Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 4 (2016), pp. 3745–3757

http://www.ripublication.com/gjpam.htm

Indonesian Financial Data Modeling and Forecasting by Using Econometrics Time Series and Neural Network

Sukono
Department of Mathematics,
Padjadjaran University, Indonesia.
E-mail: sukono@unpad.ac.id, fsukono@yahoo.com

Yuyun Hidayat
Departement of Statistics,
Padjadjaran University, Indonesia.
E-mail: yuyun.hidayat@unpad.ac.id

Suhartono
Department of Statistics,
Institute of Sepuluh November, Indonesia. E-mail: suhartono@statistika.its.ac.id.

Brodjol Sutijo
Department of Statistics,
Institute of Sepuluh November, Indonesia.
E-mail: brodjol_su@statistika.its.ac.id

Abdul Talib Bin Bon
Department of Production and Operations,
Universiti Tun Hussein Onn Malaysia, Malaysia.
E-mail: talibon@gmail.com

Sudradjat Supian
Department of Mathematics,
Padjadjaran University, Indonesia.
E-mail: sudradjat@unpad.ac.id

Abstract

In recent years, many researchers have been using neural network (NN) model as an instrument of their research. The motivation behind the application of NN is that it can be used to solve any application problem, such as pattern recognition, signal processing, process control and forecasting. Inference statistics procedure, testing hypothesis, parameter distribution and cross validation have been developed to select the best NN model. The aim of this research is to develop NN modeling in statistic structure modeling and apply it to the inflation data in Indonesia. This research is about optimizing the development of Feed Forward Neural Network (FFNN) model procedure and its application. FFNN Model and Time series model are evaluated based on RMSE, MAD and MAPE. The Best model is the one with minimum RMSE, MAD and MAPE. One of the results is a new procedure based on statistics inference, $R^2_{\text{Incremental}}$ test, in NN model especially FFNN modeling. The comparison between FFNN model and time series model shows that FFNN model is better than time series model.

AMS subject classification: 62P20, 62M10, 92B20.
Keywords: FFNN, inference statistics, $R^2_{\text{incremental}}$, ARIMAX, inflation

1. Introduction

In the beginning, neural network (NN) was designed to model the architecture of the nerves in human brain. However, researches on NN that have been widely conducted recently are, especially, motivated by a promising possibility of using NN as an intrument to overcome various application problems such as pattern recognition, signal processing, process control, and time series forecasting [5], [15]. Feedforward Neural Networks (FFNN) model is considered as an NN model that can be classified as a very flexible model group; and as such it can be used for various applications. A specific form of FFNN model that approaches the non-linear mapping using the sum of Gaussian kernels' values (activation function) is known as Radial Basis Function (RBF) networks [9], [17], [28].

Forecasting the data of time series is one field that widely uses NN model. Haykin [14], Bishop [7] and Kaashoek and Van Dijk [16], were part of researchers that first applied NN model to analyze time series data, i.e: forecasting non-linear signal raised by computer. In the research development, inferential statistical procedures were applied as well, to decide the best FFNN model. Based on Richard et al. [24], that Terasvirta dan Lin in 1993 were the first researchers that applied inferential statistical procedures to obtain the optimum number of units at the hidden layer of FFNN model with single hidden layer. Some latest articles about forming FFNN model by applying the inferential statistical procedures can be found in Kock and Terasvirta [18], Zhang [31], and Bagheri et al. [4].

The procedures for establishing the right, valid, and reliable model is the main problem in modelling. Modelling data of time series is no exception. Within the last decade, NN model has been widely used to solve problems in economic data (inflation) modelling. However, the development of NN model, especially on the procedures for establishing a suitable model that result in an optimum architecture of NN, has not much been done. Therefore, it is necessary to conduct further research to find a new procedure related to the procedure for NN modelling of inflation data, especially for inflation in Indonesia.

Based on the condition and existing problems mentioned above, especially those related to NN modelling of data inflation, this research is aimed to: (1) Study and develop an optimum procedure to find out an input variable that is suitable for NN model; (2) Study and develop an optimum procedure to find out the number of neuron unit in single hidden layer; and (3) Compare and evaluate the accuracy of forecast resulted from the application of NN model and the classical statistical model on data of inflation time series. The following are the benefits expected to be gained from this research: (i) Find a new procedure that will be able to help other researchers using NN as an instrument of their research; and (ii) Find the most suitable model for inflation data in Indonesia within a particular period of time.

