

Forecasting returns on a stock market using Artificial Neural Networks and GARCH family models: Evidence of stock market S&P 500

Nadhem Selmi*, Samira Chaabene and Nejib Hachicha

Department of Quantitative Methods, Faculty of Economics and Management of Sfax – Tunisia

CHRONICLE

ABSTRACT

Article history:

Received July 10, 2014
Accepted December 17, 2014
Available online
December 18 2014

Keywords:

Artificial neural networks
Stock market forecasting
Non-linear modeling
GARCH
Forecasts

In the area of financial stock market forecasting, many studies have focused on application of Artificial Neural Networks (ANNs). Due to its high rate of uncertainty and volatility, the stock markets returns forecasting by conventional methods became a difficult task. The ANNs is a relatively new and have been reported as good methods to predict financial stock market levels and can model flexible linear or non-linear relationship among variables. The aim of the study is to employ an ANN models to estimate and predict the dynamic volatility of the daily of S&P500 market returns. Results of ANN models will be compared with time series model using GARCH family models. The use of the novel model for conditional stock markets returns volatility can handle the vast amount of nonlinear data, simulate their relationship and give a moderate solution for the hard problem. The forecasts of stock index returns in the paper will be evaluated and compared, considering the MSE, RMSE and MAE forecasts statistic.

© 2015 Growing Science Ltd. All rights reserved.

1. Introduction

Stock market prediction (or forecasting) is one of the hottest fields of research lately due to its commercial applications owing to the high stakes and the kinds of attractive benefits that it has to offer (Majhi et al., 2007). In addition, it is regarded as a challenging task of financial time-series prediction and one of the important issues in finance. The stock market, which has been investigated by various researches, is essentially non-linear, dynamic, nonparametric, and is rather complicated environment (Tan et al., 2005). Miao et al. (2007) and Wang (2002) presented that stock market's movements are affected by many factors such as political events, firms' policies, general economic conditions, commodity price index, bank rate, bank exchange rate, investors' expectations, institutional investors' choices, movements of other stock market, psychology of investors, etc. The explanation and the understanding adequate of the stock markets return volatility constitutes a central to the study of finance. Investigating the data generating process in stock returns, recent research has suggested that the factors affecting the asset pricing behavior of investors are described by nonlinear relationships with expected returns. The predicate of stock markets returns have exhibited a strong nonlinearity

* Corresponding author.

E-mail address: nadhem.selmi@yahoo.fr (N. Selmi)

induced from an asymmetric dynamic process (Nam et al., 2002; Nam & Kim, 2003)¹. Forecasting of stock market returns in today's volatile markets became a difficult task and represents a major challenge for traditional time-series predicting. Therefore, predicting finance and economics movements is quite difficult. In the literature there are a number of methods applied to describe the behavior of time series and to accomplish this task.

In the last decade, in addition to the advent of the advancement in computer technology many studies have been included soft computing techniques. ANNs² have been popularly implemented to many areas of disciplines financial type problems such as stock markets index prediction, exchange rate prediction, bankruptcy prediction etc. This technique of soft computing can present better forecasting ability on stock market than other models, especially in nonlinear questions (Phua et al., 2003; O'Connor & Maddem, 2006; Kunhuang & Yu, 2006; Zhu et al., 2007). ANN models have been proposed as a promising alternative method to time-series forecasting. Several studies have shown that ANN could be a very useful tool for time series modeling and forecasting (Zhang et al., 1998). However this novel model of soft computing has the capability to determine the underlying functional relationships in the data and flexibly generate the forecasting models regardless of the complexity of system.

There have been various studies using ANNs in stock market prediction. These studies were performed during the past decade. One of the earliest studies, Kamijo and Tanigawa (1990) applied recurrent neural networks and Ahmadi (1990) used back propagation neural networks with the generalized delta rule to predict the stock market. Trippi and DeSieno (1992) and Choi et al., (1995) forecasted daily direction of change in the S&P500 index futures using ANNs. Choi et al. (2006) and Trippi and DeSieno (1992) predicted the daily direction of change in the S&P500 index futures using ANN. Trafalis (1999) used feed-forward ANN to forecast the change in the S&P500 index. These studies and others³, who are mainly concentrated on the application of ANNs to stock market prediction, proved neural networks to be a superior technique for forecasting in their test cases than other models, especially in nonlinear questions.

