USE OF ARTIFICIAL NEURAL NETWORKS IN THE CAPITAL MARKETS

[a]MARTA URBANÍKOVÁ, [b]MICHAELA ŠTUBŇOVÁ

Institute of Economics and Management, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovak Republic
email: [a]murbanikova@ukf.sk, [b]mstubnova@ukf.sk

Abstract: Methods of stock market analysis have undergone long-term development. In particular, stock market failures and financial crises were the driving forces behind the search for new methods and forecasts. We are currently living in a time of digital boom that has greatly helped advance science in all areas. This fact was also reflected in the area of trading on capital markets. Artificial intelligence methods are at the forefront. The paper aims to analyse the predictive ability of artificial neural networks with the application on capital markets. The results of the analysis suggest that higher prediction accuracy is achieved by networks with a lower number of time delays.

Keywords: Capital markets, Prediction, Artificial neural networks, Nonlinear autoregressive network

1 Introduction

With Brexit approaching and a slowdown in the growth of several economies, investors are increasingly motivated to analyse capital markets. In addition to fundamental and technical analysis, attention is paid to statistical analyses of time series and artificial intelligence. Methods of stock market analysis have undergone long-term development. At present, fundamental and technical analysis are the most developed, while psychological analysis is less recognized. Time-series analyses and artificial intelligence methods are also at the forefront. Fundamental analysis seeks to quantify the intrinsic value of stocks. It deals with the basic factors of global, political, economic, sectoral and corporate character, which significantly influence the exchange rate or the intrinsic value of the stock. The fundamental analysis is carried out at three levels: the macroeconomic level, which analyses the economy as a whole and its impact on stock prices; the sectoral level that examines industry specificities and their impact on equity rates; the microeconomic level that conducts a financial analysis of individual companies. Technical analysis, unlike fundamental analysis, focuses only on the capital market. By examining past and current information on volumes and movements in share prices, it seeks to predict the development of individual stocks or the overall market development. The technical analysis leads to much faster decisions. Like the fundamental analysis, it applies a structural approach. This means that it applies first to the whole market, which is mostly represented by a market index, but also to the relevant sector, where it mainly uses sectoral indices and the overall sector's share of the economy, and finally also analyses individual stocks. Technical analysis consists of an analysis of stock price graphs and an analysis of technical indicators (Momentum, MACD, RSI, SSTO, ADX, CCI, W%R, the combination of moving averages, etc.), which predict the evolution of prices shares in the future. Methods of time series analysis can be divided into two groups. One group consists of methods where the development of the modelled indicator depends only on time. These are decomposition methods, Box-Jenkinson models and spectral analysis. The second group consists of econometric modelling methods. Frequently used models of financial time series analysis include Auto-Regressive Moving Average (ARMA) model, Auto-Regressive Integrated Moving Average (ARIMA) model, Auto-Regressive Conditional Heteroscedasticity (ARCH) model, and Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH) model.

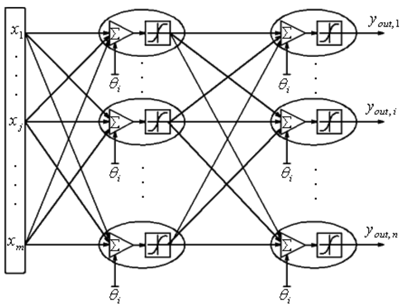
Using artificial neural networks (ANNs), it is possible to model complex high-dimensional data that we cannot describe by simple rules and classical statistical methods. Neural network modelling of systems is highly abstract, the same network can be a model for different systems without changing the structure if different data are used to learn the network. Neural networks, genetic algorithms, and fuzzy logic theory significantly contribute to the development of data analysis and modelling. They are important tools for decision support. Creation of the neural network itself, starting with data preparation, selecting the number of layers, selecting the number of neurons in each layer, choosing training methods, activation functions, etc. is still a matter of exploration and experimentation. The neural networks are gaining a lot of attention among the scientific community. One of the most important features of neural networks is their ability to universally approximate functions. It often happens that we have a very complex system whose description is almost impossible or would require too much computing time. However, we have data that enters the system and corresponding outputs. In this situation, we can use a suitable neural network that learns to behave as a monitored system. This is an important factor that determines the application of neural networks in capital markets.