2. Theoretical Background

There are two main theoretical bases that will be explained in this chapter, i.e.: the theory of time series analysis, and the theory of feedforward neural networks (FFNN).

2.1. Time Series Analysis

2.1.1 ARMA Process (Autoregressive-Moving Average)

A wide class of a stationary process can be generated by using white noise from a set of different linear equations. This fact leads to an idea of autoregressive-moving average or ARMA process.

A process $\{Y_t, t \in 0, \pm 1, \pm 2, \ldots\}$ is defined as the process of ARMA(p,q) if $\{Y_t\}$ is stationary and if for every t

$$Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{1}$$

with $\{\varepsilon_t\} \sim WN(0, \sigma^2)$. $\{Y_t\}$ is defined as the process of ARMA(p,q) with mean μ if $\{Y_t - \mu\}$ is the process of ARMA(p,q). Equation 1 can be written symbolically in a compact form as under:

$$\phi(B)Y_t = \theta(B)\varepsilon_t \tag{2}$$

with ϕ and θ as polynomial exponents p and q, i.e.:

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \tag{3}$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \tag{4}$$

and B is a backward shift operator that is defined as

$$B^j Y_t = Y_{t-j},\ j = 0, \pm 1, \pm 2, \cdots \tag{5}$$

Sing and Mishra [25], that Box dan Jenkins in 1976 introduced iterative procedures to obtain the order of the most suitable ARIMA model for data of time series that cover: identification step, parameter estimation, diagnostic examination, and forecasting. The complete methodology related to the procedure for establishing ARIMA model can be read in the books written by Tsay [29], Brockwell and Davis [8], and Liu et al. [21].

2.1.2 Transfer Function Model

Transfer Function Model is a model that describes a condition that the future forecast value of a time series (defined as output series or Y_t) is based on the past value of the time series and on one or more related time series (defined as input series or X_t) with the output series. The general form of single-input Transfer Model (x_t) and single-output Transfer Model (y_t) is:

$$y_t = \mu + \nu(B)x_t + n_t \tag{6}$$

with

μ = *constant*
y_t = *representation of stationary output series*
x_t = *representation of stationary input series*
n_t = *representation of error component (noise series) that follows*
= *a particular ARIMA model*
$\nu_(B) = \nu_0 - \nu_1 B - \nu_2 B^2 - \cdots$

Form 6 can be formulated as under as well:

$$y_t = \mu + \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + n_t \tag{7}$$

with

b = *the number of period before input series (x_t) starts to influence output series (y_t)*

$\omega_s(B) = (\omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_s B^s)$ *is the operator of order s, representing the number of previous/past observations x_t that influences y_t*

$\delta_s(B) = (1 - \delta_1 B - \delta_2 B - \cdots - \delta_r B^r)$ *is the operator of the order r, representing the number of past observations of the output series itself that influences y_t.*

The complete theory and methodology related to the procedure for establishing Transfer Function Model can be found in Tsay [29], Brockwell and Davis [8], and Dudek [10].

2.2. Feedforward Neural Networks

The most commonly applied architecture of neural networks (NN) in engineering is Multi-Layer Perceptrons (MLP), also known as Feedforward Neural Networks (FFNN). Figure 2 is an example of a special form of MLP with single hidden layer, known as

FFNN with single hidden layer In this example FFNN is described to have three inputs (i.e.: X_1, X_2, and X_3), four neuron units in the hidden layer with activation function ,and one output unit with linear ativation function [2], [7], [11].

