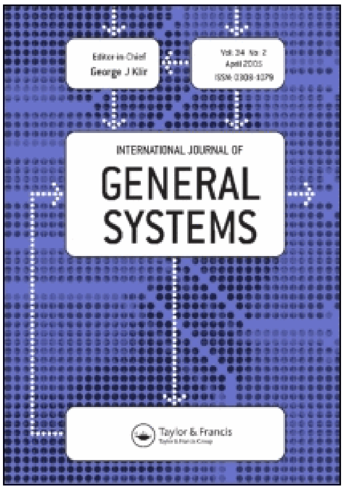

International Journal of General Systems
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713642931

Short term forecasting with support vector machines and application to stock price prediction

Huseyin Ince [a]; Theodore B. Trafalis [b]
[a] Faculty of Business Administration, Gebze Institute of Technology, Kocaeli, Turkey [b] School of Industrial Engineering, University of Oklahoma, OK, USA

To cite this Article Ince, Huseyin and Trafalis, Theodore B.(2008) 'Short term forecasting with support vector machines and application to stock price prediction', International Journal of General Systems, 37: 6, 677 — 687
To link to this Article: DOI: 10.1080/03081070601068595
URL: http://dx.doi.org/10.1080/03081070601068595

International Journal of General Systems
Vol. 37, No. 6, December 2008, 677–687

Short term forecasting with support vector machines and application to stock price prediction

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

[a]*Faculty of Business Administration, Gebze Institute of Technology, Kocaeli, Turkey;* [b]*School of Industrial Engineering, University of Oklahoma, OK, USA*

(Received 20 December 2004; revised 2 October 2006; final version received 2 May 2007)

Forecasting a stock price movement is one of the most difficult problems in finance. The reason is that financial time series are complex, non stationary. Furthermore, it is also very difficult to predict this movement with parametric models. Instead of parametric models, we propose two techniques, which are data driven and non parametric. Based on the idea that excess returns would be possible with publicly available information, we developed two models in order to forecast the short term price movements by using technical indicators. Our assumption is that the future value of a stock price depends on the financial indicators although there is no parametric model to explain this relationship. This relationship comes from the technical analysis. Comparison shows that support vector regression (SVR) out performs the multi layer perceptron (MLP) networks for a short term prediction in terms of the mean square error. If the risk premium is used as a comparison criterion, then the SVR technique is as good as the MLP method or better.

Keywords: financial time series; prediction; support vector regression; technical indicators; multilayer perceptron

1. Introduction

One of the goals of financial methods is asset evaluation which means determining the market price of an asset, predicting what this price will be in the future, and how it will move with other indicators. The behavior of an asset can be analyzed by using technical tools, parametric pricing methods or combination of these methods. Prediction of financial time series is one of the most challenging applications. Since the financial market is a complex, non stationary and deterministically chaotic system, it is very difficult to forecast by using deterministic (parametric) techniques because of the assumptions behind the parametric techniques.

The support vector regression (SVR) algorithm that was developed by Vapnik (1995) is based on statistical learning theory. SVR is based on the structural risk minimization (SRM) principle which seeks to minimize an upper bound of the generalization error rather than minimizing the empirical error. The generalization error is bounded by the sum of the empirical error and confidence interval term that depends on Vapnik–Chervonenkis (VC) dimension. In the case of regression, the goal is to construct

*Corresponding author. Email: ttrafalis@ou.edu

ISSN 0308-1079 print/ISSN 1563-5104 online

DOI: 10.1080/03081070601068595
http://www.informaworld.com

a hyperplane that lies "close" to as many of the data points as possible (Cortes and Vapnik 1995, Pontil and Verri 1997). This determines the trend line of training data, hence the name SVR.

The tendency in the field of financial forecasting is to use state variables which are fundamental or macro economic variables. Fundamental analysis uses some factors to predict the excess returns in the market. Some of these factors are price-to-earning, price-to-book, and earning momentum which is measured as the weighted average of quarterly growth rates in earning per share (EPS) over the past year. Several researchers have used these variables in order to create a winning portfolio. Gold and Lebowitz (1999) used these variables to create a computerized screening model.