Researchers are currently paying close attention to the use of soft computing methods to predict financial time series. A comprehensive overview of the published articles dealing with stock market prediction and portfolio selection models can be found in Rather, Sastry, and Agarwal (2017). The authors focused on articles in which traditional mathematical models were used and articles in which artificial intelligence-based models were used. In the study Calvacante, Brasileiro, Souza, Nobrega, and Oliveira (2016), the authors provided a comprehensive overview of articles on computational intelligence methods in financial applications, dealing with researches in preprocessing and clustering of financial data, forecasting future market movements and mining financial text information. Comparative studies of forecast accuracy of ANN and ARIMA models were conducted in papers Adebiyi, Adewumi, and Ayo (2014), Cocianu and Grigoryan (2015), Safi and White (2017), Tiwari, Bharadwaj, and Gupta (2017), reported results showed higher accuracy of ANN predictions. In the study Lahmiri (2017) a simple ANN forecasting model trained with technical analysis indicators was proposed. Forecast accuracy of the proposed model to forecast exchange rates volatilities was compared with GARCH family models and hybrid GARCH and ANN model. The choice of a suitable ANN training algorithm for networks predicting financial time series was studied in the paper Chandar, Sumathi, and Sivanandam (2015). The authors investigated the performance of ANN models with five learning algorithms and they reported high prediction accuracy of ANN models with Levenberg-Marquart training algorithm. Several authors focused on using hybrid models to increase the accuracy of their predictions. High prediction accuracy for nonlinear data was achieved with a hybrid model consisting of ARMA model, exponential smoothing (ES) model and recurrent neural network (RNN) (Rather, Agarwal, and Sastry, 2015). The hybrid model based on Support Vector Machine, Radial Basis Function, and two ANN's variants Single Layer Perceptron and Multi-layer Perceptron were proposed by Usmani, Adil, Ebrahim, and Raza (2018). The authors reported, that the combination of four machine learning techniques, is a suitable tool for forecasting the stock market behaviour. The aim of this paper is to predict stock indices using dynamic neural networks and to analyse the predictive ability of networks with different architecture and a different number of time delays.

2 Artificial neural network

ANNs, one of the soft computing methods, represent a massive parallel computing system capable of processing data containing inaccuracies, uncertainties, and noise. Pattern recognition, data recognition, image processing, stock market prediction, weather prediction, image compression, and security and loan applications were referred to neural network applications in Ibrahim (2016). The inspiration for the ANN is the human brain and its ability to learn. Learning capability enables neural networks to improve their performance and achieve accurate results. ANN consists of neurons that are simplified models of biological neurons with similar structures and functions. Neurons process the information with an activation function and transmit it via oriented weighted connections. Connected neurons are organized into layers – input layer, hidden layer, and output layer. They differ in the sources of their inputs and in the use of their outputs. The input layer processes the data of the independent variables that are inputs to the ANN and transmits it to the next network layer. Hidden layers process outputs from previous layers and transmit them to the next layer. The output layer processes the outputs of the previous hidden layer and gives the value of the dependent variable as an output (Parot, Michell, and Kristjanpoller, 2019, Yildirim, 2019). A multilayer feedforward neural network with input layer, one hidden layer, and output layer is shown in Figure 1. The signal proceeds through the network through directed connections in one direction – forward. Dynamic networks, unlike static feedforward networks, use not only the current inputs of the network but also the previous inputs and outputs to calculate the output of the network. These feedback connections allow the network to store information for further use (Mahmud and Meesad, 2016).

Figure 1: Architecture of the multilayer neural network.



Source: Elaborated by the authors.

The learning process of an ANN is done by adjusting the connection's weight values and neuron threshold values (θ). The adjusted weights of the neural network connections store acquired knowledge. Several learning laws can be applied to the neural network training process. In this study, we use Levenberg-Marquart learning algorithm, which is a variation on Backpropagation algorithm. In Li, Cheng, Shi, and Huang (2012) the Backpropagation algorithm was described in two steps. In the first step, the operating signal is propagated forward through the network layers. The error signal is computed as the difference between the real and the expected output. In the second step, the error signal is back propagated through the network. The weight values and threshold values are adjusted using the gradient descent method to minimize the error signal and therefore optimize the network performance. The disadvantage of the Backpropagation algorithm is its slow speed of convergence. To speed up the convergence, there are several variations of Backpropagation such as Levenberg-Marquart, Scaled Conjugate Gradient Descent, Quasi-Newton, One Step Secant, Variable Learning Rate Backpropagation, etc. For ANN with less than a few hundred weights, Levenberg-Marquart is the fastest converging learning algorithm (Shahbazi, Memarzadeh, and Gryz, 2016). In the study Chandar, Sumathi, and Sivanandam (2015), the best prediction of the foreign currency exchange rate was achieved with ANN trained with Levenberg-Marquart learning algorithm.