Within this architecture, the response values or output is calculated in the following way:

$$Y(X) = \beta_0 + \sum_{j=1}^{H} \beta_j \psi(\gamma_{j0} + \sum_{i=1}^{I} \gamma_{ji} X_i \tag{8}$$

with $(\beta_0, \beta_1, \ldots, \beta_H, \gamma_{10}, \ldots, \gamma_{HI})$ as network weights or MLP parameters. The non-linear function of $Y(X)$ manifests through a function ψ that is defined as activation function , that is usually a smooth function like logistic function sigmoid [12], [14], [19]:

$$\psi(Z) = \frac{1}{1 + e^{-Z}} \tag{9}$$

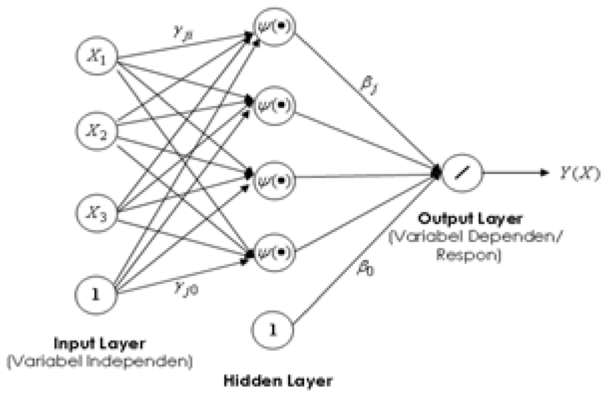

Figure 1: MLP Architecture with single hidden layer, three input units

In its development, NN architecture for time series analysis has come to an architecture that contains recurrent variable. This kind of model is further known as ARMA-RNN model or Autoregressive Moving Average Recurrent Neural Networks. Some articles related to the form of NN model containing the recurrent variable can be found in Andrawis et al. [1], Birau [6], and Erdogan and Goksu [13].

2.3. Asymptotic Characters of FFNN Estimator

Asymptotic characters of backpropagation estimator in FFNN with one hidden layer has been discussed in detail by Moreno et al. [20] and Philip et al. [21]. It has been explained previously that the main goal of network learning is to find out a solution $w*$ to a problem of optimizing arg $\min_{w \in W} Q(w)$ that is:

$$w* = \arg\min_{w \in W}(Q(w) = E((Y - f(X, w))^2/2) \tag{10}$$

with $w*$ as index of optimum network weights.

With a penalty quadratic error, learning on the network has to come to $w*$, that solves:

$$\min_{w \in W} E(Y - f(X, w))^2/2 = E([Y - E(Y|X)]^2/2) + E([E(Y|X) - f(X, w]^2/2) \quad (11)$$

To find $w*$ is a problem, that is exactly the same as to find the optimum parameters of least squares approach for $E(Y|X)$, a pre-conditioned expectation from Y given X [28]. The formal concept required to study the limit distribution (asymptotic) of $\hat{w}_n$ is the concepts of convergence in distribution. Asymptotic distribution of $\hat{w}_n$ depends on the basic character of $W*$. In general $W*$ may comprise isolated dots and/or isolated flat parts. If convergence happens to the flat part, the weight estimations $\hat{w}_n$ will have asymptotic distributions that can be analyzed using the theory about a model that is partially identified. These distributions are groupped as combined asymptotic Gaussian or "limiting mixed Gaussian" (LMG). When $w*$ is unique locally, the model is defined as identified locally and weight estimations $\hat{w}_n$ convergent to $w*$ will have normal multivariate asymptotic distribution [3], [32], [20].

White [30] has given a condition to make sure that $\hat{w}_n$ comes with has normal multivariate asymptotic distribution, that is stated/explained in the following theorem:

Theorem 2.1. (White, [30]) Supposing that (Ω, F, P), $\{Z_t\}$, W, and l are like in Theorem 3.1, whereas $\hat{w}_n \to w*$ as -P with an isolated element at W* in the interior for W.

Considered as an addition, whereas for every z in $\Re^v$, $l(z;)$, can be generated continuously until order 2 at $\int W$; whereas $E(\nabla l(Z_t, w*)'\nabla(Z_t, w*)) < \infty$; that every element of $\nabla^2 l$ is dominated at W a function that can be made integral, and that $A* = E(\nabla^2 l(Z_t, w*))$ and $B* = E(\nabla l(Z_t, w*))\nabla l(Z_t, w*)')$ are non-singular measured matrices $(s \times s)$, with ∇ and ∇^2 as notations of gradient $(s \times l)$ and Hessian operators $(stimess)$ towards w.

