Prediction of Bursa Malaysia Stock Index using Autoregressive Integrated Moving Average and Artificial Neural Network

Angie Tan Sze Yiing
Faculty of Management, Multimedia University, Cyberjaya, Selangor, Malaysia

Chan Kok Thim*
Faculty of Management, Multimedia University, Cyberjaya, Selangor, Malaysia
ktchan@mmu.edu.my (*Corresponding author)

Abstract

The FTSE Bursa Malaysia Kuala Lumpur Composite Index (KLCI) comprises the 30 largest listed companies is often used as the benchmark for the overall market performance of Malaysian stocks. Despite claims that the stock market is in fact an efficient market that follows a random walk process and therefore, is not predictable, there are evidences from previous empirical studies that argued that the stock market can be predictable. This study basically examines the ability of time series analysis and artificial intelligence system to predict the movement of stock prices. We use daily historical data from 3 January 2012 to 31 March 2014 as the base, whereas the daily forecasts will be generated for the period starting from 1 April 2014 to 31 March 2015 using Eviews 7 and MATLAB 8.5 with Neural Network Toolbox version 8.3. The predictability of the stock market is conducted using two different approaches of stochastic time series analysis and artificial intelligence system, namely the Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) methods, respectively. The findings revealed that the ANN (5-6-1) model outperformed the ARIMA (1,1,0) model with the lowest error recorded. Thus, it is concluded that the ANN model is indeed a more superior forecasting model as compared to the ARIMA model in forecasting the stock price index. Investors and financial analysts would be able to adopt the best technique to forecast accurately and gain insights to the Malaysian stock market. The knowledge and skill gained in performing forecasts using the right technique would definitely give the users an added advantage, as they would be able to get more accurate forecasts, thus, making better decisions in terms of risk and investment management.

Key Words: Forecast, ANN, ARIMA, Modeling, Stock, KLCI.

1. Introduction

Trading in the stock market indices has become remarkably popular in the major financial markets around the world (Patel and Marwala, 2006). Lawrence (1997) states that the basic motivation in predicting stock market prices is undoubtedly, financial gain. Hence, investors have been trying to conquer multiple approaches to forecast the stock market prices and its future trends so as to obtain profit from the stock market (Lee, 2010). Still, due to the noisy, non-stationary and dynamic nature of stock prices, it is undeniably challenging to forecast stock market prices accurately (Wang, Wang, Zhang and Guo, 2011).

The predictability of the stock market has been a long debate among the market participants since the earlier years (Lee, 2010). A group of researchers emerged and claimed that the stock market is an efficient market and that it follows a random walk process, thereby indicating that the stock market movement is unpredictable. This group that favors the random walk process includes Malkiel (1973), Fama (1965) and Kendall and Hill (1953). However, there is another group of researchers that opposes to the random walk hypothesis thereby arguing that the stock market is indeed predictable (Moreno and Olmeda, 2007), as can be seen in studies conducted by Keim and Stambaugh (1986), Pesaran and Timmerman (1995) and Lewellen (2004).

Many studies in forecasting stock prices have been conducted over the years (Wang, Wang, Zhang and Guo, 2012). Among the most prominent forecasting techniques are the statistical time series models such as autoregressive integrated moving average (ARIMA), and artificial intelligence (AI) models including the artificial neural networks (ANN), fuzzy logic and expert systems.

McMillan (2001) has pointed out that the forecast of stock market returns is deduced to be absolutely possible by using a range of financial and macroeconomic variables. This statement is supported by several empirical evidences in the studies done by previous researchers such as Keim and Stambaugh (1986), Fama and French (1988), Pesaran and Timmerman (1995), Kim, Shamsuddin and Lim (2011) and Lewellen (2004). In fact, there are essentially two important hypotheses that are closely related to the possibility of forecasting, namely the random walk hypothesis and the efficient market hypothesis (EMH).

Introduced by Malkiel (1973), the random walk theory states that the future values or directions of the stock prices cannot be predicted based on the past behavior because the changes in stock prices are independent of each other. On the other hand, the EMH states that an efficient market incorporates all freely available information and stock prices adjust to the new information immediately (Fama, 1970). This instantaneous adjustment of stock prices to new information is usually unpredictable, thereby leading the change in the stock prices to be absolutely random (Malkiel, 2003). Hence, it is concluded that the random walk theory is certainly related to the EMH. Fama (1970) practically classified the market efficiency into three types, namely the weak form EMH, semi-strong form EMH and strong form EMH, based on the different sets of information that includes historical prices, public information and private information respectively. It is concluded in the studies of Nassir, Ariff and Mohamad (1993) and Lim, Liew and Wong (2004) that Bursa Malaysia is essentially a stock market with weak-form efficiency.

