ISSN 2349-7807

International Journal of Recent Research in Commerce Economics and Management (IJRRCEM)
Vol. 2, Issue 2, pp: (134-140), Month: April 2015 - June 2015, Available at: www.paperpublications.org

A Combined Model between Artificial Neural Networks and ARIMA Models

Alshaimaa Ibrahim Elwasify

Assistant Lecturer at Department of Applied Statistics, Faculty of Commerce - Damietta University

Abstract: **The main objective of this study is to reach an appropriate model to predict the stock market index EGX 30. The study examined the application of the following models to predict EGX30:**

- **Artificial neural networks.**
- **ARIMA models.**
- **Combination between neural networks and time series analysis using observations. previous residuals and estimated values of ARIMA model.**

The study showed that the most appropriate model to predict the index of stock market EGX30 is the model of combination between neural networks and ARIMA models, where it gives more accurate results rather than ARIMA and ANN each separately, that is because the combination between these models combining the flexibility of the time series and the power of artificial neural networks, where one of these models compensates the shortage of the other model, further more artificial neural networks give the best prediction rather than ARIMA models in accordance with the standards of prediction accuracy MAPE and MSRE.

Keywords: **ARIMA, AR, MA, Box – Jenkins, ANN, ARIMA – ANN, combination between ARIMA and ANN.**

1. INTRODUCTION

Prediction of the future is one of the most fundamental issues used in many areas, time series is the most common statistical methods used to predict, which are used widely in many statistical and economic applications, where it is to predict the behavior of the dependent variable based on its behavior in the past.

On the other hand, there is a modern way more accurate and effective in forecasting, it can use logic in their operations rather than a fixed relationship between symbols and reactions, it is the artificial neural networks (ANN).

ANN are an appropriate way to represent the relationships between variables that differ from traditional methods, it is a system consists of a set of simple elements associated with each other to run the data dynamically to response to external input, so artificial neural networks is processing information system which simulates biological neural networks. Researchers have found that treatment methods in artificial neural networks are approaching treatment in biological networks and thus have been used in many fields, including the field of medicine, control, and other business, also ANN can treat many types of data like Linear, Non- Linear, Incomplete and Noisy data.

With the continuing need for more accurate predictions for various phenomena, the researchers are trying to combine two models instead of a single model for predicting models, where one of them offset the shortage of other. Combination between neural networks and time series models will give the most accurate results than if you use all of them alone, so the search is based on the combination of ANN and time series analysis, this will combine the flexibility of the time series with the power of ANN.

There are many ways to combine the two models, in this research we use the observation, previous residuals and estimated values of the ARIMA model to build the combination model. Therefore there will be a number of different forecasting models, that are ANN, ARIMA, and the combination model by the observation, previous residuals and estimated values of the ARIMA, the problem is to choose the appropriate model that can be used to predict EGX 30 index.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) are a family of statistical learning algorithms inspired by biological neural networks and are used to estimate or approximate functions that can depend on a large number of inputs. ANN are generally presented as systems of interconnected "neurons" which can compute values from inputs, and are capable of machine learning as well as pattern recognition thanks to their adaptive nature.

The data is processed by neural networks as follows:

1- Processed the information in simple processing elements called neurons.

2- Signals pass between neurons cross-linking lines.

3- Each link is attached by a certain weight, which is multiplied by incoming signals to neurons.

4- Each neuron applied an activation function is on the total input (the sum of the weighted input signals) to determine the output signal resulting from it.

$$netk = \sum_{i=1}^{k} W_{ki} X_i$$

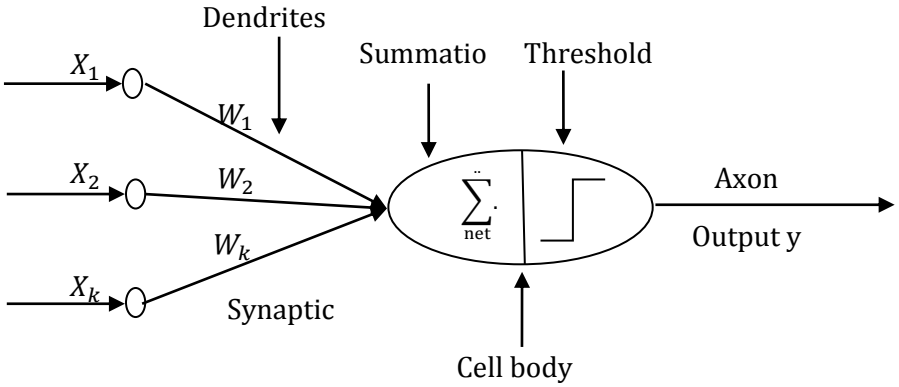

Fig.1.Simple neuron

3. ARIMA MODELS

In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series. They are applied in some cases where data show evidence of non-stationary, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationary.