Therefore, it comes to $\sqrt{n}(\hat{w}^n - w*) \to N\overset{d}{(}0, C*)$, with $C* = A^{*-1}B * A^{*-1}$. If as an addition, every element of $\nabla l \nabla l'$ is dominated at W by a function that can be made integral, finally $\hat{C}^n \to C*$ a.s. $-P$, with $\hat{C}^n = \hat{A}_n^{-1}\hat{B}_n\hat{A}_n^{-1}$, and

$$\hat{A}_n = \frac{\sum_{t=1}^n \nabla^2 l(Z_t, \hat{w}_n)}{n}, \hat{B}_n = \frac{\sum_{t=1}^n \nabla l(Z_t, \hat{w}_n)\nabla l(Z_t, \hat{w}_n)}{n}.$$

3. Research Methodology

This research comprises two kinds of activities: First, a further theoretical study related to the development of the procedure for establishing an optimum FFNN model for inflation data; Second, an applied study related to the comparative evaluation on the accuracy of FFNN forecast and time series model. Theoretical study is done by analyzing and developing several procedures connected with the establishing of FFNN model to forecast time series. On the applied study, real data, i.e.: inflation data in Indonesia; have been used.

4. Result and Discussion

4.1. The Result of Theoretical Study to Develop FFNN Model

4.1.1 Inferential Statistics from the Additional Contribution R^2

Suhartono et al. [26] introduced a new procedure that is based on inferential statistics from additional contribution. This statistcal test is constructed similarly with the one of the linear model, known as gradual significance test. The test goes through three main steps: Reduced Model, Full Model, and Establishing Statistical Test.

Whereas it is given a Reduced Model, generally it can be written in the form of:

$$Y_t = f(X_t, \hat{w}_n^{(R)}) + \varepsilon_t^{(R)} \tag{12}$$

with l_R as the number of parameters estimated. On the other hand, if, for example, it is given a Full Model that is more complex that the Reduced Model, it becomes:

$$Y_t = f(X_t, \hat{w}_n^{(F)}) + \varepsilon_t^{(F)} \tag{13}$$

with l_F as the number of the parameters estimated, and $l_F > l_R$. Therefore, the test under $H_0 : w^{*+} = 0$, or test towards the additional weight parameter in the Full Model equals zero, can be contructed through test F, that is:

$$F = \frac{(SSE_{(R)} - SSE_{(F)})/(l_F - l_R)}{SSE_{(F)}/(n - l_F)} \sim F_{(v_1=[l_F-l_R], v_2=[n-l_F])} \tag{14}$$

Statistical Test F can be formulated in the following form as well:

$$F = \frac{(SSE_{(R)} - SSE_{(F)})/(df_{(R)} - df_{(F)})}{SSE_{(F)}/df_{(F)}} \tag{15}$$

or

$$F = \frac{R^2_{incremental}/(df_{(R)} - df_{(F)})}{(1 - R^2_{(F)})/df_{(F)}} \tag{16}$$

with $R^2_{incremental} = R^2_{(F)} - R^2_{(R)}$, $df_{(R)} = n - l_R$, as free exponent of Reduced Model, and $df_{(F)} = n - l_F$ as free exponent of Full Model.

The first step of forming the FFNN model is to decide the number of neuron units in the optimum hidden layer. In this step, the decision process is based on statistical significance test F for inference $R^2_{\text{incremental}}$ with a bigger amount of neuron units. After the number of neuron units in the optimum hidden layer is obtained, the next step is to decide an optimum lag input variable. At the stage of obtaining an optimum lag input variable, the decision conducted as the first procedure is done with a forward move that is started with one lag input variable having the greatest R^2 value.

The next step is evaluating he significance of the additional contribution of lag input variable through inference $R^2_{\text{incremental}}$, with statistical test F that is done iteratively until the optimum input variable is obtained.

The implementation of the procedure for establishing the model that is being introduced here can be done simultaneously combined with including the criteria of choosing the best model, for example: Schwarz Information Criteria (SBC) at the evaluating step of the number of neuron units in the hidden layer and in the process of deciding the optimum lag input variable. In addition, the procedure for establishing FFNN model by only using the criteria of choosing the best model can be found out in [14].