The ARIMA model, or better known as Box-Jenkins methodology, is a popular approach that is widely used in analysis and forecasting, especially in predicting time series models, as it is deemed to be the most efficient technique for forecasting in the area of social science (Adebiyi, Adewumi and Ayo, 2014). Mathew, Sola, Oladiran and Amos (2013) used ARIMA model to predict the stock prices of Nokia and Zenith stock index and Nigerian Breweries Plc stock prices respectively. The ARIMA methodology has also been proved worthy in forecasting economic and social variables such as unemployment rate, interest rates, rate of failure and so on, as shown by Dobre and Alexandru (2008), Ho and Xie (1998) and Hassan (2014).

Similarly, the ANN model has been a popular forecasting technique that is frequently analyzed and applied in time series forecasting in the recent years (Zhang, 2003). The ANN approach has been proven to be capable of modeling complex non-linear problems in which no prior assumption of the relationship is needed (BuHamra, Smaoui and Gabr, 2003). The main reason for its popularity is for the fact that ANN is able to learn patterns from data and deduce appropriate solutions for it (Adebiyi *et al.*, 2014). In essence, the stock market values are best modeled with the application of expert systems with ANN given that this approach does not have standard formulas and enables the changes of the market to be easily adapted (Guresen, Kayakutlu, and Daim, 2011). For instance, Kuan and Liu (1995), Jiang (2003), Devadoss and Ligori (2013), Wang *et al.* (2011) and Yildirim, Ozsahin and Akyuz (2011) demonstrated the application of ANN models in forecasting stock prices.

2. Results

2.1 ARIMA Model

Since the ADF unit root test gives a test statistic of -2.055 which is larger than the critical value at 5% level of significance at -2.865 and the correlogram plotted on the KLCI closing price, the time series data is proven to be non-stationary at level. Hence, it is differenced once at first level to obtain stationarity. The statistical results of different ARIMA parameters for KLCI that is obtained from the estimation output generated using ordinary least squares (OLS). Since ARIMA (1,1,0) has relatively the lowest SIC, AIC and standard error of regression at 7.1219, 7.1062 and 8.4348 respectively, relatively highest adjusted R^2 of 0.99%, a significant *p*-value of 0.0113 that is lower than 5% significance level and residual Q-stats *p*-value with lag 36 at 0.106 that is higher than 5% significance level indicating the residual is white noise, it is therefore the best model in relative to other models being tested. The performance measures of accuracy of ARIMA (1,1,0) is then evaluated based on MSE, RMSE, MAE and MAPE as depicted in Table 2.1.

Table 2.1: Performance measures of accuracy of all the selected models.

	ARIMA (1,1,0)
MSE	83.4905
RMSE	9.1373
MAE	6.6522
MAPE	0.3675

2.2 ANN Model

The data series of KLCI closing price are normalized between the range of -1 to 1 before being trained, validated and tested.

2.2.1 ANN Models Without Iterations

Table 2.2 below presents the MSE recorded for each of the neural network experiment that is done without iteration at 1,000, 2,000 and 3,000 epochs respectively. From the table, it is observed that the neural network 5-6-1 (5 input neurons, 6 hidden neurons and 1 output neuron) is selected as the predictive model with the lowest MSE at all epochs as compared to the other models.

2.2.2 ANN Models With Iterations

A series of neural network experiments is also tested with 10 iterations at 50, 100 and 200 epochs respectively. The idea behind this is to attempt to improve the performance of the model. The result of the test is shown in Table 2.2. The ANN model that gives the lowest MSE of all models after 10 iterations is the 5-4-1 (5 input neurons, 4 hidden neurons and 1 output neuron) neural network.