The model is generally referred to as an ARIMA(p,d,q) model where parameters p, d, and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. ARIMA models form an important part of the Box-Jenkins approach to time-series modeling.

ISSN 2349-7807

International Journal of Recent Research in Commerce Economics and Management (IJRRCEM)
Vol. 2, Issue 2, pp: (134-140), Month: April 2015 - June 2015, Available at: www.paperpublications.org

3.1 general formula of ARIMA(p,d,q)

$$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad \text{Equation 1}$$

Where

$$w_t = \Delta^d y_t$$

Model's formula can be shortening using the backward shift operator (B) as follows:

$$\phi(B)\Delta^d y_t = \delta + \theta(B)\varepsilon_t \quad \text{Equation 2}$$

Where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad \text{Equation 3}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad \text{Equation 4}$$

3.2 Steps for building an ARIMA model

1- Model Identification:

The choice of a particular mathematical model based on some statistical measures that characterize the model for another, and the experience derived from studies and research.

2- Model Estimation:

After the nomination of one or more appropriate models to describe the viewing time series of data, we estimate parameters of the model using one of estimation methods.

3- Model Diagnostic:

It includes the residuals analysis to see how close the calculated values from the nominated model with observations, and measuring of the validity of the assumptions of the model. In the case of passing the model for these tests, we adopt it as the final model, which is used to estimate future predictions, but in the case of non-passing, we return to the first step for the appointment of a new model.

4- Predicting:

The final model is used to generate future predictions and then calculate the prediction errors that occurred.

4. PROPOSED MODEL

Combination between ANN and ARIMA models by the observation, previous residuals and estimated values of the ARIMA (Bijari and Kashei 2010, pp.2664–2675).

Assuming that the time series contains two main parts, one linear and the other is non-linear, the model will be as following:

$$Y_t = f(L_t, N_t) \quad \text{Equation 5}$$

Y_t: Observation of time series at time t.

L_t: linear part of ARIMA model.

N_t: Non − linear part of ANN.

This model is built according to the following steps:

1- Building ARIMA model for time series data y_t starting from Model Identification, to predicting stage to estimate the linear part.

$$L_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j} + e_t = \hat{L}_t + e_t \quad \text{Equation 6}$$

ISSN 2349-7807

International Journal of Recent Research in Commerce Economics and Management (IJRRCEM)
Vol. 2, Issue 2, pp: (134-140), Month: April 2015 - June 2015, Available at: www.paperpublications.org

$\hat{L}_t$: Estimation values

e_t : residuals resulting from the linear part ARIMA.

2- Building a training ANN using the observation, residuals resulting from ARIMA – which have a non-linear relationships –, and estimated values of the ARIMA model, and therefore:

$$N_t^1 = f^1(e_{t-1}, e_{t-2}, \dots, e_{t-q}) \quad \text{Equation 7}$$

$$N_t^2 = f^2(y_{t-1}, y_{t-2}, \dots, y_{t-p}) \quad \text{Equation 8}$$

$$N_t = f(N_t^1, N_t^2) \quad \text{Equation 9}$$

f^1, f^2: Non linear Function by ANN

3- So the combination model will be as follows:

$$y_t = f(N_t^1, \hat{L}_t, N_t^2)$$

$$= f(e_{t-1} \dots, e_{t-q}, \hat{L}_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$$

f: Non linear Function by ANN

The following figure illustrates that.

