Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1
© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PREDICTION OF STOCK PRICE USING ARTIFICIAL NEURAL NETWORK: A CASE OF INDONESIA

[1]TOGAR ALAM NAPITUPULU, [2]YOHANES BUDIMAN WIJAYA

[1]Senior Lecturer, Grad. School, Bina Nusantara University, Jakarta, Indonesia

[2]Former Grad. Student, Bina Nusantara University, Jakarta, Indonesia

E-mail: [1]tnapitupulu@binus.edu , [2]yohanes.bw@gmail.com

ABSTRACT

Artificial Neural Network (ANN) has been widely used in many application because of its ability to solve non-parametric problems. ANN is also recognized as a good and widely used tool in forecasting stock prices. Traditionally stock forecasting in Indonesia usually used time series analysis. This paper compared the stock forecasting result of ANTM (PT Aneka Tambang) using ANN and that of Autoregressive Integrated Moving Average (ARIMA). The results of the study showed that forecasting using ANN method has smaller error than ARIMA method.
been developed for different hybrid power system configurations for tuning the proportional-integral controller for SVC. Transient responses of different autonomous configurations show that SVC controller with its gained tuned by the ANNs provide optimum system performance for a variety of loads.

K**eywords**: *Artificial Neural Network, ARIMA, time-series, stock forecasting*

1. INTRODUCTION

Stock is the most volatile investment with high risk, but with high return to investors if carefully managed in their portfolio. In managing stocks, information on their prices is of utmost importance. Investor has many ways to predict the price of a stock and among others is using time series analysis where one of the most frequently used techniques is Autoregressive Integrated Moving Average (ARIMA).

ARIMA model combines Autoregressive and Moving Average. It is a statistical forecasting method that has to follow some assumption on autocorrelation, homoscedasticity, multicolinearity, linearity, and assumption on normality of the error terms. In addition to these basic statistical assumptions, there is the assumption on the model specification, i.e., for forecasting using ARIMA, this would be the number of lags in the explanatory independent variables and in the moving average component of the model, and the seasonality in the model. Even though violations of this latter assumption could lead to problems in forecasting precision, in general however, this can be avoided by finding the best fit model through some statistical testing and iterations. One of the limitation of statistical forecasting is this model and assumption dependent that from time to time might be violated.

A relatively new approach to forecasting stock price or forecasting in general is the Artificial Neural Network (ANN) analysis. ANN, to a certain extent, mimics the way human brain works. In this analysis, the system learns from existing data without model specification and the need for the necessary assumption in the statistical forecasting. It is simply adjust itself to the data, learn from the data and hence it is a data based forecasting. In fact this data dependent model is both its strength and its weakness. It is the strength because ANN can mimic the data exactly the way the relationship of the variables exists in the data, that is, non-linear, represented by the ANN "machine". There is no need to concern about the number of variables included as long as the data are available, and hence there is no need to concern about degrees of freedom. It is the weakness because ANN requires a relatively large and complete data that represent the how the system works both now and in the future.

It is therefore interesting to compare forecasting power of both approaches. Working hypothesis would be that providing enough data, forecasting using ANN has smaller error compared with ARIMA.

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. THEORETICAL FRAMEWORK

A . STOCK

Capital Market is the market in which have long term financial instrument, such as bonds, equities, mutual funds and derivative instruments, are traded. Capital Market serves as an alternative for a company's capital resources and public investment. It also facilitates the infrastructures needed for the selling and buying process and other related activities.

Financial instruments traded in the Capital Market are long term securities (a period of more than 1 year). They consist of stocks, bonds, warrants, rights, mutual funds, and other derivative instruments (options, futures, etc.).

Capital Market plays an important role in the economy of a country because it serves two functions all at once. First, Capital Market serves as an alternative for a company's capital resources. The capital gained from the public offering can be used for the company's business development, expansion, and others. Second, Capital Market serves as an alternative for public investment. People could invest their money according to their preferred returns and risk characteristics of each instrument. One important factor that reflects this is and of which investors based their decision in managing their portfolio is the prices of the instruments.

In daily stock trading, stock price fluctuates. Stock price is formed by the demand and supply of the stock, while the supply and demand of a stock are influenced by many factors, among others are the company and industry's performance, the macro factors (interest rate, inflation, currency rate), the non-economical factors (social and political conditions).

B. ARIMA

Business forecasting with ARIMA model or also called Box-Jenskins method developed by George Box and Gwilym Jenkins (Box & Jenkins, 1976). ARIMA is a technique to find the most suitable pattern of a group of data basically by curve fitting (Sugiarto and Harijono, 2000). ARIMA utilize past and current data to make short-term forecasting. Examples are the use of ARIMA models in forecasting stock price based on the pattern of past stock price changes (Sugiarto and Harijono, 2000).