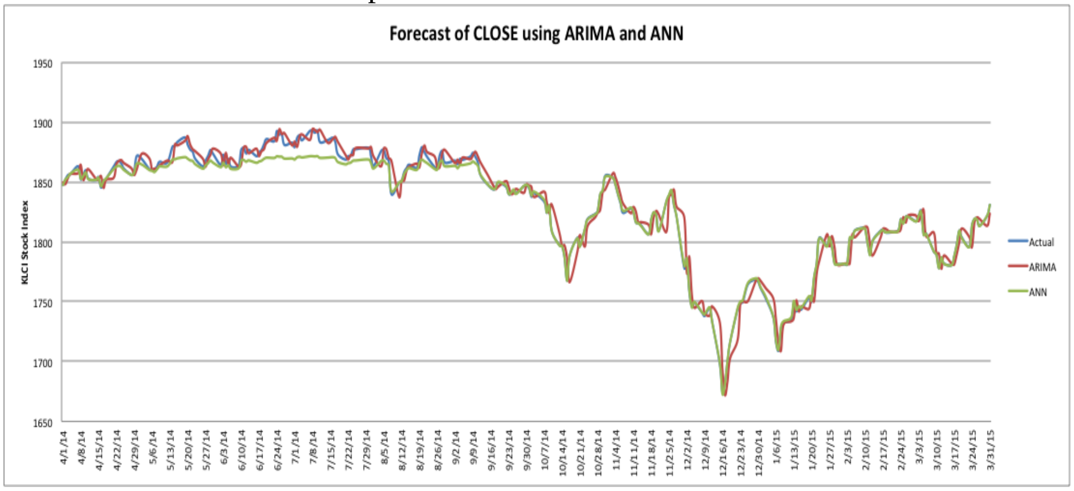
Table 2.2: Performance measures of accuracy of all the selected models.

	ANN (5-6-1)	ANN (5-4-1)
MSE	34.7891	37.1658
RMSE	5.8982	6.0964
MAE	3.2906	3.3457
MAPE	0.1759	0.1784

Since the ANN (5-6-1) model generally has a lower value of MSE, RMSE, MAE and MAPE as shown in Table 2.2, we can conclude that the ANN (5-6-1) model is a better predictive model that produces a more accurate daily forecast than the ANN (5-4-1) model.

From the Fig 2.1 below, it is observed that both the ARIMA and ANN models deviated from the actual KLCI close index with the ANN model only deviates in the beginning of the forecast period. However, towards the end of the forecast period, the ANN plot is hovering or directly above the actual KLCI close whereas there is a deviation of the ARIMA plot from the actual KLCI close index. Thus, the forecasts produced by the ANN (5-6-1) model are technically nearly the same values as the actual closing price index of the KLCI, whereas the results of the ARIMA (1,1,0) model is deviated from the actual values with an observable difference.

Figure 2.1: Graphical representation of the actual stock price index of KLCI against the forecasted stock index using ARIMA (1,1,0) and ANN (5-6-1) from 1 April 2014 to 31 March 2015.



3. Conclusion

This entire study focuses on the study of the predictability of two forecasting approaches namely the ARIMA and ANN model in predicting the daily closing price of the FTSE Bursa Malaysia KLCI over the period from 3 January 2012 to 31 March 2015. The outputs of these values provided an interesting account of the closeness between the actual and the forecasted values. Nonetheless, the accuracy of the forecast results can be further improved through the enhancement of the forecasting tools such as formation of a hybrid model combining either the ARIMA and ANN models, or even other approaches including expert systems or fuzzy logic technique. Overall, both the ARIMA and ANN approaches are indeed good predictive models and are capable of predicting the Malaysian stock market; especially in the short run, thus can be effectively engaged in risk and portfolio management.

References

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. *Journal of Applied Mathematics,* 1-7.

BuHamra, S., Smaoui, N., & Gabr, M. (2003). The Box-Jenkins Analysis and Neural Networks: Prediction and Time Series Modelling. *Applied Mathematical Modelling, 27,* 805-815.

Devadoss, A. V., & Ligori, T. A. (2013). Forecasting of Stock Prices Using Multi Layer Perceptron. *International Journal of Computing Algorithm*, *2*, 440-449.

Dobre, I., & Alexandru, A. A. (2008). Modelling Unemployment Rate Using Box-Jenkins Procedure. *Journal of Applied Quantitative Methods, 3* (2), 156-166.

Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, *51* (1), 75-80.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25* (2), 383-417.

Fama, E. F., & French, K. R. (1988). Permanent and Temporary Components of Stock Prices. *The Journal of Political Economy*, *96* (2), 246-273.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using Artificial Neural Network Models in

Stock Market Index Prediction. *Expert Systems with Applications, 38* (8), 10389-10397.