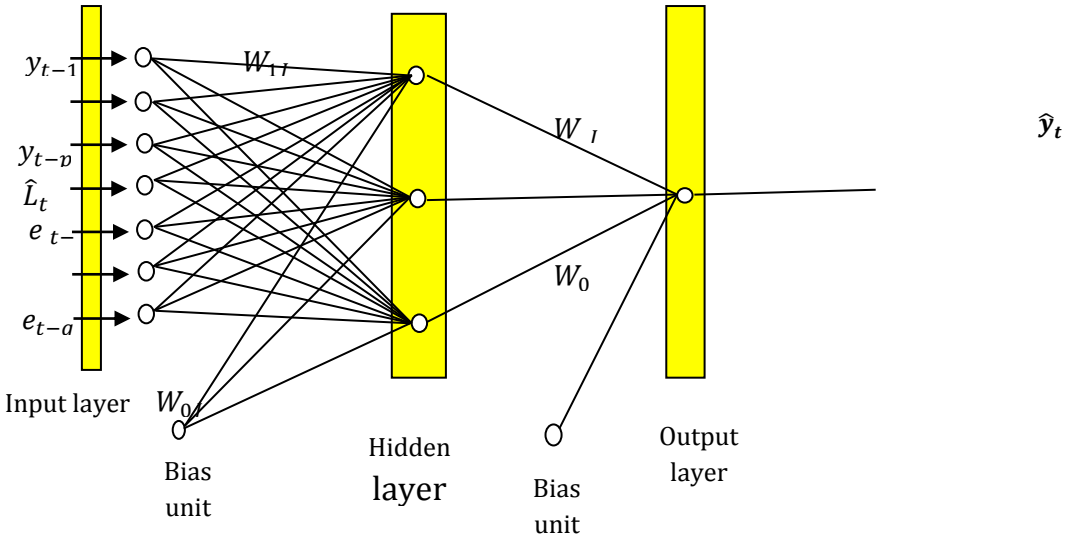

Fig.2. A neural network shows the predicted values of the combination model using observation, previous residuals and estimated values of ARIMA

5. THE APPLIED STUDY

Research focuses on predicting the value of the stock market EGX30 index. The period of The period of research was limited from October 2009 until the end of October 2010.

The appropriate ANN is MLP neural networks, it consists of three layers, an input layer is represented in the values of EGX30 y_{t-1}, a hidden layer is composed of 6 processing element, and finally the output layer is the current index values y_t, so the model is ANN(1:6:1).

The best ARIMA model after residuals analysis is ARIMA (0,2,2), and can be used in the forecasting process. The model form as follows:

$$y_t = 2y_{t-1} - y_{t-2} + e_t - .8017e_{t-1} - .1769e_{t-2}$$

After set an appropriate ARIMA model and . A neural network (MLP) was trained by observation, previous residuals and estimated values of ARIMA (0,2,2). Thus, the neural network used consists of three layers. The inputs layer ($e_{t-1}, e_{t-2}, \hat{L}_t, y_{t-1}, y_{t-2}$), a hidden layer that is composed of 6 processing elements, and finally the output layer ($\hat{y}$). The models are ARIMA- MLP (0,2,2)(5:6:1).

The observed and Predicted values of EGX30 are in the following table

TABLE 1. Predicted values of three models for EGX30

Date	ANN (1:6:1)		ARIMA(0,2,2)		ARIMA-MLP(0,2,2)(5:6:1)	
	Observed values	Predicted values	Observed values	Predicted values	Observed values	Predicted values
1-11-2010	6612.27	6626.86	6612.27	6665.77	6612.27	6625.4
2-11-2010	6654.76	6655.63	6654.76	6667.12	6654.76	6669.85
3-11-2010	6713.99	6660.31	6713.99	6668.47	6713.99	6699
4-11-2010	6764.61	6655.91	6764.61	6669.82	6764.61	6670.18
7-11-2010	6829.34	6632.11	6829.34	6673.87	6829.34	6726
8-11-2010	6837.19	6701.92	6837.19	6675.22	6837.19	6691.25

From TABLE 1. we note that:

- predicted values of ANN is close to actual values, especially in the first three days, but in the last three days the predictions is not good, where we note a wide difference between the actual values and the predicted values, but in general the differences is not significant, where the calculated χ^2 equal to (10.84), at 5 degrees of freedom and tabular χ^2 equal to (11.07).

- predicted values of ARIMA is close to actual, especially in the first four days, but in the last two days the predictions is not good, where we note a wide difference between the actual values and the predicted values, but in general the differences is not significant, where the calculated χ^2 equal (10.84), at 5 degrees of freedom and tabular χ^2 equal (11.07).

- predicted values of ARIMA $-$ MLP$(0,2,2)(5:6:1)$ is very close to actual values, and the differences is not significant, where the calculated χ^2 equal (6.21), at 5 degrees of freedom and tabular χ^2 equal (11.07).

6. COMPARISONS

The following table shows the estimation of differentiation criteria between the estimated models, the mean absolute relative prediction error (MAPE), and mean squared relative prediction error (MSRE).

TABLE 2. Estimation of differentiation criteria between the estimated models

MODEL	MAPE	MSRE
ARIMA(0,2,2)	.0129	.000230
ANN(1,6,1)	.0125	.000229
ARIMA-ANN(0,2,2)(5:6:1)	.0086	.000121

From TABLE 2., we note that:

1- The combination model gave more accurate predictions than using each model separately.

2- The best model for EGX 30 index is ARIMA $-$ MLP$(0,2,2)(5:6:1)$, where it has a lower value of MAPE and MSRE.

7. CONCLUSIONS

Accurate predictions in the future of a phenomena is the basis of economic policy success, in an attempt to estimate the model to predict the index stock EGX30, we used 270 daily observations in a form of a time series beginning from October 1, 2009 until October 31 2010 in the estimation stage and using 6 next daily observations from these data to test the model.

The following are the most important results that have been reached in this study:

Box and Jenkins method is considered the most flexible method in the construction of time-series models but resarcher need skill and experience in the first stage.