The purpose of this paper is to employ ANN approaches to predict the future stock market returns of North American S&P500. In addition, in this study this ANN model will be compared with conventional forecasting models of GARCH⁴ family. It is in fact to investigate the effectiveness of ANN for time-series modeling and forecasting. The rest of this paper is organized as follows. In section 2, the GARCH and ANN approaches are described respectively. Section 3 presents the empirical results. In section 4, conclusions and the intended direction of future research are presented.

2. Econometric Methodology

2.1. ARCH Family of Models

Following the autoregressive conditional heteroskedasticity (ARCH) model by Engle (1982), there are many extensions of ARCH model, such as GARCH, GARCH-M, EGARCH etc.; we can call them ARCH family models. ARCH process can be indicated as follows,

$$y_t = \beta y_{t-1} + \varepsilon_t \quad (1)$$

$$\varepsilon_t = \sqrt{h_t} \cdot v_t \quad (2)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (3)$$

where

¹ They identify nonlinearities induced from an asymmetric reverting property of weekly and monthly returns for U.S and Pacific Basin equity markets.

² ANN: artificial neural networks.

³ Such as Chang et al (2004); Hamid and Iqbal (2004); Liu et al (2003- 2006), Chen et al (2006) and Zhu et al (2007).

⁴ GARCH: Generalized Autoregressive Conditional Heteroscedasticity.

$$v_t \text{ iid } N(0,1), \alpha_0 > 0, \alpha_i \geq 0 \text{ et } \sum_{i=1}^q \alpha_i < 1 \quad (4)$$

This model is also recognized as the linear ARCH (q) model. With financial data it captures the information from ε_t . Bollerslev (1986) introduced an alternative model, the GARCH (p,q) model. In his model, h_t was articulated as follow,

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \theta_1 h_{t-1} + \dots + \theta_p h_{t-p}, \quad (5)$$

where $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \theta_j < 1$.

The ARCH-M model was introduced by Engle et al. (1987). While many theories in finance engage an explicit trade off among the risk and the expected return, this model is preferably suited to handling such questions in a time series background where the conditional variance may be time-varying. The model is expressed as follows,

$$y_t = \beta y_{t-1} + \gamma h_t + \varepsilon_t \quad (6)$$

People found the applications of the GARCH model limited because the conditional variance is simply linked to past conditional variances. Then the EGARCH model was modified by Nelson (1991). In this class of ARCH models, the volatility depended not only on the degree of the past surprises in returns but also on their corresponding signs. EGARCH process is indicated as follows,

$$\ln(h_t) = \alpha_0 + \sum_{j=1}^p \theta_j \ln(h_{t-j}) + \sum_{i=1}^q \alpha_i g(v_{t-i}), \quad (7)$$

$$g(v_t) = \varphi_t v_t + \left| \frac{\varepsilon_t}{\sqrt{h_t}} \right| - E \left| \frac{\varepsilon_t}{\sqrt{h_t}} \right|. \quad (8)$$

To sum up, we can find that the ARCH family models focus on structure the different models to dispose the variance ε_t in order to acquire more information from the time series data.

2.2. The Artificial neural networks model (ANNs)

In this section, we briefly describe the functional from ANNs⁵, which are considered best for forecasting stock market time series. ANNs, as computational intelligence systems, are nonlinear methods, taking an input vector X and produces output vector Y . The relationship between X and Y is determined by the network architecture. They do not require specific assumptions about the underlying model and has been used widely for approximation functions and forecasting (Hornick et al., 1989; White, 1989). For this reason, ANNs are called universal approximates that can approximate a large class of functions with a high degree of accuracy (Chen et al., 2003; Zhang & Qi, 2005).

