A Stock Market Prediction Method Based on Support Vector Machines (SVM) and Independent Component Analysis (ICA)

Hakob GRIGORYAN
Bucharest University of Economic Studies, Bucharest, Romania
Grigoryanhakob90@yahoo.com

Abstract: *The research presented in this work focuses on financial time series prediction problem. The integrated prediction model based on support vector machines (SVM) with independent component analysis (ICA) (called SVM-ICA) is proposed for stock market prediction. The presented approach first uses ICA technique to extract important features from the research data, and then applies SVM technique to perform time series prediction. The results obtained from the SVM-ICA technique are compared with the results of SVM-based model without using any pre-processing step. In order to show the effectiveness of the proposed methodology, two different research data are used as illustrative examples. In experiments, the root mean square error (RMSE) measure is used to evaluate the performance of proposed models. The comparative analysis leads to the conclusion that the proposed SVM-ICA model outperforms the simple SVM-based model in forecasting task of nonstationary time series.*

Keywords: *support vector machines, regression, independent component analysis, financial time series, stock prediction*

1 Introduction

In recent years, the fast growing financial markets opened new horizons for investors and the same time bringing new challenges for financial analysts in their efforts to make effective decisions and reduce the investment risks. Stock market is a highly dynamic and complex system since there are a great number of interacting factors that affect the future prices [1]. In fact, stock market prediction means understanding which economic and non-economic factors affect market prices in order to predict the target variables based on the analysis of historical data. Thus, many researchers in the field of economic predictions have claimed that stock market prediction is a difficult task compared with other time series analysis problems as stock data is non stationary, random and chaotic [2][3]. However, financial time series are characterized by uncertainty and noise, there is some evidence that stock markets can be predictable through the use of different methods ranging from econometric to machine learning techniques [4]. Currently, advancements in different fields of applied mathematics and information technologies have led to the development of novel prediction models based on artificial intelligence techniques. Existing research indicates that statistical techniques are useful for modelling linear problems but those fail to capture the non-linear behaviour presented in financial time series as stock markets are non-linear deterministic systems [5].

In recent years, numerous machine learning-based models have been presented for time series analysis. Among them, Support Vector Machines (SVM) is a novel technique designed to solve non-linear classification and regression problems in time series analysis. SVMs are based on the structural risk minimization principle which allows them to estimate a function by minimizing an upper bound of generalization error [6]. Due to its ability to achieve a high generalization performance and testing accuracy, SVMs have been successfully applied for time series prediction domain. In regard to financial forecasting, Trafalis and Ince (2000), Tay and Cao, (2001) introduced an application of SVM for stock market prediction, and showed promising results compared with neural network-based models. Similarly, Kim, (2003) applied SVM to forecast the stock price index and concluded that SVM can be successfully applied to stock market prediction as an alternative to neural networks. Also, Huang et al., (2005) applied SVM to forecast the movement direction of stock market, and showed that SVM has better prediction performance compared with other statistical and machine learning methods.

The main problem in stock market forecasting is the inherent noise of the financial data. By removing unwanted information from historical data we can increase forecasting accuracy and speed. Several researchers have proposed hybrid models based on SVM and feature selection techniques. For example, Cao et al, (2003) presented prediction model based on SVM technique and also three different feature extraction techniques, namely, principal component analysis (PCA), kernel PCA (KPCA) and Indepentent Component Analysis (ICA). They concluded that use of feature extraction techniques had increased prediction accuracy and among proposed techniques the best performance showed model based on KPCA and SVM techniques. Hsu et al., (2009) applied two-stage architecture for stock price prediction based on self-organizing map (SOM) and SVM. They used SOM to decompose and classify the input data and support vector machines for regression (SVR) to predict prices, and showed that their model outperforms the standard SVM-based model in stock price prediction. Lee (2009) presented a stock market trend prediction model based on support vector machine (SVM) with a hybrid feature selection method named F-score and Supported Sequential Forward Search (F_SSFS). Their experimental results showed that their proposed hybrid model outperforms neural network based model combined with other feature selection methods. Kao et al., (2013) introduced a novel combined model based on wavelet transform, multivariate adaptive regression splines (MARS), and support vector regression (SVR) to forecast stock prices, and concluded that their proposed approach outperforms other models in forecasting the stock prices.

One model can be suitable to predict a certain financial market but fail to predict another market's data as there are different factors affecting stock prices varying from one market to another. In recent years, a series of studies concerning data forecasting based on multivariate input models including different macroeconomic variables and technical indicators have been reported [15][16][17][18]. The experimental results showed that different externally determined variables based on technical and fundamental analysis are useful in prediction of stock markets data.

The main objective of this study is the investigation of the effectiveness of combined prediction model based on ICA and SVM techniques in forecasting task of noisy stock data. In such a model, first pre-processing step is used to prepare the research data and select important features by ICA method, and then SVM based prediction model is constructed based on the selected variables. The superiority of the proposed model is shown by the comparative analysis of stock market prediction model based on SVM and ICA techniques against single SVM-based prediction model without using any feature selection technique.

The remainder of this paper is organized as follows. Section 2 gives brief introduction to Support Vector Machines (SVM) and Independent Component Analysis (ICA). Section 3 describes the proposed methodology, including data collection, preparation and forecasting model. In Section 4, the experimental results together with a comparative analysis are summarized and discussed. Finally, the concluding remarks are presented in the Section 5.

2 Background

2.1 Support Vector Machines (SVM)

Support Vector Machines (SVM) is a family of learning algorithms originated as an implementation of the structured risk minimization (SRM) principle proposed by Vapnik [19].

The basic idea behind SVM model is to represent the given examples of data as points in a high-dimensional feature space and linearly separate the feature vectors by a maximum margin hyperplane. The diagram of linearly separable SVM is depicted in figure 1. The data points closest to the maximum margin hyperplane lying on the dotted line are used to determine the regression surface. This small subsets of data points are called support vectors, while the points within the ε–insensitive zone are not important in terms of the regression function and contribute to the error loss function. In time series analysis, the application of SVMs used in regression analysis is called Support Vector Regression (SVR) [20].

Given a training data set $G = \{(x_i, y_i),\ i = 1, \dots, l\} \subset X \times \mathbb{R}$, where $X \subset \mathbb{R}^n$ denotes the space of the input patterns, the SVR function can be expressed as:

$$f(x) = \omega \cdot \varphi(x) + b \quad (1)$$

where, $\omega \in X$ is a weight vector, $b \in \mathbb{R}$ is a bias and φ represents the mapping function.

The objective of the SVR is to find a function f that has the most ε deviation from the target y_i and, the most possible flat f. The problem can be solved by finding the small values of the Euclidian norm $\|\omega\|^2$ which can be achieved by solving the following optimization problem.

$$minimize \quad \frac{1}{2}\|\omega\|^2 \quad (2)$$

$$subject\ to \begin{cases} y_i - \omega \cdot \varphi(x_i) - b \le \varepsilon \\ \omega \cdot \varphi(x_i) + b - y_i \le \varepsilon \end{cases}$$

The optimization problem (2) is feasible when there is f such that $|f(x_i) - y_i| \le \varepsilon$ for all $(x_i, y_i) \in G$. When the training data is not linearly separable, slack variables ξ_i, ξ_i^* are introduced to deal with unfeasible constraints of the optimization problem (2). The optimization problem (2) can be re-formulated as:

$$minimize \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \quad (3)$$

$$subject\ to \begin{cases} y_i - w^T x_i - b \le \varepsilon + \xi_i \\ w^T x_i + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

where $\frac{1}{2}\|w\|^2$ is the regularization term preventing over-learning, $(\xi_i + \xi_i^*)$ is the empirical risk; and $C > 0$ is called a regularization constant which controls the trade-off between the empirical risk and regularization term.

The ε-insensitive loss function $|\xi|_\varepsilon$ can be described as:

$$|\xi|_\varepsilon = \begin{cases} 0 & if\ |\xi| < \varepsilon \\ |\xi| - \varepsilon & otherwise \end{cases} \quad (4)$$

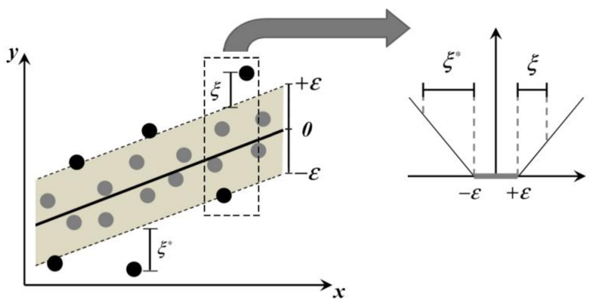

Fig 1: Support Vector Machines (SVM) and ε-insensitive loss function

The quadratic optimization problem (3) can be solved by introducing Lagrange multipliers.
Let L be the Lagrange function:

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) - \sum_{i=1}^{l}\alpha_i(\varepsilon + \xi_i - y_i + w^T x_i + b) - \sum_{i=1}^{l}\alpha_i^*(\varepsilon + \xi_i^* + y_i - w^T x_i - b) - \sum_{i=1}^{l}(\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (5)$$

where $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$ are the Lagrange multipliers. Thus the dual optimization problem corresponding to (3) is given by

$$maximize \ -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)x_i{}^T x_j - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \quad (6)$$

$$subject\ to \begin{cases}\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C]\end{cases},$$

by changing the equation $w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i$,

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i{}^T x + b \quad (7)$$

Consequently, applying Lagrange theory and the Karush-Kuhn-Tucker condition, the general SV regression function can be expressed by

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) + b \quad (8)$$

where $K(x_i, x)$ is defined as kernel function. The value of kernel function is equal to the inner product of x_i and x_j in the feature space $\phi(x_i)$ and (x_j) , such that:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (9)$$

Any function that satisfies the condition proposed by Mercer can be applied as the kernel function [19]. The most common kernel functions are Gaussian kernel and polynomial kernel functions defined as:

$$K_{gaussian} = exp(-(1/\sigma^2)(x_i - x_j)^2)$$

$$K_{polynomial} = (x_i \cdot x_j + 1)^d \quad (10)$$

where d and σ^2 are the kernel parameters [21],[22].

2.2 Independent Component Analysis (ICA)

Independent component analysis (ICA) is an unsupervised method for extracting individual signals from a multivariate signal proposed by [23]. ICA decomposes the given dataset into components so that each component is statistically independent from the others and assumed to be non-Gaussian. The basic form of ICA model is shown in figure 2.

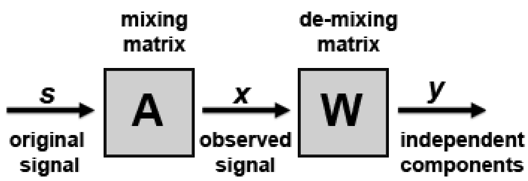

Fig. 2. Scheme of general ICA process

The original source signals *s* are mixed through the mixing matrix *A* to form the observed signal x, then the de-mixing matrix *W* transforms the observed signal into the independent components y.
Let $x_1, x_2, \dots, x_n$ be the random variables, defined by the linear combinations of the random variables $s_1, s_2, \dots, s_n$, then for any $1 \leq i \leq n$,

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n \quad (11)$$

Using vector-matrix notation and denoting the matrix A with elements $a_{i,j}$, by $A = \|a_{i,j}\|_{1 \leq i,j \leq n}$, and row vectors $x = [x_1, x_2, \dots, x_m]^T$, $s = [s_1, s_2, \dots, s_n]^T$.

The ICA model can be defined by the matrix as follows:

$$x = As = \sum_{i=1}^{n} a_i s_i \qquad (12)$$

where a_i is the ith row of unknown matrix A, and $m \geq n$

The ICA model aims to estimate the latent variables s and unknown mixing matrix A from x with the assumption that source components s_i are statistically independent and at most one of the components has a Gaussian distribution. In other words, the ICA model tends to find a de-mixing matrix W that makes the latent variables statistically independent such that,

$$y = Wx = \sum_{i=1}^{n} w_i x_i \qquad (13)$$

where $y = [y_1, y_2, \ldots, y_m]^T$ is the independent component vector, the vector w_i is the ith row of the de-mixing matrix W. The elements of the vector y are called independent components (ICs) and they are used to estimate the source components s_i.

One of the ways to determine the de-mixing matrix W is the maximization of the statistical independence of ICs such as the maximization of non-Gaussianity. The non-Gaussianity of the ICs can be measured by the negentropy which is based on the concept of the information-theoretic quantity of entropy.

The negentropy of a random variable y with m_y mean and covariance matrix Σ_y is defined by the following equation:

$$J(y) = H(y_{gauss}) - H(y) \qquad (14)$$

where y_{gauss} is a Gaussian random variable of the same covariance matrix as y and is distributed as $N(m_y, \Sigma_y)$ and H is the entropy of a random vector.

The entropy H is defined as:

$$H(y) = -\int p(y) \log p(y) d(y) \qquad (15)$$

where $p(y)$ is the the density.

The value of negentropy is always non-negative and equal to zero when y has a Gaussian distribution.

The most popular algorithm to find a maximum of the non-Gaussianity of Wx is a fixed-point algorithm called FastICA [24]. The negentropy can be approximated by:

$$[E\{G(y)\} - E\{G(v)\}]^2 \qquad (16)$$

where G is any given quadratic function and v is a Gaussian variable of zero mean.

By denoting g as the derivative of the nonquadratic function G, the one unit version of FastICA algorithm is as follows:

1. Choose an initial random vector $\boldsymbol{w}$,
2. Let $w^+ = E\{xg(w^t x)\} - E\{g'(w^t x)\}w$,
3. Let $w = w^+ / \|w\|$
4. If not converged, go back to 2

3 Research Design

3.1 The combined SVM-ICA model

This paper proposes a stock market forecasting model by combining SVM and ICA techniques (called SVM-ICA model). The proposed prediction framework consists of two stages.

In the first stage, the ICA technique is used to extract information from research data. ICA technique uses the observed data to convert original signals into separate independent components (ICs).

In the second stage, SVM technique is applied to forecast the stock prices. The features extracted by ICA technique are used as input variables to construct the prediction model.

In our prediction model, the future values of target variable are predicted by using the previous values of the same variable and sets of variables obtained from technical and fundamental analysis of the stock market.

The mathematical description of general prediction model is the following:

$$\hat{Y}_{(t+p)} = f\left(Y_t^{(d)}, X_t^{(d)}\right) \quad (17)$$

and

$$Y_t^{(d)} = \{Y_t, Y_{t-1}, Y_{t-2}, \dots, Y_{t-d+1}\} \quad (18)$$

$$X_t^{(d)} = \{X_t, X_{t-1}, X_{t-2}, \dots, X_{t-d+1}\} \quad (19)$$

where $\hat{Y}_{(t+p)}$ is the predicted closing price for the period p, d is the delay, Y_t is the closing price at the moment of time t, and by $X_t = (X_t(1), X_t(2), \dots, X_t(n))^T$ we denote the vector entries of which are the values of the indicators having influence on Y_t. In our model, we consider p=1 and, as a result of correlogram analysis, d=2.

In both testing and training phases, the aim of the proposed methodology is to obtain the ICA representation of data in each input set and, based on it, to derive the proper parameters for the SVM model. In the training phase, the ICA model was applied to represent information concerning the given collection of input data. Then the SVM model is derived using ICA representation of the input data. In the testing phase the ICA model together with the SVM parameters are re-computed based on the following computation scheme.

Let $S' = \{x_1, x_2, \dots, x_T\}$ be the set of training input data,

and $S'_N = \{x_{T+1}, x_{T+2}, \dots, x_n\}$ be the set of input testing data

then by denoting $S_{T+K} = S' \cup \{x_{T+1}, x_{T+2}, \dots, x_{T+K}\}$

The algorithm used in forecasting the new, unseen yet, samples is described as follows.

Let x_{T+K+1} be the new sample

For $k = 1, \dots, n-T$

Step 1. Compute the ICA model corresponding to the input data S_{T+K}

Step 2. Compute the SVR model to predict $\hat{Y}_{T+K}$ from (17)

Output: $\{\hat{Y}_{T+1}, \hat{Y}_{T+2} \dots \hat{Y}_n\}$

3.2. Data collection and preparation

To better understand the effectiveness of the proposed prediction model, two different experimental data sets are used in this paper, both based on real stock market data.

First data set is based on the historical weekly observations of a set of variables obtained from Bucharest Stock Exchange. The given data set covers the period from 3/9/2008 to 11/30/2014, a total of 350 cases of trading weeks. The research data set includes a total of 39 variables from which 35 variables were selected from technical analysis and 4 macroeconomic variables were obtained from fundamental analysis of OMV Petrom stock (symbol OMV). The closing price of OMV Petrom stock was used as a forecasting variable.

Second data set presents daily observations taken from Baltic Stock Exchange. The entire data set covers the period from 03/12/2012 to 12/30/2014, a total of 700 daily observations. The data set includes 35 variables obtained only from technical analysis of Tallink stock (symbol TALL1T). The closing price of Tallink stock was used as a target variable for prediction model.

The collected data samples have different scales as they come from different markets and sources. It is essential to consider data pre-processing by normalization in order to improve the training step and prediction results of the proposed model. Thus, the original data sets are normalized into the range of [0,1] using the formula given by

$$V = \frac{v - v_{min}}{v_{max} - v_{min}} \quad (20)$$

where V is the normalized data, v is the original data, v_{max} and v_{min} are maximum and minimum values of v.

4 Experimental Results and Discussion

This study integrates ICA and SVM techniques to predict the stock market closing price. The proposed SVM-ICA forecasting model was tested on two different databases.

The weekly closing price of OMV Petrom stock and daily Tallink stock closing price are used in this study as target variables. The whole data set is divided into two parts. The first part (70%) is used for training step and the second part (30%) is reserved for prediction step.

For ICA process, we consider m input time series (for the first case m=39 and for the second case m=35, based on the number of input variables that influence the closing price). Each of the time series is considered as a row formed by matrix X of size m x n (for the first case n=350 and for the second case n=700, based on the total observations of variables). The separate matrix W and the independent component matrix Y are calculated by the ICA method, where each row of Y represents an individual IC.

We adopt a fast fixed-point algorithm for ICA (FastICA) to extract independent components from collected financial time-series data. The number of ICs for two databases is as follows: 4 out of 39 variables are selected for the first data set (OMV Petrom) and 3 out of 35 variables for the second data (Tallink).

After pre-processing step, the SVM model was developed for data training. In building the prediction model, the performance of SVM depends on the accurately selected kernel function and parameters, such as, regularization constant C and loss function ε defined in section 2.1. In this experiment, the Radial Basis Function (RBF) was used as a kernel function, $K(x,x') = exp(-\gamma\|x - x'\|^2), \gamma > 0$ as it is suitable for non-linear problems. In the literature of SVM, one of the suggested methods for the choice of the parameters C and ε is based on the cross-validation via grid-search method proposed in [25]. The pair of parameters C and ε with the best cross-validation accuracy which generate the minimum forecasting is considered the best parameter set. Then, the trained SVM model with proper parameter setting is preserved and employed in the testing phase based on the algorithm presented in Section 3.1.

For building the SVM prediction model, the LIBSVM tool box was used in this study [26].

The prediction performance is evaluated in terms of root mean squared error (RMSE), defined by:

$$RMSE(T,P) = \sqrt{\frac{1}{nr}\sum_{i-1}^{nr}\left(T(i) - P(i)\right)^2} \qquad (21)$$

where $T = \left(T(1), T(2), \dots, T(nr)\right)$ is the vector of target values, $P = \left(P(1), P(2), \dots, P(nr)\right)$ is the vector of predicted values and nr is the number of data samples.

The prediction results of the proposed integrated SVM-ICA model are compared with the model based on single SVR technique without using any preprocessing step. The forecasting error is given in the table 1.

Table 1. The forecasting results using SVM-ICA model against single SVM model

Data set	RMSE	Model
Bucharest Stock Exchange	0.0729 0.022593	SVM SVM-ICA
Baltic Stock Exchange	0.12286	SVM
	0.019191	SVM-ICA

The actual closing price of OMV Petrom stock and its predicted values for 105 new samples using SVM-ICA and single SVM models are depicted in figure 3 and 4. The data set was obtained from Bucharest Stock exchange with 350 weekly observations, and the prediction results are given for 105 news samples.

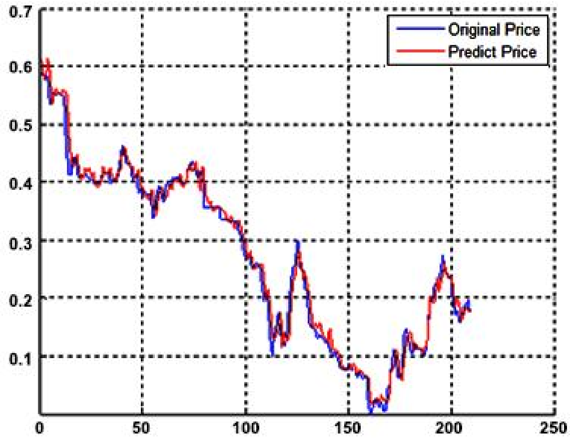

Fig. 3. Prediction results of OMV Petrom stock using SVM-ICA model, RMSE=0.022593.