Neural networks is a modern Technique that have been applied in studies and we can use it to predict future, even though they do not need to hypotheses as in statistical methods, but it is more accurate.

Combined model between ARIMA and ANN by observation, previous residuals and estimated values of ARIMA (0,2,2) have shown better performance when compared with ARIMA and ANN each separately. Thus, the best model for EGX30 index is ARIMA-ANN(0,2,2)(5:6:1).

REFERENCES

[1] Ajoy, K. (2005)."Computational Intelligence in Time Series Forecasting Theory and Engineering Applications". Springer.

[2] Aslanargun, A.,Mammadagha,M., Yazici, B. and Yolacan, S. (2007) "Comparison of ARIMA, Neural Networks And Hybrid Models in Time Series:tourist arrival forecasting" Journal of Statistical Computation and Simulation, Volume 77, No.1, (29-53).

[3] Chattopadhyay, G. and Chattopadhyay, S (2010)."univariate modeling of summer – monsoon rainfall time series: Comparison between ARIMA and ARNN". C. R. Geoscience.

[4] Edward, R. (2004). "An Introduction to Neural Networks". Visual Numerics.

[5] El – Ramly, M. (1996). " Comprehensive Study On Neural Network Applications In Mathematical Programming" Ph.D, ISSR - Cairo University.

[6] Faruk, D. O. (2009). "A hybrid neural network and ARIMA model for water quality time series prediction". Engineering Application of Artificial Intelligence.

[7] Fausett, L, (1994). "Fundamentals of Neural Networks: Architectures, Algorithms, and Applications". New York: Prentice-Hall, Inc.

[8] Hung, N., Babel, M., Tripathi,N., and Weesakul, S. (2009). "An artificial neural network model for rain fall forecasting in Bangkok, Thailand". Copernicus Publications on behalf of the European Geosciences Union. (1413-1425).

[9] Khalifa, H. A. (2005). "The Effect of Using Artificial Neural Network (ANN) On E-business Applications" .MSC. Sadat Akademy.

[10] bijari, M. and Khashei, M. (2010). "An artificial nural network (p,d,q) model for time series forecasting". Epert Systems with Applications, (479-489).

[11] bijari, M. and Khashei, M. (2010). "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting". The Impact of Soft Computing for the Progress of Artificial Intelligence. Volume 11, Issue 2, (2664–2675).

[12] Arabatzis, G., Koutroumanidis, T., and Loannou. K. and (2009). "predicting fuelwood prices in Greece with the use of ARIMA models, Artificial neural networks and a hybrid ARIMA- ANN model". Energy Policy (3627-3634).

[13] Lawrence, R. (1997). "Using Neural Network to Forecast Stock Market Price". University of Manitoba.

[14] Hyndman, R.J., Makridakis,S., and Wheelwright, S.C. (1998). "Forecasting: Methods And Applications". 3rd ed. John Wiley and Sons.

[15] Cheng Yu, H., Hsiung Tzeng, and Mei Tseng, F., G. (2002). "Combining neural network model with seasonal time series ARIMA model". Technological Forecasting and Social Change. Volume 69, Issue 1, (71–87).

[16] Nochai, R. and Nochai.T. (2006). "ARIMA Model for Forecasting Oil Palm Price" University Sains Malaysia, Penang.

[17] Pankratz, A. (1983). "Forecasting with Univariate Box-Jenkins Models: Concepts and Cases". John Wiley and Sons.

[18] Pindyck, R. S. and Rubinfeld, D.L. (1991). "Econometric Models And Economic Forecasts". McGraw-Hill, Inc.

[19] Rutka, G. (2008). "Network Traffic Prediction using ARIMA and Neural Network Model". ISSN 1392 – 1215, No.4(84) Electronics and Electrical Engineering.

[20] Saeed, N. (2000)."Forecasting of Wheat Production in Pakistan using ARIMA Models". International Journal of Agriculture &biology 1560-8530/2000/02-4-352-353.

[21] Shankar,T. N .(2008). "Neural Networks". UNIVERSITY SCIENCE PRESS.

[22] Gupta, N. D. and Smith, A. K. (2000). "Neural Networks in Business Techniques and Applications". Idea Group, Information Science Publishing.

[23] Parray,O., Diaz,A., and Suarez, c. (2009). "An ARIMA model for forecasting WI-Fi data network traffic values" AGOSTO DE, (65-69).

[24] Wallace, M. P. (2008)." Neural Networks and Their Application To Finance". Business Intelligence Journal.

[25] Zhang, G.P. (2003). "Time Series Forecasting Using A Hybrid ARIMA and Neural Network Model". Neurocomputing, Volume 50, (159 – 175).

[26] Zhang, G.P.(2004). "Neural Networks in Business forecasting''. Idea Group Publishing.