On the other hand, technical analysis, also known as charting, is used to predict future stock price movements by analyzing the past sequence of stock prices. It is based on some technical indicators and dismisses fundamental variables, economic environment, industry trend and political events (Dourra and Siy 2002). These indicators help the investors to buy or sell the underlying stock based on some rules created from the indicators. Our goal is to try to understand the relationship between the stock price and these indicators. These indicators are computed by using daily stock price and volume overtime. We will investigate the prediction of individual stock prices by applying SVR and neural networks methods.

Section 2 explains the theory of SVR. In Section 3, multilayer perceptron (MLP) networks will be discussed. Technical analysis is given in Section 4. Experiments and comparison of the SVR and MLP will be discussed in Section 5. Section 6 concludes the paper.

2. Support vector regression (SVR)

Support vector machines (SVMs) can be implemented in two different ways; specifically, regression analysis and pattern classification. Since SVMs in regression is investigated in this paper, the reader is referred to Burges (1998) and Cristiannini and Shawe-Taylor (2000) for further information on the subject, in particular on SVMs in classification. For the case of regression, a set of data points $\{(x_i, y_i)\}_{i=1}^{l}$ is given (x_i is the input vector, y_i is the desired (target) value and l is the total number of examples) drawn independently and identically from an unknown function. SVR approximates the function with three distinct characteristics (Tay and Cao 2002): (i) SVR estimates the regressor in a set of linear functions, (ii) SVR defines the regression estimation as the problem of risk minimization with respect to ε-insensitive loss function and (iii) SVR minimizes the risk based on the SRM principle.

In the ε-insensitive SVR, our goal is to find a function $f(x)$ that has ε deviation from the actually obtained target y_i for all training data and at the same time is as flat as possible. Suppose $f(x)$ takes the following form in input space,

$$f(x) = w \cdot x + b, \quad w \in X, \quad b \in \Re \tag{1}$$

If we have a w with a small norm, then we can say that f is flat. One way to do this is to minimize $\|w\|^2$ using the euclidean norm (Smola and Schölkopf 2004) subject to the constraints of the so called ε-insensitive band. One has to solve this problem in order to obtain an ε-insensitive SVR solution. Usually, we need to allow for some errors. We introduce slack variables ξ_i, ξ_i^* to cope with this situation. This case is called soft margin formulation. Specifically, we solve the following problem:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{v-1}^{\ell}\left(\xi_i + \xi_i^*\right) \quad \text{Subject to} \quad \mathbf{y}_i - wx_i - b \le \varepsilon + \xi_i$$
$$wx_i + b - y_i \le \varepsilon + \xi_i^*, \quad i = 1, \ldots, \ell \quad \xi_i, \xi_i^* \ge 0, \quad C \ge 0 \tag{2}$$

where C determines the trade-off between the flatness of the $f(x)$ and the amount up to which deviations larger than ε are tolerated.

The Lagrangian function will help us to formulate the dual problem, which will give us a quadratic programming (QP) problem formulation. By using the dual problem formulation, we can decrease the variables and the size of the problem becomes smaller.

In the non-linear case, we need to map the input space into the feature space through a feature map ϕ and try to find a regression hyperplane in the feature space. A nonlinear regression problem in input space becomes a linear regression problem in feature space using the concept of the kernel method (Figure 1). The dual problem allows the problem data, x_i and x_j, to be expressed in terms of inner products. Hence, in order to allow for nonlinear regression analysis, the only change that has to be made is rewriting the dot product in the feature space as

$$k(x_i, x_j) = \phi(x_i)\cdot\phi(x_j) \tag{3}$$

Then, we need to solve the following problem (Vapnik 1995):

$$\max -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\left(\lambda_i - \lambda_i^*\right)\left(\lambda_j - \lambda_j^*\right)K(\mathbf{x}_i, \mathbf{x}_j) - \varepsilon\sum_{i=1}^{l}\left(\lambda_i + \lambda_i^*\right) + \sum_{i=1}^{l} y_i\left(\lambda_i - \lambda_i^*\right)$$
$$\text{Subject to} \ \sum\left(\lambda_i - \lambda_i^*\right) = 0 \quad \lambda_i, \lambda_i^* \in (0, C) \tag{4}$$