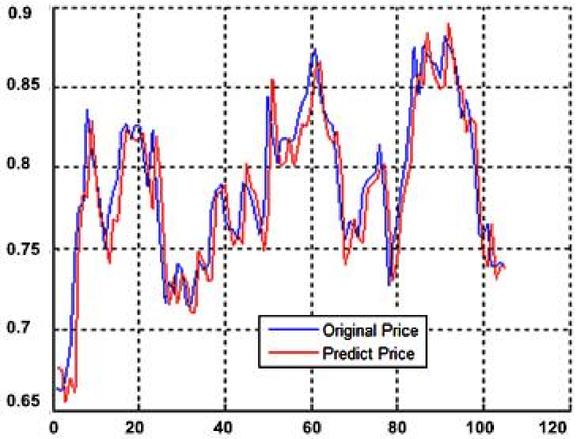

Fig. 4. Prediction results of OMV Petrom stock using single SVM model without any pre-processing step, RMSE=0.0729.

For the data set obtained from Baltic Stock exchange with 700 daily observations, the prediction results of the closing price of Tallink stock for 210 news samples yet not seen are depicted in the figure 5 and 6.

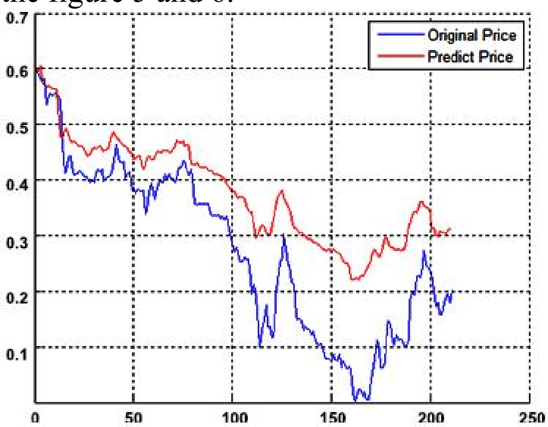

Fig. 5. Prediction results of Tallink stock using SVM-ICA model, RMSE=0.019191

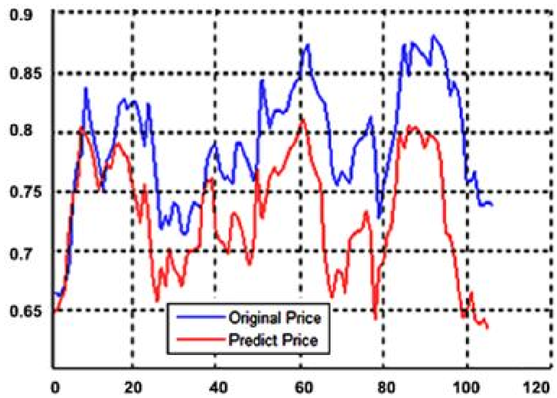

Fig. 6. Prediction results of Tallink stock using single SVM model without any pre-processing step, RMSE=0.12286.

5 Conclusion

Currently, stock market prediction is one of the challenging tasks of time-series analysis domain. In recent years, different prediction models have been presented using novel machine learning techniques. SVM is a new but promising approach for financial predictions. In this research, the combined stock market prediction model based on ICA and SVM techniques was examined. The main idea was to demonstrate the effectiveness of the hybridization of two methods in forecasting task of noisy data. This study used ICA to select input variables from technical and fundamental analysis, and SVM model to forecast the closing price. To show the effectiveness of the proposed methodology, two different data sets were used in the experiments. In addition, comparative analysis has been conducted against the model that uses only SVM technique without pre-processing step. The results obtained from the hybridized SVM-ICA model showed that ICA method effectively improved forecasting results from the point of view RMSE measure. This study allows us to conclude that SVM technique is an effective method for stock market prediction when it is combined with feature selection techniques. Moreover, experimental results obtained from proposed two-stage model encourage us to use other pre-processing methods for gaining better experimental results.