Newly, ANNs have been successfully applied to predict important financial markets, like for example indexes, S&P500 and others (Chen, 1994; Huarng & Yu, 2006; Huang et al., 2007; Yu & Huarng, 2008) and others applications (Jain & Kumar, 2007; Majhi et al., 2009). The greatest advantage of an ANNs over other modeling techniques is their capability to model complex non-linear relationships in the observations with a high degree of accuracy (Chen et al., 2003; Wong et al., 2000). Among types of neural networks that are endowed to this property, they are the multilayer perception (MLP), radial

⁵ The first artificial neuron was proposed in 1943 by the neurophysiologist Warren McCulloch and in the 1990s the use of neural networks for identifying or predicting time series became widespread.

basis function (RBF) networks and recurrent networks, which are widely used in time-series forecasting, because they are capable of resolving a wide of problems (Sarle, 2002). But the multi-layer feed forward model is the most popular neural networks used in forecasting problems.

A multilayer perceptron neural network model, developing by the back-propagation algorithm, is grouped in layers and a set of layers forms a network. There are three kinds of layers: inputs, representing the external data into the artificial neural networks, output, which depends on the information out and hidden layers that, appears between the input and output. For input layers, it receives the inputs and directs them to the hidden layers. For time series forecasting and modeling, one hidden layer feed forward network is the most widely used (Zhang & Qi, 2005). The MLP neural network is illustrated in Fig. 1. This figure exhibits an example of a simple feed-forward network for time series analysis with one hidden layer. The input layer has five nodes, and the hidden layer has three. The input nodes are connected forward to each and every node in the hidden layer, and these hidden nodes are connected to the single node in the output layer. We call the network a 5-3-1 feed-forward network.

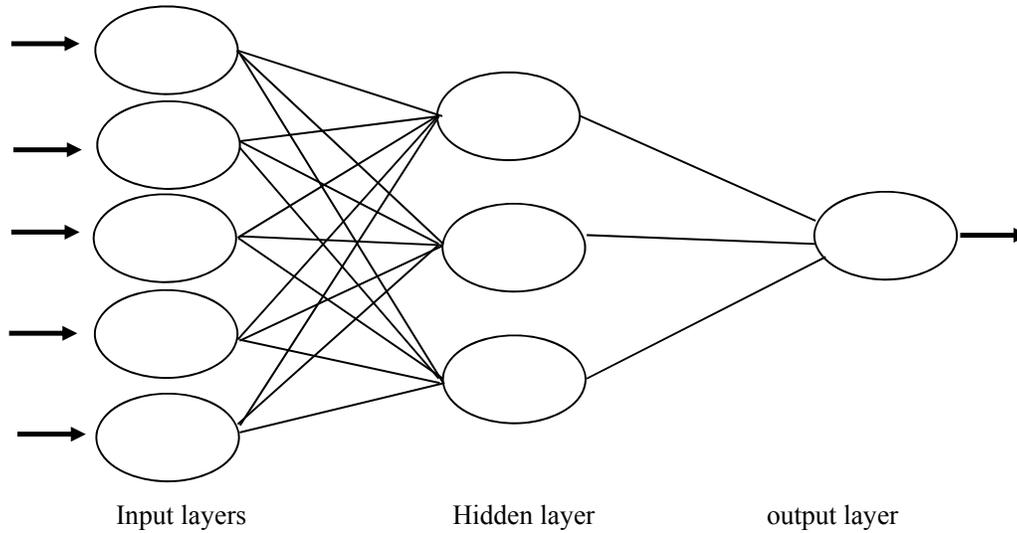


Fig. 1. An example of a Multilayer Perceptron (MLP) with one hidden layer for time series forecasting. The relationship between the output y_t and the inputs $(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ has in general the following mathematical representation:

$$y_t = w_0 + \sum_{j=1}^q w_j g \left(w_{0j} + \sum_{i=1}^p w_{ij} y_{t-i} \right) + \varepsilon_t, \quad (9)$$