The Levenberg-Marquart learning algorithm combines the gradient descent method and the Gauss-Newton method. The non-linear function is minimized with a numerical solution (Gavin, 2019). According to Yu and Wilamowski (2011), the learning rule of Levenberg-Marquart learning algorithm is given by:

$$w_{k+1} = w_k - (\boldsymbol{J}_k^T \boldsymbol{J}_k + \mu \boldsymbol{I})^{-1} \boldsymbol{J}_k \boldsymbol{e}_k, \tag{1}$$

where w_{k+1} and w_k are components of wight vector $\boldsymbol{w}$; μ is a combination coefficient with positive value; $\boldsymbol{I}$ is the identity matrix; $\boldsymbol{e}_k$ is a vector of training errors defined as $\boldsymbol{e}_k = y_k - \widehat{y_k}$, where y_k are target values and $\widehat{y_k}$ are output values; $\boldsymbol{J}$ is the Jacobian matrix defined as:

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \cdots & \frac{\partial e_{1,1}}{\partial w_N} \\ \frac{\partial e_{1,2}}{\partial w_1} & \frac{\partial e_{1,2}}{\partial w_2} & \cdots & \frac{\partial e_{1,2}}{\partial w_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_{1,M}}{\partial w_1} & \frac{\partial e_{1,M}}{\partial w_2} & \cdots & \frac{\partial e_{1,M}}{\partial w_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_{P,1}}{\partial w_1} & \frac{\partial e_{P,1}}{\partial w_2} & \cdots & \frac{\partial e_{P,1}}{\partial w_N} \\ \frac{\partial e_{P,2}}{\partial w_1} & \frac{\partial e_{P,2}}{\partial w_2} & \cdots & \frac{\partial e_{P,2}}{\partial w_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_{P,M}}{\partial w_1} & \frac{\partial e_{P,M}}{\partial w_2} & \cdots & \frac{\partial e_{P,M}}{\partial w_N} \end{bmatrix}, \tag{2}$$

where N is a number of weights, M is a number of outputs and P is a number of patterns.

Considering the dependence of the financial time series values on its previous values, it is appropriate to use dynamic neural networks for its prediction. In this study, a nonlinear autoregressive network (NAR) is used. It is a recurrent dynamic network based on a linear autoregressive model (AR). The architecture of NAR network consists of feedback connections and tapped delay lines (TDL). The TDL unit returns a vector consisting of input data from the current time-step to the required number of time delays (Stokes and Abou-Zaid, 2012). When using supervised learning, the actual previous values of the time series are known and can be used to replace the feedback connections. After modification of the network to feedforward architecture, static Backpropagation in the network training process can be used, which reduces the time required for calculations. After training of the ANN, it is possible to iterate the prediction of several time steps ahead using the original network architecture containing the feedback connections. Leaving a purely feedforward architecture, the network would be able to predict only one-step-ahead. Equation (3) is the definition of the NAR model (Shahbazi, Memarzadeh, and Gryz, 2016).

$$y(k) = F(y(k-1), y(k-2), \dots, y(k-d)\,), \tag{3}$$

where $y(k)$ is the value of the financial time series at time k and d is a number of time delays.

3 Methodology and data

The analysis is conducted by examining the accuracy of the prediction of the stock index Financial Times Stock Exchange 100 Index (FTSE 100). The stock index consists of 100 companies with the highest market capitalisation listed on the London Stock Exchange. The robust model is developed by using historical prices from January 2, 2009, to August 30, 2019, which represents 2693 values. The tracking period is shown in Figure 2.

Figure 2: FTSE 100 prices from January 2, 2009, to August 30, 2019.

Source: Elaborated by the authors based on data from https://www.investing.com/indices/uk-100-historical-data.

To simplify the problem of the outliers in the network, normalisation of the network's input data was proposed by Shahbazi, Memarzadeh, and Gryz (2016). We normalise the prices of FTSE 100 stock index according to Markechová, Stehlíková, and Tirpáková (2011):

$$U = \frac{X-\mu}{\sigma}, \tag{4}$$

where U is the normalized variable, X is the original variable, μ is the mean value of the variable and σ is the standard deviation of the variable. After pre-processing, we divide the data into training set, validation set and testing set in a ratio 70:20:10.