4.2. Empirical Studies on the Forming of FFNN Model through Inference $R^2_{\text{incremental}}$

Simulation data generated are the data that follow the model of Exponential Smoothing Transition Autoregressive (ESTAR), i.e.:

$$Y_t = 6.5Y_{t-1}\dot{e}xp(-0,25Y_{t-1}^2) + u_t \tag{17}$$

with $u_t \sim IIDN(0,05^2)$ The application of this procedure is started with FFNN involving six lag input variables $(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-6})$, one constant input and six neuron units in the hidden layer. The results of optimization to obtain the number of neuron units in the hidden layer can bee seen in Table 1 and 2.

Table 1. The result of deciding the number of optimum units in the hidden layer using backward procedure

Steps	The number of units in hidden layer	SBC	R^2	R^2_{incr}	PCA
1	6 units (6 inputs)	-62,1157	0,98668	*	
	without h1		0,79924	0,18744	0,127
	without h2		0,07076	0,91592	0,843
	without h3		0,95844	0,02824	-0,043
	without h4		0,44489	0,54179	-0,371
	without h5		0,66586	0,32082	-0,365
	without h6		0,98010	0,00658	0,017
2	4 units (1 input)	-122,833	0,97545	*	
	without h1		0,95258	0,02288	-0,014
	without h2		0,22184	0,75361	0,737
	without h3		0,40196	0,57350	-0,675
	without h4		0,88360	0,09185	0,041
3	2 units (1 input)	-137,764	0,97246	*	
	without h1		0,36845	0,60588	0,762
	without h2		0,24701	0,72732	-0,610

Table 2. The results of deciding optimum input unit with backward procedure based on Kaashoek and Van Dijk [15], [16]

Steps	Lag unit input	SBC	R^2	R^2_{incr}	PCA
1	6 (lag 1-6)	-62,1157	0,98668	*	
	without lag 1		0,00476	0,98192	0,997
	without lag 2		0,96040	0,02628	0,042
	without lag 3		0,97614	0,01053	0,032
	without lag 4		0,97417	0,01251	-0,014
	without lag 5		0,96432	0,02235	-0,029
	without lag 6		0,97417	0,01251	-0,045

The result of the implementation of further optimization, i.e.: step 2 and 3, can be seen in Table 1. Step 2 of the optimization shows that the 1st and 4th neuron units in the hidden layer can be excluded from the model. Hence, the optimum architecture of FFNN with this data simulation implementing the backward procedure is FFNN with one unit input Y_{t-1}, and two neuron units in the hidden layer or FFNN(1,2,1).

4.3. Study on the Modeling of Indonesian Inflation Data

Inflation data in Indonesia used in this study are those monthly data that were observed starting from January 2009 until April 2015, consisting of 76 observations. Modeling process was done on the first 72 data (known as data training in NN model); while the last 4 data were used to evaluate and to compare the forecast accuracy (data testing). The inflation data can be seen graphically in Figure 2. From this figure, it can be explained that the data has a relatively stationary pattern with a little seasonal variation.

The comparative evaluation of the forecast accuracy to choose the best model is done by implementing the value of Mean Squares of Error (MSE), and the forecast error ratio (represented by MSE). The complete result can be seen in Table 3.

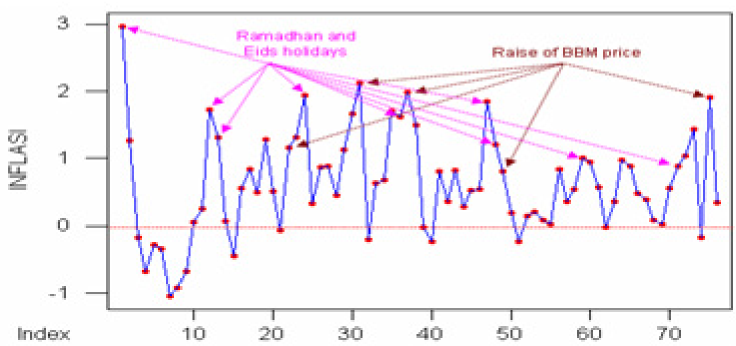

Figure 2: Plot of Time Series of Inflation in Indonesia, January 2009.