where w_j ($j=1, 2, \dots, q$) are model parameters often called connection weights; w_{ij} ($i=1, 2, \dots, p; j=1, 2, \dots, q$) and p is the number of input nodes; q is the number of hidden nodes. ε_t is a residual i.i.d. $g(\cdot)$ function determines the connections between nodes of the hidden layer, and is used as the hidden unit activation function to enhance the nonlinearity of the model. The most commonly used are the logistic and the hyperbolic tangent functions:

$$g(x) = \text{Sig}(x) = \frac{1}{1 + \exp(-x)}, \quad (10)$$

$$g(x) = \text{Tanh}(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}, \quad (11)$$

In fact, the ANN model of Eq. (9), performs a nonlinear functional mapping from the past observations to the future value y_t , i.e.

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \theta) + \varepsilon_t, \quad (12)$$

where, $f(\cdot)$ function is determined by the network studies and connection weights. θ is a vector of all parameters. Getting up and development with Neural networks involves several steps: (Gradijevic & Yong, 2000).

1. Specifying the structure of the neural network or selection of the input and output variables, number of hidden layers, and number of hidden neurons, training algorithm, activation function and initial weights.
2. Division of the data set into three sets, network training, validation and testing sets. Sixty percent of the data set using to training, thirty percent was used for validation and the rest to test.
3. Estimation of the predict output. The input values of the last 10 percent of the observation were then used by the trained to generate output, i.e. stock market returns forecasting in this study.
4. Evaluation of the forecast performance of the ANN, and comparison of it to other models using RMSE or others measures.
5. Steps 1 to 4 were repeated until the error goal was not reached.

Over the last decade, neural networks represent a recent approach to time series forecasting that are a newly proliferating technique in desktop quantitative analysis. They have been found to be a viable contender to various traditional time series models (Gradojevic & Yang, 2000; Chen et al., 2005; Jain & Kumar, 2007).

2.3. Forecast Evaluation criterions

Three criterions will be used to evaluate the model forecasting aptitude. The first is the root mean squared error (RMSE). The formula for RMSE is:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}, \quad (13)$$

where \hat{y}_t is the predicted value, y_t is the actual value.

The second is the mean absolute percent error (MAPE). The average error in proportion terms for each point forecast, MAPE is given by:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \times 100 \right|. \quad (14)$$

The third is the Theil disparity Coefficient (Theil IC). Theil IC value is always among 0 and 1, and a smaller value designates the error among the predicted value and the actual value is smaller. Theil IC is given by:

$$Theil\ IC = \frac{\sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}}{\sqrt{\frac{\sum_{t=1}^n \hat{y}_t^2}{n}} + \sqrt{\frac{\sum_{t=1}^n y_t^2}{n}}} \quad (15)$$

3. Empirical analysis

The daily of S&P500 futures price data are obtained from Reuters Information System. The data span the period from 8 July 1996 to 26 May 2006 (2488 observations). We take data from 8 July 1996 to 25 May 2005 as in-sample data sets (2236 observations), and we take the data from 27 May 2005 to 26 May 2006 as out-of-sample data sets (252 observations) which are used to evaluate the performance of the predictions.