The models are constructed in the MATLAB R2019b software. To predict the stock index, we develop and validate NAR networks with tansig transfer function and Levenberg-Marquadt learning algorithm. The network architecture is optimized using different numbers of the hidden layer's neurons and different number of time delays. We construct networks with the time delay from 1 to 10 days and 5, 10, 15, 20, 25, 30 neurons in the hidden layer for each value of time delay. We construct, train, validate and test 60 neural networks. To evaluate the network performance, it is possible to use the determination coefficient (R^2), the mean square error (MSE), the root mean square error ($RMSE$), the mean absolute error (MAE) and the mean absolute percentage error ($MAPE$).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{5}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}, \tag{7}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{8}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \tag{9}$$

where n is the number of data points, y_i are target values, $\hat{y}_i$ are output values and $\bar{y}$ is the mean of target values.

In the study, we use the MSE as a measure of the quality of network performance. The MSE includes the variance of the predictor and its bias.

4 Results and discussion

For constructed neural networks, the MSE values are calculated for training, validation and test data set and for the whole data set. When selecting the network with the best prediction ability, we take into account the MSE values for the validation data set as proposed by Moghaddam, Moghaddam, and Esfandyari (2016). The networks with the most accurate predictions for each value of time delay are listed in Table 1.

Tab. 1: MSE values for constructed networks with different architecture

Time delays	Hidden neurons	MSE			
		Train	Valid	Test	All
1	20	6.13393E-03	4.28847E-03	3.95746E-03	5.54718E-03
2	25	6.24435E-03	3.67726E-03	4.09539E-03	5.51659E-03
3	20	6.25611E-03	3.76361E-03	4.26495E-03	5.55876E-03
4	5	6.27595E-03	3.89500E-03	4.70441E-03	5.64261E-03
5	30	5.88020E-03	4.27361E-03	5.23377E-03	5.49410E-03
6	5	6.12434E-03	4.21384E-03	4.56855E-03	5.58627E-03
7	20	6.09874E-03	4.18674E-03	4.48561E-03	5.55513E-03
8	30	5.78057E-03	4.22786E-03	5.72543E-03	5.46463E-03
9	20	5.56488E-03	4.05982E-03	4.25169E-03	5.13231E-03
10	5	6.11457E-03	4.07677E-03	4.75840E-03	5.57145E-03

Source: Elaborated by the authors.

The minimum MSE value for the validation data set has the NAR network with 2 days time delay and 25 neurons in the hidden layer. The data in Table 1 suggests that higher prediction accuracy is achieved by networks with a lower number of time delays. Setting the optimal network architecture and number of time delays needs to be adapted to each financial time series.

5 Conclusion

Neural networks are often referred to as a black box because we cannot know the internal structure of the system in detail. We put only a few assumptions on the internal structure of the system, which is modelled by the black box, which will make it possible to describe the behaviour of the system by a function that does the input-output transformation. Neural networks are useful when coincidence plays a significant role in the modelling process and where deterministic dependencies are so complex and interconnected that we cannot separate and analytically identify them. They are therefore suitable for modelling complex, often irreversible, strategic decisions. The characteristics of ANNs predetermine them for the analysis of financial time series, which is also confirmed by the results of the performed analysis. The dynamic NAR network with a lower number of time delays is proven to be a suitable tool for predicting the FTSE 100 stock index.

Literature:

1. Rather AM, Sastry VN, Agarwal A. Stock market prediction and Portfolio selection models: a survey. OPSEARCH. 2017; 54(3): 558-579. eISSN 0975-0320 doi: 10.1007/s12597-016-0289-y
2. Calvacante RC, Brasileiro RC, Souza VLF, Nobrega JP, Oliveira ALT. Computational Intelligence and Financial Markets: A Survey and Future Directions. Expert Systems With Applications. 2016; 55: 194-211. ISSN 0957-4174 doi: 10.1016/j.eswa.2016.02.006
3. Adebiyi AA, Adewumi AO, Ayo CK. Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. Journal of Applied Mathematics. 2014: 1-7. eISSN 1687-0042 doi: 10.1155/2014/614342
4. Cocianu C, Grigoryan H. An Artificial Neural Network for Data Forecasting Purposes. Informatica Economică. 2015; 19(2): 34-45. ISSN 1453-1305 doi: 10.12948/issn14531305/19.2.2 015.04
5. Safi S, White A. Short and long-term forecasting using artificial neural networks for stock prices in Palestine: A comparative study. Electronic Journal of Applied Statistical Analysis. 2017; 10(1): 14-28. e-ISSN 2070-5948 doi:10.1285/i20705948v10n1p14
6. Tiwari S, Bharadwaj A, Gupta S. Stock Price Prediction Using Data Analytics. In 2017 IEEE International Conference on Advances in Computing, Communication and Control (ICAC3); 2017; Mumbai; Mumbai: Fr Conceicao Rodrigues Coll Engn; 2017. p. 1-5. eISBN 978-1- 5386-3852-1 doi: 10.1109/ICAC3 .2017.8318783
7. Lahmiri S. Modeling and predicting historical volatility in exchange rate markets. Physica A: Statistical Mechanics and its Applications. 2017; 471: 387-398. ISSN 0378-4371 doi: 10.1016/j.physa.2016.12.061
8. Chandar SK, Sumathi M, Sivanandam SN. Forecasting of Foreign Currency Exchange Rate Using Neural Network. International Journal of Engineering and Technology. 2015; 7(1): 99-108. ISSN 0975-4024
9. Rather AM, Agarwal A, Sastry VN. Recurrent neural network and a hybrid model for prediction of stock returns. Expert Systems with Applications. 2015; 42(6): 3234-3241. ISSN 0957-4174 doi: 10.1016/j.eswa.2014.12.003
10. Usmani M, Adil SH, Ebrahim M, Raza K. Predicting Market Performance with Hybrid Model. In 2018 3RD INTERNATIONAL CONFERENCE ON EMERGING TRENDS IN ENGINEERING, SCIENCES AND TECHNOLOGY (ICEEST); 2018; Karachi;Karachi: IQRA University, 2018. p. 1-4. eISBN 978-1-5386-8249-4 doi: 10.11 09/ICEEST.2018.8643327