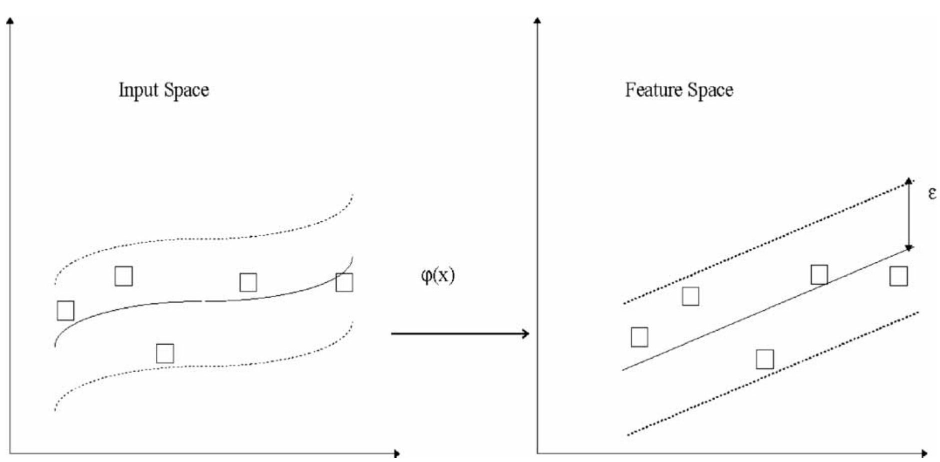

Figure 1. Kernel method for support regression problem (Smola and Schölkopf 2004).

At the optimal solution, we obtain

$$w^* = \sum_{i=1}^{l} \left(\lambda_i - \lambda_i^*\right)\phi(x_i), \quad \text{and} \quad f(x) = \sum_{i=1}^{l} \left(\lambda_i - \lambda_i^*\right)K(x_i, x) + b \tag{5}$$

SVR results in a rather simple QP problem. Since the problem is convex, there are no local maxima and several optimization algorithms will be able to find the optimal solution. The name support vector is derived from those points in the input space, which touch ("support") the decision function. More formally, if one of the two Lagrangian multipliers (λ_j or λ_j^*) is nonzero, then, according to complementary slackness the corresponding data sample ($x_{j,}$ y_j) fulfills one of the primal constraints in (2) as an equality. Thus, all samples that yield either a strictly positive λ_j or a strictly positive λ_j^* are support vectors.

Any symmetric positive semi-definite function, which satisfies Mercer's conditions can be used as a kernel function in the SVR context (Burges 1998, Smola and Schölkopf 2004). Polynomial and radial basis funtion (RBF) kernel functions are very common. An RBF kernel function is defined as

$$K(x, y) = \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \tag{6}$$

Usually, we have more than one kernel to map the input space into feature space. The question is which kernel functions provide good generalization for a particular problem. We could not say that one kernel outperforms the others. Therefore, one has to use more than one kernel function for a particular problem. Some validation techniques such as bootstrapping, and cross-validation can be used to determine a good kernel (Smola and Schölkopf 2004).

In order to solve the SVR optimization problem in Equation (4), one needs to solve a QP problem. Despite its many advantages, the disadvantage is that the size of the matrix of the QP problem is directly proportional to the number of training points so that standard QP packages cannot be used even for moderately large data sets. Recently, several algorithms have been developed to solve this problem. For example, one widely used algorithm for solving the SVR QP problem is the Advanced Working Set Algorithm (Smola and Schölkopf 2004). The implementation of this method on the SVR case with a convex cost function can be found in Collobert and Bengio (2001). Another algorithm, sequential minimum optimization (SMO), is proposed by Platt (1999) that uses the chunking to the extreme by iteratively selecting subsets only of size 2 and optimizing the target function with respect to them. The optimization subproblem can be solved analytically without any quadratic optimization routine. This algorithm is developed for the pattern recognition case, but one can use this algorithm for the SVR case with small changes. Derivation of SMO for SVR can be found in Platt (1999). Suykens and Vandewalle (1999) proposed a modified version of SVM for classification called least square SVM (LSSVM) which resulted in a set of linear equations instead of a QP problem. Other training algortihms can be found in Cortes and Vapnik (1995), Osuna *et al.* (1997), Lee and Mangasarian (2001) and Trafalis and Ince (2002).

3. Artificial neural networks

Neural networks have been used in function estimation such as stock price prediction, option price modeling and currency exchange rate estimation. An artificial neural network

(ANN) is a learning machine that is designed to model the way in which the brain performs particular tasks. A common ANN is the multi-layer perceptron (MLP) that is constituted of a set of interconnected basic processing units (neurons) organized in layers (Figure 2). Each neuron produces its output by taking a linear combination of the input signals and transforming it using a function called the activation function. The weights of this linear combination are associated with the numeric connections (synaptic weights) linking the neuron with all of the neurons belonging to the upper layer.

The output of a neuron as a function of the input signals can thus be written as

$$y_j = f\left(\sum_{i=1}^{m} w_{ji} x_i - b\right), \quad (7)$$

where y_j is the output of the generic neuron belonging to layer j; x_i are the input signals to the neuron; w_{ji} is the synaptic weight associated with the connection between the generic neurons belonging to layers j and i, respectively; b is the bias term (another neuron weight); f is the activation function.

The most widely used weight updating algorithm for the training of MLP networks is the so called "error backpropagation" (EBP). The basic idea of the back propagation learning algorithm consists of the repeated application of the rule for computing the influence of each weight in the network with respect to an arbitrary error function. More details of the EBP algorithm can be found in Haykin (1994).

4. Technical indicators of financial time series

According to the efficient market hypothesis, past price movements in a competitively traded financial market do not help predicting future prices. However many recent articles question the efficient market hypothesis and support the notion that stock market excess returns can be predicted by publicly available information (Fama and French 1995,

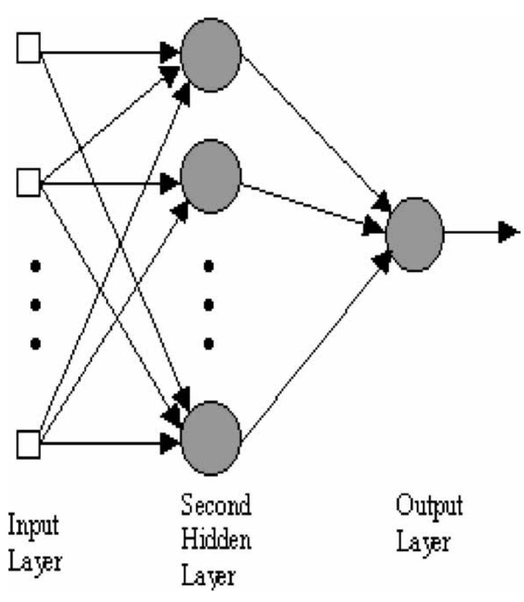

Figure 2. An MLP ANN architecture.

Pesaran and Timmerman 1995, Gencay 1996, Gold and Lebowitz 1999). Although it is still commonly believed that the US stock market is semi-strong efficient, financial research indicates that security prices do not reflect all publicly available information (Walker and Hatfield 1996). Because of this, investor uses computerized screening programs that take advantage of advances in information technology to efficiently select stock portfolios.

Technical analysis is essentially search for predictable patterns in stock prices. Technical trading rules aim in general to identify the initiation of new trends (Tian *et al.* 2002). Some of the rules include filter rules (buy when the price rises by a given proportion above recent price) and moving average crossovers (buy when a shorter moving average penetrates a longer moving average from below). For each rule, the analyst chooses the time horizon over which troughs and peaks are to be identified and moving averages calculated, as well as the threshold before a decision made (Beechey *et al.* 2000, Tian *et al.* 2002). Although, the majority of professional traders use technical analysis, most academics, until recently, had not recognized the validity of these methods.

Most of the technical trading rules are simple and fairly inexpensive to implement. One would therefore not expect such rules to generate excess profits in an efficient market. Most academic studies of technical analysis concluded that technical analysis is not useful. In the last few years, increasing evidence that a relatively simple set of technical trading rules posses significant forecast power for equity returns has renewed interest in technical analysis (Brock *et al.* 1992, Tian *et al.* 2002).

Following technical indicators have been used to explain the dynamics of a stock price. We will consider six indicators as inputs in the model: exponential moving average (EMA), volume (V), relative strength index (RSI), Bollinger bands (BB), moving average convergence divergence (MACD) and chaikin money flow (CMF). They can be used for short term and long term investment strategies. Usually, three or more indicators can be used together for identifying the trend of the market.

5. Experiments

There have been several papers discussing the use of non parametric techniques for financial forecasting. For example neural networks, time series ARIMA models have been applied to stock index and exchange rate forecasting (Tsaih 1999, Leigh *et al.* 2002). Similarly, SVR has been applied to stock price forecasting and option price prediction (Trafalis *et al.* 2003).

The combination of technical indicators and fundamentals are used for effective and efficient portfolio management. Portfolio managers use different techniques to identify which stocks they have to put in their portfolio and which ones they should sell. Generally, they focus on long term portfolio management. Here, we will focus on short term portfolio management. In this paper, short term is defined as one or two weeks.

We have used the daily stock prices of ten companies traded in NASDAQ. These companies are selected from different sector. The dataset contains daily stock prices from January 1, 2000 to January 31, 2006. The training set contains 1300 observations and the testing/validation set consists of 200 examples.

We propose two simple models. In the first model, we assume that the current stock price s_t depends on the previous EMA, Volume, RSI, BB, MACD and CMF. Specifically,

$$S_t = f(\mathrm{EMA}_{t-1}, V_{t-1}, \mathrm{RSI}_{t-1}, \mathrm{BB}_{t-1}, \mathrm{MACD}_{t-1}, \mathrm{CMF}_{t-1}), \tag{8}$$

The second model is

$$S_t = f(\mathrm{EMA}_{t-1}, \mathrm{EMA}_{t-2}, V_{t-1}, V_{t-2}, \mathrm{RSI}_{t-1}, \mathrm{RSI}_{t-2}, \mathrm{RSI}_{t-3}, \mathrm{BB}_{t-1}, \mathrm{BB}_{t-2}, \mathrm{MACD}_{t-1}, \mathrm{MACD}_{t-2}, \mathrm{CMF}_{t-1}, \mathrm{CMF}_{t-2}). \quad (9)$$

Since most of these indicators are computed by using daily prices of underlying stock, they are highly correlated with each other. Therefore, it should be wise not to use any parametric techniques such as multiple regression models. Instead of parametric models, non parametric data driven model, SVR and MLP will be applied with different settings. Furthermore, time series model, namely auto regressive integrated moving average (ARIMA) will be used to compare performance of SVR and MLP with a pure time series model. Our goal is to develop stock prediction model based on the relationship between stock price and certain technical indicators. SVR, MLP and ARIMA methods are compared with each other in order to identify the best method for each stock prediction model given in Equations (8) and (9).

Since SVR and MLP methods are data driven methods and have some free parameters, we have employed 10-fold cross-validation techniques to determine those parameters. Radial basis kernel function (RBF) is chosen for SVR. After 10-fold cross-validation, we determine the spread (σ) and trade-off (C) values for each stock shown in Table 1. Also, Table 2 shows the best network structure for each stock.

After determining the free parameters and the network architecture for each dataset, performances of the proposed techniques are compared in terms of risk (profit) and root mean square error. In order to compute the risk associated with each techniques, we have used to trading strategies. These strategies can be defined as follows (Yao and Tan 2000):

Strategy 1: If $(\hat{x}_{t+1} - \hat{x}_t) > 0$ then buy the underlying stock else sell.
Strategy 2: If $(\hat{x}_{t+1} - x_t) > 0$ then buy the underlying stock else sell.

x_t is the actual stock price at time t, and $\hat{x}_{t+1}$ is the prediction of the stock price by SVM, MLP or ARIMA models. We assume there is no transaction cost. The risk associated with these trading strategies is computed by using the Sharpe ratio which is defined as the excess return divided by a risk measure (Sharpe 1966). If this ratio is higher,

Table 1. Ten-fold cross-validation results for model 1 and 2 to determine the free parameters (C and σ) of the RBF kernel function used in SVR.

	Model 1		Model 2	
Stock name	Trade-off (C)	Spread (σ)	Trade-off (C)	Spread (σ)
SONS	3812	14,063	3988	14,507
MSFT	1464	13,924	1128	14,733
YHOO	2352	10,951	34,657	11,745
INTC	2691	12,081	2294	11,217
GEOI	2420	19,455	1737	11,740
ABLE	1370	16,603	989	11,298
CSCO	3008	14,525	16,674	11,998
AMZN	1445	15,635	1227	12,572
MCEL	3328	11,542	7826	12,162
FCEL	1290	14,021	1484	11,298

Table 2. Ten-fold cross-validation results for model 1 and 2 to determine the architecture of MLP networks.

Stock name	Model 1	Model 2
SONS	6-36-1	13-6-1
MSFT	6-13-1	13-13-1
YHOO	6-13-1	13-13-1
INTC	6-17-1	13-13-1
GEOI	6-22-1	13-63-1
ABLE	6-25-1	13-14-1
CSCO	6-22-1	13-14-1
AMZN	6-15-1	13-27-1
MCEL	6-25-1	13-22-1
FCEL	6-26-1	13-54-1

then the risk adjusted return is also higher. It can be computed as

$$s = \frac{\mu - c}{\sigma}$$

where μ is the expected return from the trading strategy, σ is standard deviation of the return and c is one plus a risk-free rate. Tables 3 and 4 show the Sharpe ratio of the models 1 and 2 for two strategies. SVR methods have the best risk adjusted return for strategy 1 (5 out of 10 stock) and strategy 2 (6 out of 10 stock), which means that the SVR method outperforms the other two techniques for model 1.

We have similar findings for model 2 that is given in Equation (9). SVR has better results for strategy 1, and it is as good as MLP for strategy 2 (Table 4). Different trading strategies produce different results. Therefore, we cannot say that one tehnique is better than the other technique.

RMSE is used to evaluate the predictive power of the forecasting techniques. Table 5 shows the comparison of the SVR, MLP and ARIMA techniques in terms of the RMSE statistics for testind period. For each forecasting model, SVR outperforms the MLP and

Table 3. Risk associated with trading strategy 1 and 2 for model 1. Risk premium is computed by using Sharpe ratio for testing set. Bold ones are the best model for each stock.

	Strategy 1			Strategy 2		
Stock name	SVR	MLP	ARIMA	SVR	MLP	ARIMA
SONS	**1.5863**	1.2913	−2.1966	**0.6853**	−3.6659	0.5146
MSFT	0.5469	**0.8146**	0.4745	−4.6177	−2.5208	**1.3428**
YHOO	−**1.0361**	−7.1008	−4.9580	**1.8578**	0.1697	0.4888
INTC	−4.5305	−3.8364	−**1.2412**	−5.8006	−4.6658	−**3.0721**
GEOI	−2.1392	0.3554	**1.6952**	−**0.5294**	−1.7321	−0.6800
ABLE	**2.2134**	0.0660	−1.7807	−**1.9854**	−3.0388	−3.9627
CSCO	−**2.4650**	−3.6338	−3.9196	**5.4112**	−4.7368	−5.3831
AMZN	2.0148	**2.3675**	1.6511	−1.5000	−1.3339	**0.3514**
MCEL	−3.0016	−**2.6625**	−4.4959	**1.4398**	0.3533	−0.5962
FCEL	**2.0134**	−3.4469	0.0621	0.4258	**1.0489**	−1.6935

Table 4. Risk associated with trading strategy 1 and 2 for Model 2. Risk premium is computed by using Sharpe ratio for testing set. Bold ones are the best model for each stock.