References

[1] Gujarati, D. N. (2003). Basic Econometrics. 4th. *New York: McGraw-Hill.*

[2] Yudong, Z., & Lenan, W. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. Expert Systems with Applications, 36(5), 8849–8854.

[3] Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, *6*(3), 205-213.

[4] Tsang, P. M., Kwok, P., Choy, S. O., Kwan, R., Ng, S. C., Mak, J., ... & Wong, T. L. (2007). Design and implementation of NN5 for Hong Kong stock price forecasting. *Engineering Applications of Artificial Intelligence*, *20*(4), 453-461.

[5] Hill, T., O'Connor, M., & Remus, W. (1996). Neural network models for time series forecasts. *Management science*, *42*(7), 1082-1092.

[6] Vapnik, V. N., & Vapnik, V. (1998). *Statistical learning theory* (Vol. 1). New York: Wiley.

[7] Trafalis, T. B., & Ince, H. (2000, July). Support vector machine for regression and applications to financial forecasting. In *ijcnn* (p. 6348). IEEE.

[8] Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega*, *29*(4), 309-317.

[9] Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1), 307-319.

[10] Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, *32*(10), 2513-2522.

[11] Cao, L. J., Kok Seng Chua, W. K. Chong, H. P. Lee, and Q. M. Gu. "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. " *Neurocomputing* 55, no. 1 (2003): 321-336.

[12] Hsu, S. H., Hsieh, J. P. A., Chih, T. C., & Hsu, K. C. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications*, *36*(4), 7947-7951.

[13] Lee, M. C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications*, *36*(8), 10896-10904.

[14] Kao, L. J., Chiu, C. C., Lu, C. J., & Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems*, *54*(3), 1228-1244.

[15] Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Applications*, *38*(5), 5311-5319.

[16] de Oliveira, F. A., Nobre, C. N., & Zarate, L. E. (2013). Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index–Case study of PETR4, Petrobras, Brazil. *Expert Systems with Applications*, *40*(18), 7596-7606.

[17] Cocianu, C. L., & Grigoryan, H. (2015). An Artificial Neural Network for Data Forecasting Purposes. *Informatica Economica*, *19*(2), 34-45.

[18] Grigoryan, H. Stock Market Prediction using Artificial Neural Networks. Case Study of TAL1T, Nasdaq OMX Baltic Stock. *Database Systems Journal BOARD*, 14.

[19] Vapnik V.N. (2000). The Nature of Statistical Learning Theory, Springer, New York.

[20] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression.*Statistics and computing*, *14*(3), 199-222.
[21] Catalina Cocianu, Luminita State, *Kernel-Based Methods for Learning Non-Linear SVM*, Economic Computation and Economic Cybernetics Studies and Research, No. 1/2013, ISSN 0424-267X, pp. 41-60
[22] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in Proc. 2th Berkeley Symp. Mathematical Statistics and Probabilistics. Berkeley, CA: Univ. California Press, 1951, pp. 481–492.
[23] Herault, J., & Jutten, C. (1986, August). Space or time adaptive signal processing by neural network models. In *Neural networks for computing* (Vol. 151, No. 1, pp. 206-211). AIP Publishing.
[24] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Transactions on Neural Networks 10 (1999) 626–634.
[25] Hsu, C.-W., Chang, C.-C., & Lin, C.-J., (2004). A practical guide to support vector classification. Technical report. Department of Computer Science and Information Engineering, National Taiwan University.
[26] Chang, C. C., & Lin, C. J. (2001). LIBSVM: a Library for Support Vector Machines.
[29] Cocianu, C., State, Luminita, & Vlamos, P. (2009). NEURAL IMPLEMENTATION OF A CLASS OF PCA LEARNING ALGORITHMS. *Economic Computation and Economic Cybernetics Studies and Research*,*43*(3), 141-153.

Hakob GRIGORYAN has graduated the Faculty of Cybernetics of the State Engineering University of Armenia (Polytechnic) in 2011. In 2014 he graduated the Faculty of Informatics and Mathematics of the University of Bucharest with a specialization in Database and Web Technologies. At the present, he is earning his PhD degree in Economic Informatics at Bucharest University of Economic Studies, coordinated by Professor Catalina-Lucia Cocianu. His PhD thesis is "Machine Learning-Based Techniques for Financial Data Analysis and Forecasting Purposes".