11. Ibrahim D. An overview of soft computing. Procedia Computer Science. 2016; 102: 34-38. doi: 10.1016/j.procs.2016.09.366
12. Parot A, Michell K, Kristjanpoller WD. Using Artificial Neural Networks to forecast Exchange Rate, including VAR-VECM residual analysis and prediction linear combination. Intelligent System in Accounting Finance and Management. 2019; 26(1): 3-15. ISSN 1099-1174 doi: 10.1002/isaf.1440
13. Yildirim H. Property Value Assessment Using Artificial Neural Networks, Hedonic Regression and Nearest Neighbors Regression Methods. Selcuk University Journal of Engineering, Science and Technology. 2019; 7(2): 387-404. ISSN 2147-9364 doi: 10.15317/Scitech.2019.207
14. Mahmud MS. Meesad P. An innovative recurrent error-based neuro-fuzzy system with momentum for stock price prediction. Soft Computing. 2016; 20(10): 4173-4191. ISSN 1433-7479 doi: 10.1007/s00500-015-1752-z
15. Li J, Cheng J, Shi J, Huang F. Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement. In: Jin D, Lin S, editors. Advances in Computer Science and Information Engineering. Advances in Intelligent and Soft Computing. vol. 169, Berlin: Springer, Berlin, Heidelberg; 2012. p. 553-558. ISBN 978-3-642-30223-7 doi:10.1007/978-3-642-30223-7_87
16. Shahbazi N, Memarzadeh M, Gryz J. Forex Market Prediction Using NARX Neural Network with Bagging. MATEC Web of Conferences. 2016; 68. doi: 10.1051/matecconf/20166819001
17. Gavin HP. The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. Department of Civil and Environmental Engineering, Duke University. 2019: 1-15
18. Yu H, Wilamowski BM. Levenberg-Marquardt training. In Wilamowski BM, Irwin JD, editors. The Industrial Electronics Handbook, Intelligent Systems. Boca Raton: CRC Press, Taylor & Francis Group; 2011. p. 12-1–12-16. ISBN: 9781439802830 doi: 0.1201/9781315218427-12
19. Stokes A, Abou-Zaid A. Forecasting foreign exchange rates using artificial neural networks: A trader's approach. International Journal of Monetary Economics and Finance. 2012; 5(4): 370-394. ISSN 1752-0479 doi: 10.1504/IJMEF.2012.052502
20. Fusion Media Limited. FTSE 100 (FTSE) [Internet]. Fusion Media Limited; 2019 [updated 2019 Sep 13; cited 2019 Sep 13] . Available from: https://www.investing.com/indices/uk-100-historical-data
21. Markechová D, Stehlíková B, Tirpáková A. Štatistické metódy a ich aplikácie. Nitra: Univerzita Konštantína Filozofa v Nitre; 2011. 534 p. ISBN 978-80-8094-807-8
22. Moghaddam AH, Moghaddam MH, Esfandyari M. Stock market index prediction using artificial neural network. Journal of Economics, Finance and Administrative Science. 2016; 21(41): 89-93. ISSN 2077-1886 doi: 10.1016/j.jefas.2016.07.002

Primary Paper Section: A

Secondary Paper Section: AH