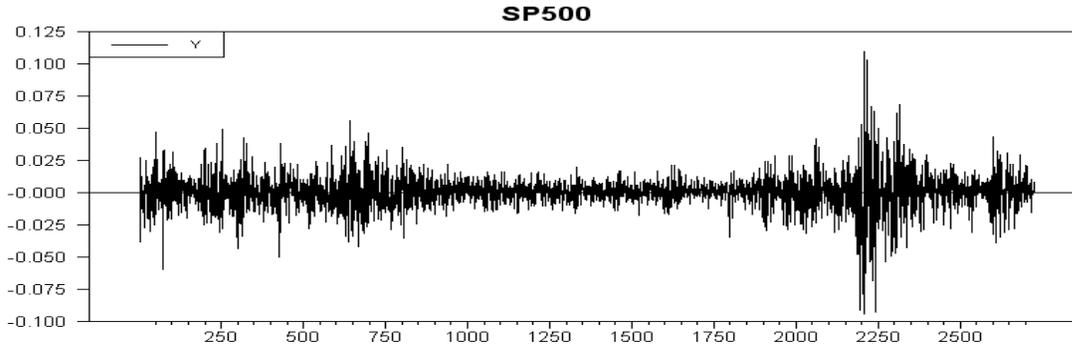


Fig. 2. Plots logarithm returns data series of S&P500 futures price

The main equation is regression equation and Akaike's Information Criterion (AIC) rule is used to determine the lag lengths. Finally, we select 3 lags after testing many times. The residual equation we selected are ARCH (1), GARCH (1,1), ARCH (1)-M, GARCH (1,1)-M, EGARCH(1,1) for Bollerslev's (1988) idea, in most applications $p = q = 1$ is found to suffice. On the other hand, the ANN model uses trial and error to determine the network architecture of (12) by minimizing the forecasting error. First, we process the forecast using the ARCH family models. Table 1 gives the forecasting result. According to the forecast evaluating criteria, we can find that the ARCH(1)-M performs best.

Table 1

Forecast Result of ARCH Family Models

	RMSE	Rank	MAPE	Rank	Theil IC	Rank
ARIMA	7.214	3	1.342	2	0.0072	1
ARCH(1)	7.211	1	1.348	6	0.0074	3
GARCH(1,1)	7.213	2	1.346	5	0.0073	2
EGARCH(1,1)	7.216	5	1.340	1	0.0075	4
ARCH(1)-M	7.215	4	1.343	3	0.0077	5
GARCH(1,1)-M	7.217	6	1.345	4	0.0079	6
ANN()	6.38	4	1.26	4	0.0003	5

Therefore, we form the hybrid model using the best ARCH(1)-M model combined with the ANN model. Table 2 gives the forecasting results. Simultaneously, we also process the forecast using the ARIMA-ANN hybrid model. According to the forecast evaluation criteria, we can find that the ARCH (1)-M-ANN model performs better than the ARIMA-ANN model, and also performs the best in all the given models.

Table 2

Forecast Result of Hybrid Models

Hybrid Models	RMSE	Rank	MAPE	Rank	Theil IC	Rank
ARCH(1)-M-ANN	7.193	1	1.323	1	0.0064	1
ARIMA-ANN	7.191	2	1.328	2	0.0065	2

According to the NMSE, MAPE, Thiel IC evaluation criterion, we can discover that the ARCH(1)-M-ANN hybrid model is greater to ARIMA, ARCH(1), GARCH(1,1), EGARCH(1,1), ARCH(1)-M, GARCH(1,1)-M and ARIMA-ANN models for S&P500 futures price forecasting.

4. Conclusions

Commodity index futures represent a useful investment vehicle for speculators and hedgers in stock markets. In this study, we have suggest a hybrid model combining ARCH-M model and ANN model to predict S&P500 futures price. In terms of the empirical results, we find that the model combining ARCH-M and ANN models performs the best on the selected criteria.

In this paper, the auto-regressive integrated moving average models have been applied to propose a new hybrid method for improving the performance of the artificial neural networks to time series forecasting. In our proposed model, based on the Box-Jenkins methodology in linear modeling, a time series is considered as nonlinear function of several past observations and random errors.

Therefore, in the first stage, an auto-regressive integrated moving average model is used in order to generate the necessary data, and then a neural network is used to determine a model in order to capture the underlying data generating process and predict the future, using preprocessed data. Empirical results with three well-known real data sets indicate that the proposed model can be an effective way in order to yield more accurate model than traditional artificial neural networks. Thus, it can be used as an appropriate alternative for artificial neural networks, especially when higher forecasting accuracy is needed.