Hassan, J. (2014). ARIMA and Regression Models for Prediction of Daily and Monthly Clearness Index. *Renewable Energy*, *68*, 421-427.

Ho, S. L., & Xie, M. (1998). The Use of ARIMA Models for Reliability Forecasting and Analysis. *23rd International Conference on Computers and Industrial Engineering*, *35* (1), 213-216.

Isenah, G. M., & Olubusoye, O. E. (2014). Forecasting Nigerian Stock Market Returns Using ARIMA and Artificial Neural Network Models. *CBN Journal of Applied Statistics*, *5* (2), 25-48.

Jiang, H. (2003). Forecasting Stock Market with Neural Networks. *UMI Microform*.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting Returns in the Stock and Bond Markets. *Journal of Financial Economics*, *17*, 357-390.

Kendall, M. G., & Hill, A. B. (1953). The Analysis of Economic Time-Series-Part I: Prices. *Journal of the Royal Statistical Society. Series A (General)*, *116* (1), 11-34.

Kim, J. H., Shamsuddin, A., & Lim, K.-P. (2011). Stock Return Predictability and the Adaptive Markets Hypothesis: Evidence from Century-Long U.S. Data. *Journal of Empirical Finance*, *18* (5), 868-879.

Kuan, C.-M., & Liu, T. (1995). Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks. *Journal of Applied Econometrics*, *10* (4), 347-364.

Lawrence, R. (1997, December 12). Using Neural Networks to Forecast Stock Market Prices. *Department of Computer Science.*

Lee, A. (2010, June). A Stock Forecasting Framework for Building and Evaluating Stock Forecasting Models. *Thesis of Faculty of Graduate Studies and Research.*

Lewellen, J. (2004). Predicting Returns with Financial Ratios. *Journal of Financial Economics*, *74* (2), 209-235.

Lim, K.-P., Liew, V. K.-S., & Wong, H.-T. (2004). Weak-Form Efficient Market Hypothesis, Behavioral Finance and Episodic Transient Dependencies: The Case of the Kuala Lumpur Stock Exchange. *EconPapers*, 2-26.

Malkiel, B. G. (1973). *A Random Walk Down Wall Street.* New York, United States of America: W. W. Norton & Company, Inc.

Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives, 17* (1), 59-82.

Mathew, O., Sola, A. F., Oladiran, B. H., & Amos, A. A. (2013). Prediction of Stock Price Using Autoregressive Integrated Moving Average Filter (ARIMA (P,D,Q)). *Global Journal of Science Frontier Research Mathematics and Decision Sciences*, *13* (8), 79-88.

McMillan, D. G. (2001). Nonlinear Predictability of Stock Market Returns: Evidence from Nonparametric and Threshold Models. *International Review of Economics and Finance, 10* (4), 353-368.

Moreno, D., & Olmeda, I. (2007). Is the Predictability of Emerging and Developed Stock Markets Really Exploitable? *European Journal of Operational Research, 182* (1), 436-454.

Napitupulu, T. A., & Wijaya, Y. B. (2013). Prediction of Stock Price Using Artificial Neural Network: A Case of Indonesia. *Journal of Theoretical and Applied Information Technology*, *54* (1), 104-109.

Nassir, A. M., Ariff, M., & Mohamad, S. (1993). Weak-Form Efficiency of the Kuala Lumpur Stock Exchange: An Application of Unit Root Analysis. *Pertanika Journal of Social Sciences and Humanities*, *1* (1), 57-62.

Patel, P. B., & Marwala, T. (2006). Forecasting Closing Price Indices Using Neural Networks. *IEEE Conference on Systems, Man, and Cybernetics,* 2351-2356.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of Stock Returns: Robustness and Economic Significance. *The Journal of Finance*, *50* (4), 1201-1228.

Wang, J.-Z., Wang, J.-J., Zhang, Z.-G., & Guo, S.-P. (2011). Forecasting Stock Indices with Back Propagation Neural Network. *Expert Systems with Applications, 38* (11), 14346-14355.

Zhang, G. P. (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing, 50,* 159-175.

Yildirim, I., Ozsahin, S., & Akyuz, K. C. (2011). Prediction of the Financial Return of the Paper Sector with Artificial Neural Networks. *BioResources*, *6* (4), 4076-4091.