In general, this model is formulated as an ARIMA (p, d, q) where p is the degree of autoregressive (AR), d is the degree of difference (seasonality) and q is the degrees moving average (MA) in the model. AR model is a model that shows that the dependent variable is influenced by the same dependent variable in the earlier periods (lags), while the independent variable of MA model is the residual value (error) in prior periods. AR and MA models combined into ARIMA model.

ARIMA modeling consists of three basic steps, namely identification, assessment & testing phase, and the diagnostic phase. Identification stage is the search phase of the degrees p, d, q, with the help autocorrelation and partial autocorrelation. Assessment phase is the phase to estimate AR and MA parameters contained in the model. In general this estimation can use the estimation method that is the ordinary least square squares (OLS) method. Diagnostic stage is the stage to ascertain whether the model obtained fit the data, to avoid the possibility of a more appropriate ARIMA model to the data. This diagnostic stage is seen based on the value of the error or white noise. If the model is not suitable then the process is repeated from the beginning.

C. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is an information processing system that has characteristics similar to biological neural networks. According to Fauset (1994), a neuron in the neural network as analogous to biological neurons in which had three types of components that are part of the understanding of artificial neurons: dendrites, synapses and axons. A Dendrites usually get a signal from another neuron, signals in the form of an electrical impulse is transmitted via a synaptic gap with the help of chemical processes. These chemical processes modify the incoming signal, where the chemical process works closely with the neural network activation function. The second component, the synapses or cell body, is the sum of the incoming signals. Synapses is obtained from the activation process of existing dendrites via a known path, Axon. This path is symbolized by the neural network weights, where weight is what distinguishes the value of connections from each existing path.

According to Gonzalez (2000), ANN is a mathematical model that is structured like the way the human brain works to identify patterns in a number of variables. Analytically an ANN can be written in an equation such as in (1).

Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1
© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Y = h\left(\sum_{j=1}^{J} \alpha_J g\left(\sum_{i=1}^{I} \beta_{ij} X_i\right)\right). \quad (1)$$

Summations represent dendrites, while functions g and h represent axons and α and β represent synapses. Equation (1) then represent an ANN with I input neurons, one hidden layer with J neurons, and one output Y. This model or "machine" specified in (1) is uniquely determined by α and β. The objective then is to estimate them (or having the "machine" learned) by minimizing the sum of squared of errors (SSE) between the output and the actual data until a specified level of convergence is achieved. Let Z_t be the actual t-th state of nature. Then the minimization can be stated as in (2).

$$\min_{\alpha_j \beta_{ij}} SSE \sum_{t=1}^{T}\left[Z_t - h\left(\sum_{j=1}^{J} \alpha_j g\left(\sum_{i=1}^{I} \beta_{ij} X_i\right)\right)\right]^2. \quad (2)$$

For this study, the activation functions g(.) and h(.) are chosen to be sigmoid and the back-propagation method. The inputs to the model are one and two years lag GDP, population growth, inflation rate, exchange rate, and political situation.

According TKacz and Hu (1999), ANN can be used in conditions when the linear model cannot solve the problem, or difficult to see relationships between events.

ANN approach consists of 3 stages: training, testing and forecasting. In the first stage, ANN model will be input with past data. This data will make the ANN model become intelligent and suitable with the problem. After ANN get enough learning from the data, this model will come to the next stage, to be tested with another certain set of data. After reaching small enough error rate, the model will be used for future value or data.

One of ANN method is a back propagation. A back-propagation ANN, is trained to perform specific tasks. During the training period, the teacher evaluates whether the ANN's output is correct. If it's correct, the neural weightings that produced that output are reinforced; if the output is incorrect, those weightings are replaced.

In many studies, ANN has proven to be an efficient tool for non-parametric data model in the form of non-linear function, such as business forecasting, credit scoring, bond rating, business failure prediction, medicine, pattern recognition and image processing

3. PREVIOUS RESEARCH

Research conducted by Bambang et. al (1999) using feed-forward artificial neural network technique on Gresik and Gudang Garam stock with data from October 1, 1998 until January 8, 1999, with 10 hidden layer and a coefficient of learning rate with an error value at 0.0001, produced the conclusion that artificial neural network technique can be used to calculate the movement of the stock in the next few days.

Kalyvas and Manchester (2001) prediction of FTSE 500 and S & P 500 using the ANN indicated that ANN can predict the movement of stocks good enough compared Auto Regressive method.

Eliyani (2007) concluded that the backpropagation method is a suitable method to predict the movement of stock.

Research conducted by Dutta et. al. (2006) on Bombay Stock Exchange, with three hidden layers indicated that ANN can be used to help predict stock price movements with reasonable accuracy, i.e., RMSE 4.82 percent.