References

- Ahmadi, H. (1990). Testability of the arbitrage pricing theory by neural networks. *Proceedings of the International Conference on Neural Networks*, 385- 393.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis*, 9(2), 121-131.
- Chen, C. H. (1994). Neural networks for financial market prediction. *Proceedings of the IEEE International Conference on Neural Networks*, 2, 1199-1202.
- Chen, A., Leung, M. T., & Hazem, D. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers and Operations Research*, 30, 901-923.
- Choi, H., Lee, E., & Lee, C. (2006). Neural Network Deinterlacing Using Multiple Fields. *Lecture Notes in Control and Information Sciences*, 345, 970-975.
- Choi, H., Lee, C., & Rhee, M. W. (1995). Trading S&P 500 stock index futures using a neural network. *The Third Annual International Conference on Artificial Intelligence Applications on Wall Street*, 63-72.
- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50, 987-1008.
- Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica: Journal of the Econometric Society*, 391-407.
- Gradojevic, N., & Yang, J. (2000). The Application of Artificial Neural Networks to Exchange Rate Forecasting: The Role of Market Microstructure variables, *Working Paper n° 2000-23, Bank of Canada*.
- Hamid, S. A., & Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index futures prices. *Journal of Business Research*, 57(10), 1116-1125.
- Hornick, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal

- Approximators. *Neural Networks*, 2, 359-366.
- Huang, W., Lai, K. K., Nakamori, Y., Wang, S. Y., & Yu, L. (2007). Neural networks in finance and economics forecasting. *International Journal of Information Technology and Decision Making*, 6, 113-140.
- Huang, K., & Yu, T. H. K. (2006). The application of neural networks to forecast fuzzy time series. *Physica A: Statistical Mechanics and its Applications*, 363(2), 481-491.
- Kamijo, K., & Tanigawa, T. (1990). Stock price pattern recognition: A recurrent neural network approach. *Proceedings of the International Joint Conference on Neural Networks*, 215-221.
- Kim, N. H., Yoo, S. K., & Lee, K. S. (2003). Polygon reduction of 3D objects using Stokes' theorem. *Computer methods and programs in biomedicine*, 71(3), 203-210.
- Kunhuang, H., & Yu, T.H.K. (2006). The application of neural networks to forecast fuzzy time series. *Physical A: Statistical Mechanics and Its Applications*, 363(2), 481-491.
- Majhi, R., Panda, G., Sahoo, G., Dash, P. K., & Das, D. P. (2007, September). Stock market prediction of S&P 500 and DJIA using bacterial foraging optimization technique. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on* (pp. 2569-2575). IEEE.
- Miao, K., Chen, F., & Zhao, Z. G. (2007). Stock Price Forecast Based on Bacterial Colony RBF Neural Network [J]. *Journal of Qingdao University (Natural Science Edition)*, 20 (2), 50-54 (in Chinese).
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 59(2), 347-370.
- O'Connor, N., & Maddem, M. G. (2006). A neural network approach to predicting stock exchange movements using external factors: Applications and innovations in intelligent network to investment analysis. *Financial Analysts Journal*, 78-80.
- Tan, T. Z., Quek, C., & Ng, G. S. (2005). discovery using complementary learning fuzzy neural learning memory structures. *Neural Networks*, 18(5-6), 818-825.
- Trippi, R. R., & DeSieno, D. (1992). Trading equity index futures with a neural network. *The Journal of Portfolio Management*, 19(1), 27-33.
- White, H. (1989). Some asymptotic results for learning in single hidden-layer feedforward network models. *Journal of the American Statistical Association*, 84(408), 1003-1013.
- Yu, T. H. K., & Huang, K. H. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications*, 34, 2945-2952.
- Zhang, G. P., & Qi, G. M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, 160, 501-514.
- Zhang, D., Patuwo, E., & Hu, Y.M. (1998). Forecasting with artificial neural networks the state of the art. *International journal of forecasting*, 14(1), 35-62.
- Zhu, X., Wang, H., Xu, L., & Li, H. (2007). Predicting stock index increments by neural networks: The role of trading volume under different horizons. *Journal of Expert Systems with Applications*, 34(4), 3043-3054.