective performances are compared depending on the type of the inputs: