SECTION II. ECONOMIC-MATHEMATICAL MODELING

MICEX INDEX FORECASTING: THE PREDICTIVE POWER OF NEURAL NETWORK MODELING AND SUPPORT VECTOR MACHINE

Agata M. Lozinskaia

ORCID ID: 0000-0001-8723-275X, Researcher ID: L-6971-2015
E-mail: AMPoroshina@gmail.com
National Research University Higher School of Economics
27, Lebedeva st., Perm, 614070, Russia

Victor A. Zhemchuzhnikov

ORCID ID: 0000-0001-6513-3611, Researcher ID: D-2384-2017
E-mail: viktr5909@mail.ru
National Research University Higher School of Economics
27, Lebedeva st., Perm, 614070, Russia

The ability to predict the dynamics of financial instruments is an important topic for financial market players. In the context of large and heterogeneous information, there is a need to use effective methods to data processing for management decision-making. In particular, machine learning techniques are becoming very popular in financial modeling. The aim of this paper is to forecast the Russian stock price index by using machine learning methods such as neural network modeling and support vector machine and to examine their predictive power. For economic and mathematical modeling, we use statistical and analytical information on the dynamics of the *MICEX* stock price index, fundamental and technical indicators of the stock market for the period 2002–2016 years. For computational experiments, we use training, testing and validation datasets and software for machine learning in Python. The predictive power of the methods is estimated on the validation data with using both traditional indicators of mathematical statistics (such as absolute and relative prediction error) and count coefficient of determination. We found the use of a longer time period for the *MICEX* index that corresponds to large training data set in neural network modeling has led to training error reduction. The predictive power of support vector machine on validation data set is higher comparing with neural network modeling. However, that difference in prediction metrics is not significant. The development of a methodology for filtering input data and trading strategy based on machine learning algorithms are possible directions for further research.

Keywords: forecasting, financial time series, machine learning, neural network modeling, support vector machine, computer experiment, stock market, MICEX index, returns.

ПРОГНОЗИРОВАНИЕ ИНДЕКСА ММВБ: ПРЕДСКАЗАТЕЛЬНАЯ СИЛА МЕТОДА НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ И МЕТОДА ОПОРНЫХ ВЕКТОРОВ

Агата Максимовна Лозинская

ORCID ID: 0000-0001-8723-275X, Researcher ID: L-6971-2015

Электронный адрес: AMPoroshina@gmail.com

Национальный исследовательский университет «Высшая школа экономики»

614070, Россия, Пермь, ул. Лебедева, 27

Виктор Андреевич Жемчужников

ORCID ID: 0000-0001-6513-3611, Researcher ID: D-2384-2017

Электронный адрес: viktr5909@mail.ru

Национальный исследовательский университет «Высшая школа экономики»

614070, Россия, Пермь, ул. Лебедева, 27

Возможность прогнозирования динамики финансовых инструментов представляет собой актуальную задачу для участников финансового рынка. В условиях большого потока разнородной информации возникает потребность в использовании эффективных методов их обработки для выработки оперативных управленческих решений. В частности, все большее распространение в финансовом моделировании получают методы машинного обучения. Цель работы заключается в моделировании прогноза российского биржевого индекса с помощью таких методов машинного обучения, как метод нейросетевого моделирования и метод опорных векторов, и исследовании их предсказательной силы. Информационную базу экономико-математического моделирования составили статистические и аналитические данные о динамике индекса ММВБ, фундаментальных и технических индикаторов фондового рынка за 2002–2016 гг. Результаты компьютерных экспериментов выполнены на обучающей, тестирующей и подтверждающей выборке с использованием соответствующих библиотек машинного обучения на языке Python. Оценка предсказательной силы методов осуществлялась на подтверждающей выборке с использованием традиционных показателей математической статистики (абсолютной и относительной ошибки прогноза) и счетного коэффициента детерминации. Установлено, что использование более длительного временного промежутка для индекса ММВБ и, соответственно, большего числа наблюдений, при обучении нейронной сети способствовало уменьшению ошибки обучения. На подтверждающей выборке отмечена более высокая предсказательная точность результатов, полученных с помощью метода опорных векторов в сравнении с методом нейросетевого моделирования. Однако обнаруженная разница в соответствующих показателях качества прогноза незначительна. Возможные направления дальнейших исследований включают в себя разработку методологии фильтрации входных данных и создание торговой стратегии на основе алгоритмов машинного обучения.

Ключевые слова: прогнозирование, финансовые временные ряды, машинное обучение, нейросетевое моделирование, метод опорных векторов, компьютерный эксперимент, фондовый рынок, индекс ММВБ, доходность.

Introduction

The *MICEX* stock price index is the major indicator of the Russian stock market behavior which is calculated based on 50 most liquid Russian companies stocks. Especially, investors are interested in the *MICEX* index prediction to develop an optimal trading strategy, including futures trading. However, this task is complicated under the almost continuous data stream. For data stream processing, modern computing systems and machine learning techniques can be employed. The key feature of machine learning techniques is their ability to remember existing data relationships and to adapt to new ones. Neural networks and support vector machines became wi-despread in the financial time series modeling among different machine learning techniques.

Note that according to the Efficient Market Hypothesis (*EMH*), the asset price fully reflects all currently available public information. It means that it is impossible to predict a financial instrument behavior better than using the random walk model because all available information is already reflected in the asset price. However, the efficiency of financial markets continues to be studied because *EMH* is difficult to be tested [1].

This paper examines the predictive power of machine learning methods, such as neural network and support vector machine, in the *MICEX* index returns forecasting. Since the Russian stock market is an emerging market, and therefore, dynamically changing, verification and updating of existing empirical results are an important topic in the field of modeling and analyzing financial markets. This work represents one of the first known in the academic literature attempts to apply support vector machine for predicting the Russian stock index.

Literature review

One of the earliest works on the application of neural network modeling to the prediction of the financial instruments behavior (*IBM* stocks returns) belongs to H. White [2]. He demonstrates the high predictive power of a neural network to forecast of the financial instruments behavior in his later works, for example [3]. The advantages of a neural network in comparison with the classical time series models such as *ARIMA* include the ability to model time series forecast automatically, the lack of subjectivity in choosing the best model, flexibility and nonlinearity, etc. [4].

At the same time H. White [2] and the authors of recent studies, for example, R. Jammazi and S. Aloui [3], N.A. Valiotti and V.L. Abbakumov [4], I. Kaastra et al. [5] pointed out that the practical application of the neural network is associated with two main problems: the selection of independent variables and the neural network architecture selection. The first problem is not only typical for neural networks but it is an important step in the development of any financial time series prediction model. An automatic "filter-

ing" of irrelevant variables is possible through the update of weights in neural networks during training process [6]. As it was shown by R. Pakath et al. [7], an increase in a number of irrelevant variables can lead to a deterioration of prediction. Expert knowledge is commonly used to find the optimal set of factors. In particular, factors that influence the financial time series behavior are usually divided into two categories: technical and fundamental. Technical indicators are based on the different transformations of the financial instrument values, such as lagging, calculation of moving average, etc. They are presented in Table. 1. Fundamental variables reflect factors describing the fundamental state of the economy, such as macroeconomic indicators.

In this paper expert knowledge, especially the set of factors that determines the Russian stock market behavior is based on the results of previous studies. Generally, in previous papers, fundamental indicators are used. For example, the growing dependence of the Russian stock market from external factors, mostly from the US stock market (for example, the *MSCI* index for US behavior and the rate on short-term Treasury bonds), has been demonstrated in the paper by S.A. Anatolyev [8]. The author used the regression analysis on *MSCI* index for Russia (1995–2005 years) and his findings relevant to the study by A.A. Peresetsky [9] on an explanation of the *MICEX* index daily returns (2001–2011 years). By using *ARIMA* and *GARCH* models, A.A. Peresetsky found a statistically significant effect for the one-day lagged *MICEX* index returns, the one-day lagged American *S&P500* index returns, the one-day lagged *WTI* oil price and the Japanese *NIKKEI* index returns. Results of the papers [8] and [9] signal that not only certain indicators of American and Japanese stock market impact on the Russian stock market but also there is the lagged effect of the *MICEX* index returns on its current dynamics. Additionally, A.A. Peresetsky [9] using regression with a moving window found that the dependence of the Russian stock market on oil prices was weakened.

Table 1

Technical variables in studies on neural network modeling of financial instruments

Authors	Dependent variable	Time period, country	Independent variables
H. White [1]	*IBM* stock returns	1974–1980 years, US	Stock returns with lags from 1 to 5 days
F. Tay, C. Lijuan [10]	Relative percentage difference of *S&P500* futures prices for 5 days	1989–1993 years, US	5-day exponential moving average, dependent variable with a 10-minute, 15-minute and 20-day lags
Y. Zhang, Z. Yudong, W. Lenan [11]	*S&P500* index	1998–2008 years, US	10, 20 and 30-day exponential moving averages, relative strength index with a lag of 9 and 14 days, *PROC*, *CPACC*, *HPACC*, *ADO*, *STI*
J. Yao, C. Tan, H. Poh [12]	*KLCI* index	1984–1991 years, Malaysia	*KLCI* index with lags 1 and 2 days, 5, 10 and 50-day moving averages, momentum, Relative Strength Index, fast stochastic, slow stochastic
Y. Kara, M. Boyacioglu, Ö. Baykan [13]	*ISE* National 100 index	1997–2007 years, Turkey	10-day moving average and weighted moving average, momentum, fast stochastic, slow stochastic, Relative Strength Index, convergence/divergence moving averages, *MACD* indicator, Williams Percent Range, Chaikin Oscillator, accumulation/distribution indicator, Commodity Channel index
K. Kim [14]	The first difference of *KOSPI* index	1989–1998 years, South Korea	Momentum, fast stochastic, slow stochastic, rate of price change, Williams Percent Range, Accumulation / Distribution indicator, 5 and 10-day disparity, price oscillator, Commodity Channel Index, Relative Strength Index
E. Schöneburg [15]	Stock prices of *BASF*, Mercedes and Commerzbank	February–April 1989, Germany	Stock prices with lags from 1 to 10 days
V.N. Bugorskiy [16]	*RTS* index	2003–2008 years, Russia	Economic and political events that influenced *RTS* index price changes (for example, the arrest of Yukos head etc.)

The second problem of practical application of neural network modeling is related to the selection of neural network architecture. Basically, it is the selection of connection type between neurons, a number of hidden layers and number of neurons in hidden layers. The currently known number of connection types is finite and relatively small. Feed-forward neural networks are widely used for financial time series modeling that can approximate any measurable function given a sufficient number of hidden neurons [17]. To choose the number of hidden layers and the number of neurons, empirical rules are mostly used. For example, a large number of hidden layers will lead to neural network overfitting problem as it was shown in the paper [5]. For this reason, one or two hidden layers are sufficient in most cases. Furthermore, when the neural network with two hidden layers demonstrates unsatisfactory results, it is necessary to reanalyze input data before adding an additional hidden layer.

An excessive number of hidden neurons may also cause an overfitting problem [18]. This leads to an inability for the neural network to predict adequately any data other than a training sample.

There are numerous problems of neural network modeling practical application faced by researchers along with the above-mentioned problems. These are issues of data pre-processing, the selection of learning algorithm and learning parameters, and the selection of a criterion to choose the best model that is discussed in the methodological part of this research.

Support vector machine (*SVM*) [19] is an alternative method to a neural network that

is used to predict financial time series. Initially, *SVM* was used for classification, but later on, the regression support vector regression as the modification of the support vector machine was developed to predict financial time series. The detailed description of support vector machine is presented in the methodology section of this paper. Earlier studies [10; 20] have shown that overfitting is unlikely to occur with the support vector machine and the learning process is guaranteed to converge to a global minimum in comparison with a neural network. A risk of underfitting with *SVM* remains, but it is relatively small. It usually occurs when a large number of degrees of freedoms in the neural network [14] or very small values of learning algorithm parameters are used. Support vector machine is quite successfully applied to financial time series prediction, and the results of empirical studies demonstrate its higher predictive power in comparison with the neural network [10; 14].

The current paper continues the existing studies in the area of financial time series forecasting and examines the predictive power of neural network modeling and support vector machine. The methodology section of the research is based on the existing in modern literature expert knowledge about the factors that determine dynamics of financial time series, as well as theoretical and empirical rules for the implementation of the above-mentioned machine learning methods. The distinctive features of the study can be summed up as follows: we analyze a longer time period and a wide range of fundamental and technical indicators; we perform the sensitivity analysis of the obtained neural network modeling results to a high instability of the Russian stock market that previously found in the literature [8; 9].

Data

The data used in this study is the daily *MICEX* stock price index for the period from January 15, 2002, to April 25, 2016[1], and includes 3559 observations. For the analyzed time period, we collect information on such fundamentals as the *S&P500* index and *DAX* index, exchange rate for Russian ruble to US dollar and Brent crude oil price[2]. In addition, we use information from websites of Central Bank of Russia[3] and US Department of Treasury[4] on the dynamics of 1-month Moscow Interbank Offered Rate and 3-month Treasury Bill rates respectively. We select technical and fundamental indicators taking into account that they are frequently used in prior research; they have the ability to explain the behavior of individual indicators of Russian stock market that previously found in the literature [8; 9]; we are able to calculate them for the current research data.

The descriptive statistics of selected variables are presented in Table 2. Technical indicators include the following:

1) Average of *n*-th order:

$$\frac{1}{n}\sum_{i=0}^{n-1} C_{t-i}, \qquad (1)$$

where C_{t-i} is the closing price at time *t-i*, n − the smoothing interval length (in days);

2) Exponential Moving Average (*EMA*):

$$\alpha \cdot C_t + (1-\alpha) \cdot EMA_{t-1}, \qquad (2)$$

where α − the smoothing constant $\left(\alpha = \dfrac{2}{n+1}\right)$, the first value is $EMA_0 = C_0$;

3) Relative Strength Index (*RSI*):

$$100 - \frac{100}{(1+RS)}, \qquad (3)$$

where $RS = \dfrac{EMA(n)\ no\ U}{EMA(n)\ no\ D}$, U − the positive price changes, D − the negative price changes;

4) Momentum:

$$(C_t - C_{t-i}), \qquad (4)$$

where C_t is the closing price at time *t*, C_{t-i} is the closing price at time *t-i*;

5) Fast stochastic:

$$\frac{C_t - LL_n}{HH_n - LL_n}, \qquad (5)$$

where LL_n − the lowest price for *n* days, HH_n − the highest price for *n* days;

---

[1] Official site of "Finam". Available at: http://www.finam.ru/ (accessed 03.03.2017).

[2] Official site of "Finam". Available at: http://www.finam.ru/ (accessed 03.03.2017).

[3] Official site of Central Bank of Russian Federation. Available at: http://www.cbr.ru/ (accessed 03.03.2017).

[4] US Treasury Department official site. Available at: http://www.treasury.gov/ (accessed 03.03.2017).

6) Slow stochastic is calculated as a mean of fast stochastic values for *n* days.

The data is not cleared from the trends and outliers because a neural network is ro-

bust to them [21]. Normalization is used and described further.

Table 2

The definition of variables and descriptive statistics

Variable, unit	Mean	Median	Standard deviation	Min	Max
Dependent variable					
Price of *MICEX* index, pct.	1198.26	1392.40	494.35	250.71	1969.91
Technical variables					
5-day moving average, pct.	1197.31	1392.42	494.11	256.47	1952.63
10-day moving average, pct.	1196.13	1394.12	493.92	255.87	1931.15
15-day moving average, pct.	1194.96	1393.25	493.79	247.78	1921.16
5-day exponential moving average, pct.	1197.31	1393.04	493.92	256.47	1949.07
10-day exponential moving average, pct.	1196.14	1395.29	493.55	254.86	1932.84
15-day exponential moving average, pct.	1194.97	1394.88	493.23	249.79	1920.68
Momentum for 3 days, pct.	1.43	3.26	38.82	–286.68	224.56
Momentum for 6 days, pct.	2.83	5.97	54.49	–392.61	261.57
RSI for 9 days	88.91	92.70	16.59	–381.28	98.72
RSI for 14 days	88.51	91.70	10.01	–3.52	98.11
Fast stochastic for 14 days, %	58.94	64.46	30.93	0	100
Slow stochastic for 3 days, %	58.94	64.82	28.95	0	100
Fundamental variables					
Price of *S&P500* index, pct.	1349.26	1277	350.28	682.55	2130.82
Price of *NIKKEI255* index, pct.	12550.55	11389.35	3399.34	7054.98	20868.03
Price of *DAX* index, pct.	6441.98	6202.82	2213.98	2202.96	12374.73
Ruble/US Dollar exchange rate	33.13	30.36	10.77	23.13	83.59
Brent crude oil price, US Dollars	71.67	68.93	31.14	17.94	146.09
1-month Moscow Interbank Offered Rate, %	9.11	7.35	4.21	4.33	27.83
3-month Treasury Bill rates, %	1.27	0.235	1.61	–0.02	5.05

Note: pct. is percentage point.

Methodology

Neural network modeling

To predict *MICEX* index a feedforward neural network is used. It is trained using a backpropagation algorithm and that is also known as MultiLayer Perceptron (*MLP*). The main stages of neural network design to financial time series prediction are discussed, for example, in the papers [16; 18]. Hyperbolic tangent is used as the activation function of the neural network. We normalize the dependent variable and independent variables to the interval [-1; 1] that corresponds to the range of the activation function. The normalization is based on the following transformation [5]:

$$x_{normalized} = TF_{\min} + (TF_{\max} - TF_{\min})\frac{(x - x_{\min})}{(x_{\max} - x_{\min})}, (6)$$

where x – the initial value of variable, $x_{normalized}$ – the normalized value of variable, $TF_{\min}$ – the lower interval limit, i.e. –1, $TF_{\max}$ – the upper interval limit, i.e. 1, $x_{\min}$ – the maximum value of the initial variable, $x_{\max}$ – the minimum value of the initial variable.

To reduce the risk of overfitting problem the data is divided into training (60%) and testing (20%) samples [20]. Training and testing samples are used to find the optimal parameters of the learning algorithm and neural network architecture. Additionally, validation sample is used (20%) to evaluate the predictive performance of the neural network.

We use grid search as an empirical way to find the optimal learning parameters of neural network. For this purpose, the neural network is trained using different learning

parameters and the best neural network with minimum the Mean Squared Error (*MSE*) of the forecast for the test sample is chosen.

$$MSE = \frac{1}{T_1} \sum_{t=1}^{T_1} (y_t - \hat{y}_t)^2, \qquad (7)$$

where y_t – the *MICEX* actual price at time *t*, $\hat{y}_t$ – the predicted *MICEX* price at time *t*, T_1 – the length of the test sample.

The number of learning epochs (training iterations) is fixed at all training stages and equals to 800. This study varies the following learning parameters: the number of hidden layers (1 or 2), the learning speed (*LR*, Learning Rate) (from 0.1 to 0.9 [5]) and Momentum (*M*, (from 0 to 1 [22]).

LR controls the size of the steps for the method of gradient descent on the surface of a target error function (error surface). Too high value for this parameter can lead to missing of the minimum and too low value can slow down the training process. In this paper, the advanced gradient descent method with an additional parameter learning *M* is also used [22]. According to the paper [23], advanced method allows learning algorithm to converge faster to a minimum with a lower risk of failing into a local minimum.

For neural network modeling, Keras framework in Python language is employed. The obtained results for training and testing samples indicate that the optimal parameters of learning algorithm are *LR* = 0.09, *M* = 0.8 and the neural network topology consists of 1 hidden layer with 18 neurons. The value of *MSE* on the test sample equals to 0.0002.

According to the literature [8; 9], the Russian stock market is highly volatile. We test whether using financial time series data for a longer time period in neural network training leads to complication of data relationships detection by a neural network (due to a change in pre-existing data relationships, its disappearance, and appearance of new ones) and decrease in the predictive power. We train the neural network by using historical data for the 15-year period (2002–2016 years), the 10-year period (2007–2016 years) and the 5-year period (2012–2016 years). *MSE* value tends to increase with a decrease in the sample size and equals to 0.001, 0.015 and 0.052, respectively. For this reason, the historical data for the 15-year period is used and includes 3559 observations. It does not follow from the obtained results that we do not find support for the instability of the Russian stock market. It rather means that in this paper to increase the prediction accuracy of the neural network a large dataset should be used. We cannot, however, confirm and test that using longer than 15 years time period can lead to a decrease of the predictive performance as in the study [24].

Support vector machine

SVM[5] is a machine learning method that allows to transfer an input vector in multidimensional space and to evaluate a linear regression in this space. As well as neural network, *SVM* uses fewer data assumptions in comparison with classical regression models (for example, using data normalization is enough in most cases) and is robust to nonstationarity in the analyzed financial time series. This method also includes training and testing processes to find an optimal configuration of *SVM*. Typically, the training process is associated with quadratic programming problem that is discussed in the work [20].

Key training parameters are chosen experimentally based on the results of training and testing on the same samples that are used for neural network modeling. Data normalization is performed using the transformation (6). The main training parameters include parameter *C*, regularization parameter ε and kernel type. A kernel is a function allows transferring the input vector to *n*-dimensional space, where *n* is the number of observations. We use the Gaussian kernel with one parameter of width σ^2 that is widely applied in literature. Support vector machine has fewer adjustable learning parameters in comparison to a neural network that allows finding faster optimal values for them. In the case when we additionally find an optimal kernel type the search for optimal learning parameters does not become very complicated because the number of kernel type is finite. The development of algorithms for finding optimal learning parameters *C* and ε, as

---

[5] This term is commonly used in the literature. However, in this paper, we employ an extension of *SVM* for the regression – *SVR*, Support Vector Regression.

well as the kernel, remains an area that requires further investigation [20]. To implement *SVM* framework for machine learning in Python language Scikit-learn is used.

For selection of optimal learning parameters, grid search is applied. For this purpose, each training parameter is varied in the range from 10^{-3} to 10^3. The obtained results for training and testing samples indicate that the optimal parameters of learning algorithm are $C = 2$, $\varepsilon = 0.01$, $\sigma^2 = 0.001$. *MSE* value for testing sample is 0.0006.

Prediction performance

We compare prediction performance of neural network and support vector machine on validation sample based on two groups of metrics. The first group of metrics consists of Mean Squared Error (*MSE*) (7), Root Mean Squared Error (*RMSE*), Mean Absolute Error (*MAE*) (8) and Mean Absolute Percentage Error (*MAPE*) (9). *RMSE* increases faster when the share of large absolute errors is higher.

$$MAE = \frac{1}{T_2} \sum_{t=1}^{T_2} |y_t - \hat{y}_t|, \qquad (8)$$

$$MAPE = \frac{1}{T_2} \sum_{t=1}^{T_2} \frac{|y_t - \hat{y}_t|}{|y_t|}, \qquad (9)$$

where T_2 – the length of the validation sample.

The second group of metrics consists of coefficient of determination R^2 (10) and counted coefficient of determination cR^2 (11).

$$R^2 = 1 - \frac{\sum_{t=1}^{T_2} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{T_2} (y_t - \bar{y})^2}, \qquad (10)$$

$$cR^2 = \frac{TP + TN}{T_2}, \qquad (11)$$

where *TP* – the number of true positive cases, i.e. correctly predicted positive changes of the *MICEX* index, *TN* – the number of true negative cases, i.e. correctly predicted negative changes of the *MICEX* index, T_2 – the length of the validation sample.

$$cR^2 = \frac{1}{T_2} \sum_{t=1}^{T_2} \begin{cases} 1, if\ (y_t - y_{t-1})\ has\ the\ same\ sign\ as\ \hat{y}_t - y_{t-1}, \\ 0,\ otherwise. \end{cases} \qquad (12)$$

where y_{t-1} – the actual value of the *MICEX* index at time *t-1*.

The aim of forecasting is not only to achieve high average accurate prediction of *MICEX* index, but also to predict price change direction correctly. The second point is important, for example, for an investor who intends to trade futures, because the decision to purchase or sell the asset in this case is taken in advance. In other words, if the methods predict a rise in the asset price tomorrow, an investor will enter into a contract to purchase the asset today for the purpose of selling it tomorrow. If the actual price of the asset actually rises tomorrow, an investor will benefit, even if the growth is less than predicted. Otherwise, an investor will bear the loss.

In this context, the practical application of machine learning methods for predicting financial time series in order to make investment decisions should have a measure of a model ability to predict a direction of price movement. For example, we use counted coefficient of determination cR^2. The numerator in the formula (11) represents the total number of correctly predicted directions of *MICEX* price change. The direction change is a sign of the absolute change in the index price. If forecasted direction is equivalent to actual direction change, the observation equals to 1, and to 0 otherwise (12). Therefore, counting R^2 shows the percentage of correctly predicted *MICEX* price changes among all index changes. This indicator in itself is similar to an economic forecasting error, proposed in the paper [16], and Hit Ratio and Directional Symmetry (*DC*), presented in the papers [25] and [14], respectively.