Table 4.3 Summary of the forecast comparison dynamically

Forecast Model	MSE (data *testing*)	MSE Ratio (towards FFNN with input like ARIMAX)
▪ ARIMA(1,0,0)(1,0,0)^{11}	0.6826480	3.02
▪ ARIMAX	0.2407240	1.07
▪ FFNN with input like ARIMA	0.4711709	2.08
▪ FFNN with input like ARIMAX	0.2261001	1.00

5. Conclusion

This research has been able to identify a new procedure for choosing FFNN model that can be applied to forecast time series. This becomes the major contribution of this research. The procedure implements Statistical Test F on $R^2_{\text{incremental}}$ with a forward scheme started by deciding the number of units in the hidden layers, folowed by deciding/choosing the optimum input variable. Empirical comparative results related to the forecast accuracy between FFNN model and time series model on the inflation data shows that FFNN model provides a better result than the time series model. However, it is applicable under the condition that the early data processing in FFNN model is done correctly.

Acknowledgement

The authors thank to Faculty of Mathematics and Natural Sciences, Padjajaran University, who have given the facilities to do the research and the publication through their *academic leadership grant (ALG)*.

References

[1] R.R. Andrawis, A.F. Atiya, and H. El-Shishiny. Forecast Combinations of Computational Intelligence and Linear Models for the NN5 Time Series Forecasting Competition. International Journal of Forecasting, 27 (2011), 672–688. www.elsevier.com/locate/ijforecast.

[2] R. Adhikari and R. K. Agrawal. A Homogeneous Ensemble of Artificial Neural Networks for Time Series Forecasting. International Journal of Computer Applications (0975–8887), Volume 32, No. 7, October 2011.

[3] K.G. Akintola, B.K. Alese and A.F. Thompson. Time Series Forecasting with Neural Network: A Case Study of Stock Prices of Intercontinental Bank Nigeria. IJRRAS 9 (3) December 2011. www.arpapress.com/Volumes/Vol9Issue3/IJRRAS_9_3_16.pdf

[4] M. Bagheri, A. Soltani, and L. Torki. Identifying rational speculative Bubbles in Tehran Stock Exchange Total Index Using Neural Network. Applied mathematics in Engineering, Management and Technology 2(3), 2014: 151–157

[5] S. Benkachcha, J. Benhra, and H. El Hassani. Seasonal Time Series Forecasting Models Based on Artificial Neural Network. International Journal of Computer Applications (0975–8887), Volume 116, No. 20, April 2015.

[6] F.R. Birau. Forecasting Financial Time Series Based On Artificial Neural Networks. IMPACT: International Journal of Research in Business Management (IMPACT: IJRBM), ISSN(E): 2321-886X; ISSN(P): 2347–4572, Vol. 2, Issue 4, Apr 2014, 63–66. © Impact Journals.

[7] Bishop, C.M. (1995). Neural Network for Pattern Recognition. Oxford: Clarendon Press.

[8] Brockwell, P.J. and Davis, R.A. (1991). Time Series: Theory and Methods, 2nd edition. New York: Springer Verlag.

[9] C.L. Cocianu and H. Grigoryan. An Artificial Neural Network for Data Forecasting Purposes. Informatica Economic, vol. 19, no. 2/2015.

[10] G. Dudek. Generalized Regression Neural Network for Forecasting Time Series with Multiple Seasonal Cycles. adfa, p. 1, 2011. Springer-Verlag Berlin Heidelberg 2011.

[11] S. F. Crone, M. Hibon, and K. Nikolopoulos. Advances in Forecasting with Neural Networks Empirical Evidence from the NN3 Competition on Time Series Prediction. International Journal of Forecasting 27 (2011) 635–660. www.elsevier.com/locate/ijforecast

[12] J. P. Donate, and X. Li. Time Series Forecasting by Evolving Artificial Neural Networks with Genetic Algorithms, Differential Evolution and Estimation of Distribution Algorithm. Neural Comput & Applic (2013) 22:11–20. DOI 10.1007/s00521-011-0741-0. Springer-Verlag London Limited 2011.

[13] O . Erdogan, and A. Goksu. Forecasting Euro and Turkish Lira Exchange Rates with Artificial Neural Networks (ANN). International Journal of Academic Research in Accounting, Finance and Management Sciences, Vol. 4, No.4, October 2014, pp. 307–316, E-ISSN: 2225-8329, P-ISSN: 2308-0337. 2014 HRMARS. www.hrmars.com

[14] S. Haykin. Neural Networks. A Comprehensive Foundation. Second Edition. Prentice-Hall, Inc. New Jersey. 1999.