	Strategy 1			Strategy 2		
Stock name	SVR	MLP	ARIMA	SVR	MLP	ARIMA
SONS	**1.19792**	−1.46726	−2.19662	**2.80284**	−2.58553	0.51463
MSFT	**1.35187**	0.80705	0.47451	−0.48011	**2.07998**	1.34284
YHOO	**1.70587**	−1.33823	−4.95796	**2.01222**	−0.04081	0.48881
INTC	−4.23502	−**0.71304**	−1.24115	−5.03456	−4.20251	−**3.07209**
GEOI	−1.73611	0.63758	**1.69520**	**2.18232**	−2.86777	−0.68002
ABLE	−**1.75049**	−4.51655	−1.78067	**0.08230**	−1.99021	−3.96268
CSCO	−**1.47745**	−3.30816	−3.91956	−1.15257	−3.44856	−5.38312
AMZN	1.69058	**1.88736**	1.65113	−3.48956	**2.76652**	0.35136
MCEL	−**1.69580**	−4.66445	−4.49595	−3.26953	−**0.23004**	−0.59622
FCEL	−3.64613	**1.13745**	0.06213	0.41489	**2.33361**	−1.69350

ARIMA techniques. Furthermore, we can conlude that the forecasting accurary of a techniques does not guarantee a better profit.

6. Conclusions

Several papers (Fama and French 1995, Pesaran and Timmerman 1995, Gencay 1996, Walker and Hatfield 1996, Gold and Lebowitz 1999) discussed the efficiency of US equity market and suggested that excess returns can be made by using publicly available information. However, financial research indicates that security prices do not reflect all publicly available information (Walker and Hatfield 1996). Our motivation is based on these findings, and is a part of the more general attempt to apply data mining techniques in order to discover hidden patterns in financial time series.

We have presented an approach for short term stock price prediction based on technical indicators. Two different models have been used to explain the relationship between stock price and technical indicators. Usually, these indicators help the investors/traders to identify a pattern in a stock price chart. We have observed a nonlinear and highly correlated relationship between current and previous indicators and underlying

Table 5. RMSE error of testing set for model 1 and 2 for each techniques. Bold ones are the best RMSE values.

	Model 1			Model 2		
Stock name	SVR	MLP	ARIMA	SVR	MLP	ARIMA
SONS	**0.237024**	0.248295	0.46593	0.331708	**0.183784**	0.46593
MSFT	**0.409455**	0.421117	1.208301	0.467078	**0.345463**	1.208301
YHOO	1.960879	**1.597523**	3.334722	**2.902069**	2.985166	3.334722
INTC	0.663811	**0.557087**	1.668769	**1.004055**	1.435188	1.668769
GEOI	**0.842642**	1.001138	3.399409	**2.835715**	3.030869	3.399409
ABLE	**0.92343**	1.629718	2.226838	**1.596917**	1.778303	2.226838
CSCO	**0.327995**	0.412449	1.070846	**0.769111**	0.928634	1.070846
AMZN	**2.715294**	2.860663	4.264398	**4.023659**	4.371194	4.264398
MCEL	**0.11354**	0.12119	0.275135	**0.163628**	0.174023	0.275135
FCEL	**0.379534**	0.443738	0.836351	0.545375	0.530562	0.836351

stock prices. This relationship has been exploited by using those indicators as the input variables for SVR and MLP techniques. Then, we have compare these methods with a pure time series technique which is called ARIMA.

SVR method and MLP networks are data driven techniques and they both have some free parameters (for example, one would choose the kernel function and its parameters before the algorithm starts and the topology of the MLP networks has to be determined). The problem of finding the optimal parameters and network structures in an automated manner is still an open-problem. Therefore, we performed 10-fold cross-validation technique to determine the free parameters and network architectures.

These three techniques were compared with each other in terms of excess return and RMSE criteria. The results showed that performance of a forecasting technique depend on the trading strategy.