Empirical results

Fig. 1 and Fig. 2 display that both neural network and support vector machine predictions demonstrate high prediction accuracy for validation sample. Predicted values of *MICEX* price do not deviate substantially from real values. Most of the data point on correlation fields (Fig. 3, Fig. 4) are located along a straight line at 45 degrees angle that confirms this finding.
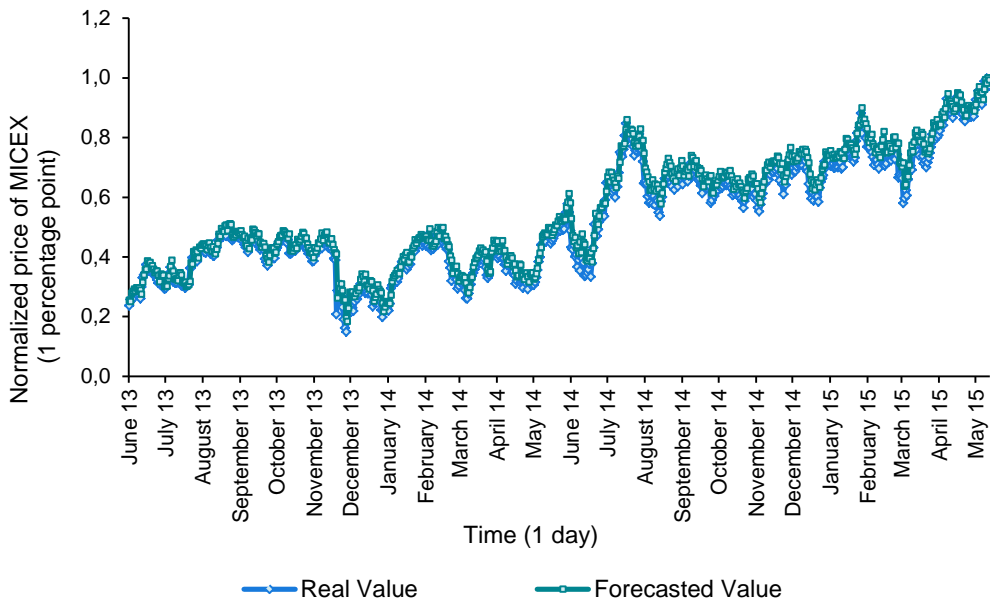
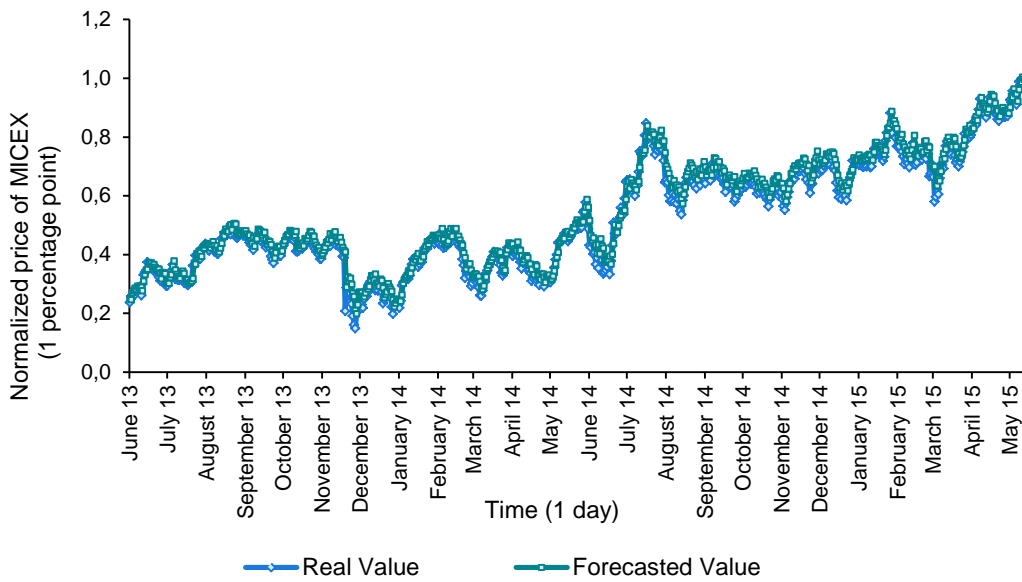**Fig. 1.** *MICEX* price predicted by neural network

Fig. 2. *MICEX* price predicted by support vector machine
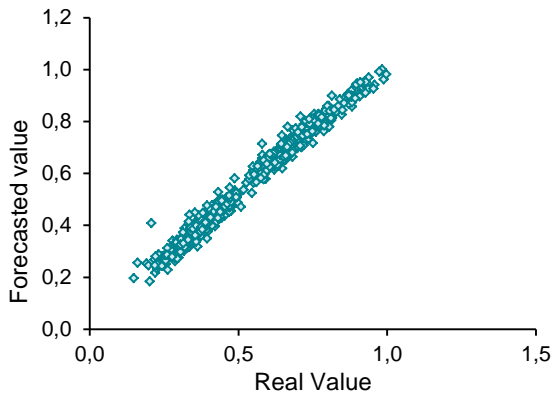