4. STOCK FORECASTING

This research used the daily closing value ANTM (PT Aneka Tambang) share in November 2007 to December 2007.

	A	B	C	D	E	F	G	H	I	J	K
1	Data	Close (T)		Data	Close (T)		Data	Close (T)		Data	Close (T)
2	1-Nov-07	3650		15-Nov-07	4250		29-Nov-07	4675		13-Dec-07	4450
3	2-Nov-07	3550		16-Nov-07	4225		30-Nov-07	4625		14-Dec-07	4175
4	5-Nov-07	3675		19-Nov-07	4425		3-Dec-07	4575		17-Dec-07	4075
5	6-Nov-07	3875		20-Nov-07	4450		4-Dec-07	4400		18-Dec-07	4075
6	7-Nov-07	4050		21-Nov-07	4350		5-Dec-07	4650		19-Dec-07	4275
7	8-Nov-07	4475		22-Nov-07	4650		6-Dec-07	4475		26-Dec-07	4425
8	9-Nov-07	4525		23-Nov-07	5000		7-Dec-07	4475		27-Dec-07	4475
9	12-Nov-07	4225		26-Nov-07	4900		10-Dec-07	4525			
10	13-Nov-07	4200		27-Nov-07	5050		11-Dec-07	4600			
11	14-Nov-07	4350		28-Nov-07	4800		12-Dec-07	4475			

Figure 1. Closing Value of ANTM (Nov-Dec 2007)

With the help of expert modeler available on SPSS, identification and parameter estimation step can be done automatically. By using the expert modeler, identification obtained as ARIMA (0,1,2) that has a value of p = 0, the value d = 1 and q = 2, with RMSE = 54 218.

Model Description

			Model Type
Model ID	Close (T)	Model_1	ARIMA(0,1,2)

Figure 2. *ARIMA Model*

For ANN, Data years 2004-2007 was prepared in excel format specified variable by variable. Open,

Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1
© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

high, low, close (t-1), mining IDX, JCI and the gold price are the input variables. Whereas for variable output is variable Close (T). This data were divided into two parts, the first 900 data were used in the training process and the remaining 74 data were used in the forecasting process. Mathlab organised the training data and divided into 70 percent for training, 15 percent for validation, and 15 percent for testing.

Determination of hidden layer number, number of neuron and learning rate is performed iteratively by trial and error. This trial and error compares whether there is reduction in errors that occur with the changes.

ANN architecture that was selected is using the 8 neurons for input, one hidden layer with eight neurons in the hidden layer and one neuron in the output (results). Eight input variables that have been normalized were used as data in the input layer neurons. Data close (T) was normalized and was used as a data target/output.

For Transfer function between input-hidden layer and hidden-output layer, we used tan sigmoid. The result of calculation produced MSE of 0.000183.

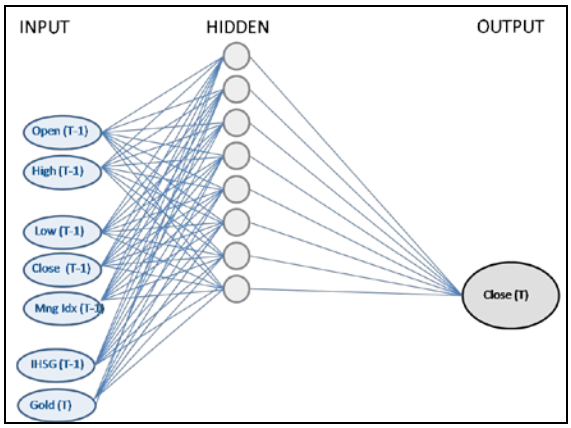

Figure 3. Neural Network Architecture

From the prediction results of ARIMA and ANN prediction results in September 2007 to December 2007, calculated for each sum square value with the actual value. ESS result with ARIMA is 284.95 and ESS result with ANN prediction is 170.40.

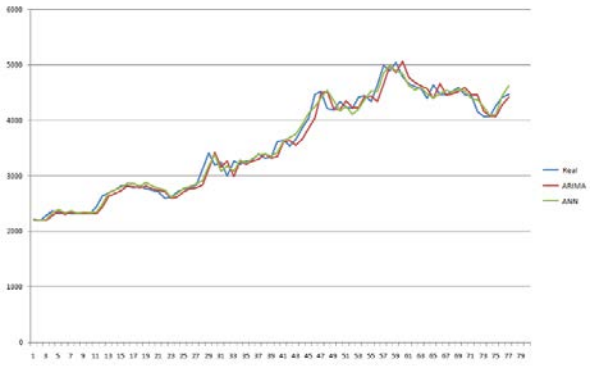

Figure 4. Real VS Arima VS ANN result

5. CONCLUSION

ANN forecasting method has the ability to overcome problems of non-linear data through learning and recognition of the pattern in the data, which cannot be done by ARIMA model. ARIMA only see the price of factors in a few days, and try to predict by fitting the model to existing data, while the ANN was formed more towards the biological way of thinking, whereby the model is inherent in the data, irrespective of the nature of the data, such as the distribution of data.