References

Beechey, M., Gruen, D. and Vickery, J., 2000. The efficiency market hypothesis: a survey. Research Discussion Paper (Reserve Bank of Australia).

Brock, V., Lakonishok, J. and LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance* (December), 1731–1764.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern classification. *Data Mining and Knowledge Discovery*, 2(2), 121–167.

Collobert, R. and Bengio, S., 2001. Svmtorch: support vector machines for large-scale regression problems. *Journal of Machine Learning Research*, 1, 143–160.

Cortes, C. and Vapnik, V., 1995. Support vector networks. *Machine Learning*, 20, 273–297.

Cristianini, N. and Shawe-Taylor, J., 2000. *An Introduction to Support Vector Machines*. Cambridge, UK: Cambridge University Press.

Dourra, H. and Siy, P., 2002. Investment using technical analysis and fuzzy logic. *Fuzzy Sets and Systems*, 127, 221–240.

Fama, E.F. and French, K.R., 1995. Size and book-to-market factors in earnings and returns. *The Journal of Finance*, 50(1), 131–155.

Gencay, R., 1996. Non-linear prediction of security returns with moving average rules. *Journal of Forecasting*, 15, 165–174.

Gold, S.C. and Lebowitz, P., 1999. Computerized stock screening rules for portfolio selection. *Financial Services Review*, 8(2), 61–70.

Haykin, S., 1994. *Neural Networks: A Comprehensive Foundation*. New York: MacMillan Publishing Company.

Lee, Y.-J. and Mangasarian, O.L., 2001. Rsvm: reduced support vector machines. *CD of Proceedings of the First SIAM International Conference on Data Mining*, 89.

Leigh, W., Paz, M. and Purvis, R., 2002. An analysis of a hybrid neural network and pattern recognition technique for predicting short-term increases in the NYSE. *The International Journal of Management Science*, 30, 69–76.

Osuna, E., Freund, R. and Girosi, F., 1997. Training support vector machines: an application to face detection. *Proceedings of Computer Vision and Pattern Recognition'97*, 130–136.

Pesaran, H.M. and Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. *The Journal of Finance*, 4, 1201–1228.

Platt, J.C., 1999. Fast training of support vector machines using sequential minimal optimization. In: B. Schölkopf, C. Burges and A. Smola, eds. *Advances in Kernel Methods*, Cambridge: The MIT Press.

Pontil, M. and Verri, A., 1997. Properties of support vector machines. Technical report. Massachusetts Institute of Technology, Artificial Intelligence Laboratory.

Sharpe, W.F., 1966. Mutual fund performance. *Journal of Business*, 119–138.

Smola, A. and Schölkopf, B., 2004. A tutorial on support vector regression. *Statistics and Computing*, 14, 199–222.

Suykens, J.A.K. and Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural Processing Letters*, 9(3), 293–300.

Tay, F.E.H. and Cao, L.J., 2002. Modified support vector machines in financial time series forecasting. *Neurocomputing*, 48, 847–861.

Tian, G.G., Wan, G.H. and Gua, M., 2002. Market efficiency and the returns to simple technical trading rules: new evidence from US equity market and Chinese equity markets. *Asia-Pacific Financial Markets*, 9, 241–258.

Trafalis, T.B. and Ince, H., 2002. Benders decomposition technique for support vector regression. *IJCNN'02, Proceedings of the 2002 International Joint Conference on Neural Networks*. Vol. 3. New York: IEEE, 2767–2772.

Trafalis, T.B., Ince, H. and Mishina, T., 2003. Support vector regression in option pricing. *Proceedings of Conference on Computational Intelligence and Financial Engineering (CIFer 2003)*, March 20–23. Hong Kong.

Tsaih, R., 1999. Sensitivity analysis, neural networks and the finance. 3830–3835.

Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. New York: Springer Verlag.

Walker, M.M. and Hatfield, G.B., 1996. Professional stock analysts' recommendations: implications for individual investors. *Financial Services Review*, 5(1), 13–29.

Yao, J. and Tan, C.H., 2000. A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, 34, 79–98.