Fig. 3. Correlation field between forecasted by neural network *MICEX* price of and real *MICEX* price
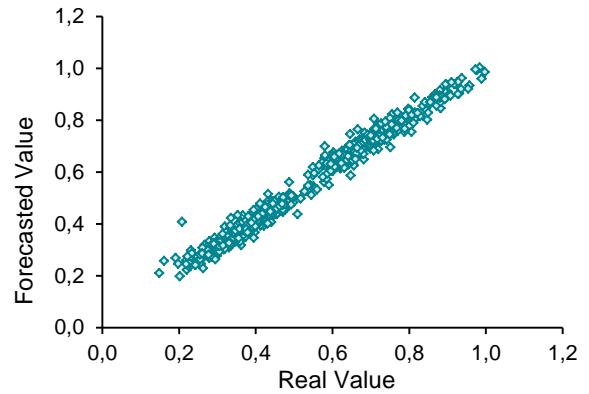
Fig. 4. Correlation field between forecasted by support vector machine *MICEX* price of and real *MICEX* price

57

To compare more precisely the prediction performance of the methods we have calculated several prediction metrics presented in Table 3.

Table 3

Prediction metrics for *MICEX* stock price index (validation sample)

Prediction metrics	Neural network	Support vector machine
MSE	0.0013	0.0009
RMSE	0.0358	0.0308
MAE	0.0287	0.0237
MAPE	0.0613	0.0514
R^2	0.9633	0.9728
cR^2	0.5099	0.5141

Firstly, *RMSE* value is close to *MAE* value that means we have few observations with a noticeable difference in predicted and real values. In this study, we obtain a lower *MAE* value compared with the results of L. Cao and F. Tay [25], who forecasted the absolute value of *S&P500* index. In their research, *MAE* value has a range (0.3496–0.6347) for a neural network and (0.3403–0.3706) for *SVM* on the testing sample.

Secondly, cR^2 equals to 0.51 and can be interpreted that the methods predict correctly only a half of all *MICEX* price direction changes. This result is comparable to the range of values for similar to count coefficient of determination prediction metrics (Hit Ratio [25] and Directional Symmetry (*DC*) [14]) obtained by previous studies. For example, Hit Ratio in forecasting of the first difference of *KOSPI* index in the study [25] ranged from 0.5198 to 0.5783 for *SVM*. *DC* for the testing sample in the study [14] varied from 0.392 to 0.4975 for a neural network and from 0.4623 to 0.4772 for SVM. Additionally, the papers [14; 25] point out that to reach high prediction accuracy of the direction change in stock price index is quite difficult.

In the case of investment in *MICEX* futures given $cR^2 = 0.51$, the strategy generates a near-zero profit. However, this problem requires further investigation by modifying in neural network modeling *MSE* function by reducing weights of neurons in case of incorrect direction change prediction and increasing them otherwise. For this purpose, new neural network architecture should be found.

Thirdly, *SVM* has higher prediction accuracy than neural network. This finding is consistent with the previous studies for the US stock market [10; 25] and Korean stock market [14]. Despite the higher prediction accuracy of *SVM*, the difference in corresponding prediction metrics is quite small and methods are based on various assumptions. *SVM* is robust to overfitting problem and always converges to the global minimum. Neural network has more variety in its design and modification that allows using it in conjunction with other forecasting tools [26]. The choice of the preferred prediction method depends on many factors, including computational costs.

Conclusion

This paper forecasts the Russian *MICEX* stock price index using neural network modeling and support vector machine. We use historical data of MICEX prices, as well as certain fundamental and technical indicators for the period 2002–2016 years. Keras and Scikit-learn frameworks in Python language are employed to perform computer experiments. Based on the analysis of certain prediction metrics, we conclude that both methods demonstrate high predictive power. Support vector machine outperforms neural network, but the difference in predictive metrics is not substantial. This finding is relevant to previous studies for different international financial markets. Further research includes the development of a methodology for filtering and transformation input data with using additional machine learning algorithms, and the development of guidelines for a choice of a trading strategy based on the machine learning results.

References

1. Malkiel B. The efficient market hypothesis and it's critics. *The Journal of Economic Perspectives*, 2003, vol. 17, pp. 59–82.

2. White H. Economic prediction using neural networks: the case of IBM daily stock returns. *IEEE International Conference on Neural Networks*, 1988, vol. 2, pp. 451–458.

3. Jammazi R., Aloui C. Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, 2012, vol. 34, pp. 828–841.

4. Valiotti N.A., Abbakumov V.L. Kolichestvennoe otsenivanie posledstvii upravlencheskikh reshenii na osnove neirosetevykh modelei [Quantitative evaluation of the effects of management decisions on the basis of neural network models]. *Prikladnaia informatika* [Applied Informatics], 2013, no. 5 (47), pp. 6–13. (In Russian).