With ARIMA, forecasting relatively take less time and easier because there is no need for trial and error. With ANN sometimes takes a long time to do trial and error technique, i.e., in determining hidden layers and neuron structure, including suitable learning rate for better MSE value.

The conclusion of this forecasting is ANN method was more accurate than the ARIMA method. This is shown from the ESS result with ARIMA is 284.95 with ANN is 170.40.

For further research it is interesting to look at the possibility of combining ANN and ARIMA or any other statistical forecasting techniques.

REFRENCES:

[1] Eliyani, 2007, Peramalan Harga Saham Perusahaan Menggunakan Artificial Neural Network Dan Akaike Information Criterion, Universitas Muhammadiyah, Gresik

[2] Darmanwan, Dwi Putra, 2005, *Peramalan Bisnis Dengan Metode Box-Jenkins(ARIMA)*, Universitas Udayana, Bali

[3] Kalyvas, Efstathios 2001, *Using Neural Networks And Genetic Algorithms To Predict*

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Stock Market Returns, Department of Computer Science, University of Manchester

[4] Muis, Saludin 2008, Meramal Pergerakan Harga Saham Menggunakan Pendekatan Model ARIMA, Indeks Tunggal & Markowitz, Graha Ilmu, Yogyakarta.

[5] Gonzales, M. E. Q. Neural networks and Mental Representation: An essay on Harmony and Rationality. In: Trans/ Form/Ação, São Paulo, v. 14, p.93-108, 1991.

[6] Dulta, Goutam dan Pankaj Jha dan Arnab Kumar Laha dan Neeraj Mohan, Artificial Neural Network Models for Forecasting Stock Price Index in The Bombay Stock Exchange, Journal of Emerging Market Finance,5:3 (2006)

[7] McHelis, Paul D,Neural Network in Finance: Gaining Predictive Edge in the market, Elsevier academy press,New York, 2005

[8] Hui, S.C dan M.T Yap dan P.Prakash, A Hybrid Time Lagged Network for Predicting Stock Prices, School of Applied Science, Nanyang Technological University, Singapore

[9] Lubis, Chairisni dan Eddy Sutedjo dan Bowo Setiadi, 2005, Prediksi Harga Saham dengan Menggunakan Algoritma Hybrid Neural Network, Fakultas Teknologi Informasi: Universitas Tarumanegara, Jakarta

[10] Demuth, H., Beale, M., Neural Network Toolbox, For Use with MATHLAB, TheMathWorks, 2001.

[11] Bambang, Budi dan Rochani J. Widodo dan Iftikar Z. Sutalaksana, Teknik Jaringan Syaraf Tiruan Feedforward untuk Prediksi Harga Saham pada Pasar Modal Indonesia, Program Pascasarjana teknik Kendali: Institut Teknologi Bandung, Bandung

[12] Wang, Jung Hua dan Jia Yann Leu, Stock Market Trend Prediction Using Arima-Based Neural Network, Department of Electrical Engineering: National Taiwan University, Taiwan, 1996

[13] Hendy Fakhruddin dan *Tjiptono* Darmadji. Pasar Modal Di Indonesia. Jakarta: Salemba Empat, *2001*.

[14] Siang, Jong Jek, Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan Mathlab, Andi Yogyakarta, Yogyakarta, 2009.

[15] Arima pada website
http://www.wahana-statistika.com/index.php/Analisis-Deret-Waktu/arima-autoregressive-integrated-moving-average.html

[16] Data Saham pada www.idx.com

[17] Neural network pada website : http://www.Wikipedia.org/neural_network

[18] Obligasi pada http://id.wikipedia.org/wiki/Obligasi

[19] ARIMA pada http://en.wikipedia.org/wiki/Autoregressive_moving_average_model

[20] Jaringan syaraf tiruan pada www.ilmukomputer.com

[21] Kirkpatrick dan Dahlquist Technical Analysis: The Complete Resource for Financial Market Technicians (Financial Times Press, 2006), page 3.

[22] Murphy, John J., Technical Analysis of the Financial Markets (New York Institute of Finance, 1999), pages 1-5,24-31.

[23] Haykin, Simon, Neural Networks: A Comprehensive Foundation", Prentice Hall , New Jersey,1998.

[24] Fausett, L., Fundamental of Neural Networks, Architecture, Algorithms, and Applications, Prentice Hall, 1994.

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195