[15] Kaashoek, J.F. and Van Dijk, H.K., (2001). Neural Networks as Econometric Tool. Report EI 2001–05, Econometric Institute Erasmus University Rotterdam.

[16] Kaashoek, J.F., and Van Dijk, H.K. (2002). Neural Network Pruning Applied to Real Exchange Rate Analysis. Journal of Forecasting, 21, pp. 559–577.

[17] M. Khashei, and M. Bijari. An Artificial Neural Network (p, d,q) Model for Time-series Forecasting. Expert Systems with Applications 37 (2010) 479–489. journal homepage: www.elsevier.com/locate/eswa.

[18] A.B. Kock and T.Terasvirta. Forecasting the Finnish Consumer Price Inflation Using Artificial Neural Network Models and Three Automated Model Selection Techniques. Finnish Economic Papers, Volume 26, Number 1, Spring 2013.

[19] N. Kumar, and G. K. Jha. A Time Series ANN Approach for Weather Forecasting. International Journal of Control Theory and Computer Modeling (IJCTCM), Vol.3, No.1, January 2013.

[20] H. Li, D. Pi and M. Jiang. Forecasting Chaotic Time Series with Wavelet Neural Network Improved by Particle Swarm Optimization. International Journal of Hybrid Information Technology, Vol.6, No.6 (2013), pp. 369–380. http://dx.doi.org/10.14257/ijhit.2013.6.6.33.

[21] N. Liu, V. Babushkin, and A. Afshari. Short-Term Forecasting of Temperature Driven Electricity Load Using Time Series and Neural Network Model. Journal of Clean Energy Technologies, Vol. 2, No. 4, October 2014.

[22] J. J. M. Moreno, A. P. Pol, and P. M. Gracia. Artificial Neural Networks Applied to Forecasting Time Series. Psicothema 2011. Vol. 23, no. 2, pp. 322–329, ISSN 0214 - 9915 CODEN PSOTEG. www.psicothema.com. Copyright 2011 Psicothema.

[23] A. A. Philip, A. A. Taofiki, and A. A. Bidemi. Artificial Neural Network Model for Forecasting Foreign Exchange Rate. World of Computer Science and Information Technology Journal (WCSIT), Vol. 1, No. 3, 110–118, 2011, ISSN: 2221-0741.

[24] K. Richard, Waititu A., and Wanjoya A. Artificial Neural Network Application in Modeling Revenue Returns from Mobile Payment Services in Kenya. American Journal of Theoretical and Applied Statistics, 2014; 3(3): 60–64.

[25] A. Singh and G.C. Mishra. Application Of Box-Jenkins Method And Artificial Neural Network Procedure For Time Series Forecasting Of Prices. STATISTICS IN TRANSITION new series, Spring 2015, Vol. 16, No. 1, pp. 83–96.

[26] Suhartono, Subanar and Guritno, S. (2006). Model Selection in Neural Networks by Using Inference of $R^2_{\text{Incremental}}$, PCA, and SIC Criteria for Time Series Forecasting, JOURNAL OF QUANTITATIVE METHODS: Journal Devoted to The Mathematical and Statistical Application in Various Fields, Vol. 2, No. 1, 41–57.

[27] A. Tarsauliya, S. Kant, R. Kala, R. Tiwari, and A. Shukla. Analysis of Artificial Neural Network for Financial Time Series Forecasting. International Journal of Computer Applications (0975-8887), Volume 9, No.5, November 2010.

[28] K. C. Tseng, O. Kwon, and L. C. Tjung. Time series and neural network forecast of daily stock prices. Investment Management and Financial Innovations, Volume 9, Issue 1, 2012.

[29] R.S. Tsay. Analysis of Financial Time Series. Second Edition. John Miley & Sons, Inc, Publication. Hoboken, New Jersey. 2005.

[30] White, H. (1999). Asymptotic Theory for Econometricians. Academic Press Inc., New York.

[31] G. P. Zhang. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Neurocomputing 50 (2003), 159–175. www.elsevier.com/locate/neucom.

[32] G. P. Zhang, and V.L. Berardi. Time Series Forecasting with Neural Network Ensembles: an Application for Exchange Rate Prediction. Journal of the Operational Research Society (2001) 52, 652–664. www.palgrave-journals.com/jors.