5. Kaastra I., Iebeling K., Milton B. Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 1996, vol. 10, no. 3, pp. 215–236.

6. Dayhoff J.E. *Neural network architectures: An introduction*. New York, Van Nostrand Reinhold Publ., 1990. p. 259.

7. Pakath R., Ramakrishnan P., Zaveri J.S. Specifying critical inputs in a genetic algorithm-driven decision support system: An automated facility. *Decision Sciences*, 1995, vol. 26, no. 6, pp. 749–771.

8. Anatol'ev S. A 10-year retrospective on the determinants of Russian stock returns. *Research in International Business and Finance*, 2008, vol. 22, no. 1, pp. 56–67.

9. Peresetsky A.A. What determines the behavior of the Russian stock market. *MPRA paper 41508*, University Library of Munich Publ., 2011. Available at: https://mpra.ub.uni-muenchen.de/41508/1/MPRA_paper_41508.pdf (accessed 17.02.2017).

10. Tay F.E.H., Lijuan, C. Application of support vector machines in financial time series forecasting. *Omega*, 2001, vol. 29, no. 4, pp. 309–317.

11. Zhang Y., Yudong Z., Lenan W. Stock market prediction of S&P500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications*, 2009, vol. 36, no. 5, pp. 8849–8854.

12. Yao J., Tan C.L., Poh, H.L. Neural networks for technical analysis: a study on KLCI. *International Journal of Theoretical and Applied Finance*, 1999, vol. 2, no. 02, pp. 221–241.

13. Kara Y., Boyacioglu M.A., Baykan Ö.K. 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*, 2011, vol. 38, no. 5, pp. 5311–5319.

14. Kim K.J. Financial time series forecasting using support vector machines. *Neurocomputing*, 2003, vol. 55, no. 1, pp. 307–319.

15. Schöneburg E. Stock price prediction using neural networks: A project report. *Neurocomputing*, 1990, vol. 2, no. 1, pp. 17–27.

16. Bugorskiy V.N., Sergiyenko A.G. Neirosetevoe modelirovanie faktorov, vliiaiushchikh na volatil'nost' tsennykh bumag [Neural network modeling of factors affecting the volatility of securities]. *Prikladnaia informatika* [Applied Informatics], 2009, no. 5 (23), pp. 16–26. (In Russian).

17. Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators. *Neural Networks*, 1989, vol. 2, no. 5, pp. 359–366.

18. Iasnitskii L.N. *Intellektual'nye sistemy* [Intelligent systems]. Moscow, Laboratoriia Znanii Publ., 2016. 221 p. (In Russian).

19. Cortes C., Vapnik V. Support-Vector Networks. *Machine Learning*, 1995, vol. 20, no. 3, pp. 273–297.

20. Haykin S. *Neironnye seti: polnyi kurs* [Neural networks: A comprehensive foundation]. Moscow, OOO «I.D. Willams» Publ., 2006. 1103 p.

21. Trippi R.R., Turban E. *Neural networks in finance and investing: Using artificial intelligence to improve real world performance*. New York, McGraw-Hill Inc., 1992, p. 400.

22. Sapuan S.M., Mujtaba I.M. *Composite materials technology: neural network applications*. Berlin, CRC Press, 2009, p. 368.

23. Qian N. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 1999, vol. 12, no. 1, pp. 145–151.

24. *Walczak S.* An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of Management Information Systems*, 2001, vol. 17, no. 4, pp. 203–222.

25. Cao L., Tay, F.E. Financial forecasting using support vector machines. *Neural Computing & Applications*, 2001, vol. 10 (2), pp. 184–192.

26. Evans C., Pappas K., Xhafa F. Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. *Mathematical and Computer Modelling*, 2013, vol. 58, pp. 1249–1266.

Information about the Authors

Lozinskaia Agata Maksimovna – Candidate of Economic Sciences, Senior Lecturer of the Department of Economics and Finance, Research Fellow of the Research Group for Applied Markets and Enterprises Studies of the Laboratory of Interdisciplinary Empirical Studies, National Research University Higher School of Economics (27, Lebedeva st., Perm, 614070, Russia; e-mail: AMPoroshina@gmail.com).

Zhemchuzhnikov Viktor Andreevich – Undergraduate Student of Academic program "Economics", National Research University Higher School of Economics (27, Lebedeva st., Perm, 614070, Russia; e-mail: viktr5909@mail.ru).

Сведения об авторах

Лозинская Агата Максимовна – кандидат экономических наук, старший преподаватель департамента экономики и финансов, научный сотрудник сектора эмпирического анализа рынков и компаний научно-учебной лаборатории междисциплинарных эмпирических исследований, Национальный исследовательский университет «Высшая школа экономики» (Россия, 614070, г. Пермь, ул. Лебедева, 27; e-mail: AMPoroshina@gmail.com).

Жемчужников Виктор Андреевич – студент 4-го курса образовательной программы бакалавриата «Экономика», Национальный исследовательский университет «Высшая школа экономики» (Россия, 614070, г. Пермь, ул. Лебедева, 27; e-mail: viktr5909@mail.ru).