FORECASTING NIGERIA FOREIGN EXCHANGE USING ARTIFICIAL NEURAL NETWORKS

Alhassan[1], J. K.; Misra[2], S.; Ogwueleka[3], F.; & Inyiama[4], H. C.
[1, 3, 4] School of Information and Communication Technology
Federal University of Technology, Minna, Nigeria
[2]Atilim University, Ankara, Turkey
E-mail: jkalhassan@yahoo.com
Phone No: +234-803-596-1620

Abstract
Artificial neural networks (ANNs) are computing models for information processing and pattern identification. They grow out of research interest in modeling biological neural systems, especially human brains. An ANN is a network of many simple computing units called neurons or cells, which are highly interconnected and organized in layers. Each neuron performs the simple task of information processing by converting received inputs into processed outputs. In past two decades, ANN has been applied in Economics, Finance and other sectors. The foreign exchange market assists international trade and investment by enabling currency conversion. In this study we applied a time-delayed neural network model for forecasting daily foreign exchange rate of a US Dollar to Naira for Nigeria by using Artificial Neural Network (ANN) methodology on the basis of daily data for September 2011 to February 2012. We compared ANN with Single Exponential Smoothening (SES) and Autoregressive-Integrated-Moving-Average (ARIMA) models, the ANN forecasting tool proved to be more accurate than the SES and ARIMA as it had a smaller root mean squared error of 0.6995 as compared to the root mean squared error of the SES which was 0.9890 and ARIMA which was 0.7880. More research work can be carried out by comparing ANN with other available forecasting tools.

Key words: Artificial Neural Networks, Forecasting, Foreign Exchange Rate, Single Exponential Smoothening

Introduction
To forecast mean to predict. Forecasting is a process that produces a set of results by given a set of variables. Usually, the variables are past data. Fundamentally, forecasting assumes that future occurrences are depended, at least in fraction, on currently observable or long-ago events. Its assumption is that a number of aspects of the precedent patterns will keep into the future. Artificial Neural Networks (ANNs) are computing models for information processing and pattern identification. Artificial Neural Networks (ANNs) are branch of Artificial Intelligence (AI) systems. They are capable of correlating input and corresponding output data (Dayhoff & DeLeo 2001), (Piuri & Scotti, 2004), (Perantonis, *et al*, 1998) and (Huang & Ma, 1999), as a result they have become powerful tool in various applications. There are many organizations or institutions, for example: CZECH National Bank (Marek *et al*, 2005), Bank of Canada (Greg & Hu, 1999), Bank of Jamaica (Serju, 2002), to mention few, that are currently using their forecasting models based on ANN methodology for predicting various macroeconomic indicators. The capability of neural networks in modeling linear time series has been studied and confirmed by a number of researchers (Hwang, 2001), (Medeiros & Pedreira, 2001), (Zhang, 2001). There are numerous research works carried out using ANN (Nakamura, 2006), (Haider & Adnan, 2007), (Alhassan & Sanjay, 2011) to mention few.

The form of exchange rate for the global decentralized trading of international currencies is the foreign exchange rate market. Apart from weekends, financial centers around the world function as anchors of trading between a wide range of different types of buyers and sellers around the clock. It is the foreign exchange market that determines the relative values of different currencies (O'Sullivan & Steven 2003). The foreign exchange market assists international trade and investment by enabling currency conversion. It also supports direct speculation in the value of currencies, and the carry trade speculation based on the interest rate differential between two currencies.

The main purpose of this paper is to forecast daily foreign exchange rate of US Dollar to Naira for Nigeria by using ANN methodology on the basis of available daily data from September 2011 to February, 2012. We also compare the forecast performance of the ANN model with that of Single Exponential Smoothening (SES) and Autoregressive-Integrated-Moving-Average (ARIMA) based models, because of problem of data over-fitting with SES and ARIMA. It is observed that forecasts based on ANN are more precise than those based upon SES and ARIMA models. The rest of the paper is organized as follows: Neural network methodology is presented in section 2. Section 3 provides data and methodology. Forecasting evaluation is discussed in section 4 and the last section is on the summary of our findings.

Literature Review

Artificial Neural Networks (ANN)
Usually called "neural network" (NN), is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data (Zhang, 2004).

Like human beings, ANN learns by experience. An ANN is configured for a specific application, such as pattern recognition and time series forecasting, through a learning process. In biological systems, learning involves adjustments to the synaptic connections that exist between the neurons. The basic building block of a brain and the neural network is the neuron. Figure 1 shows the human neuron.

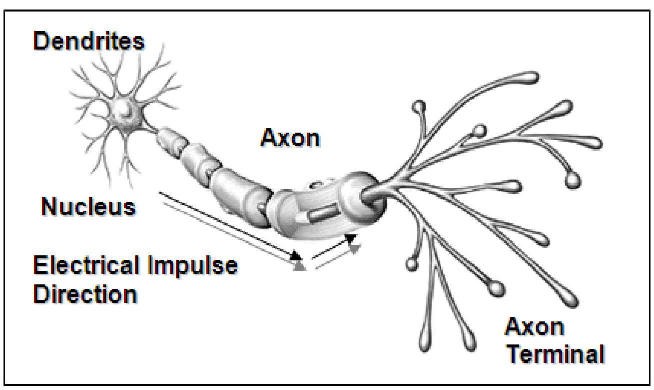
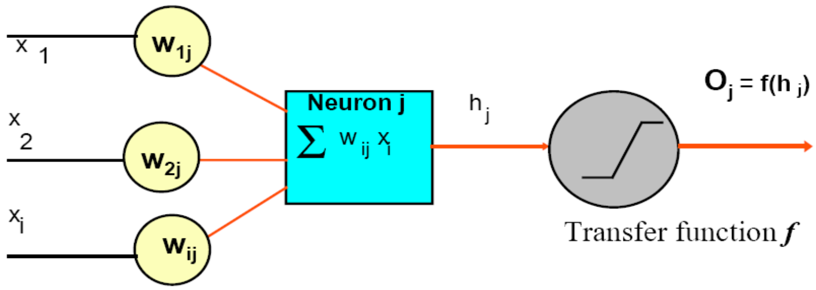


Figure 1: Biological Model of Human Neuron (Adapted from Haykin & Simon, 1994)

According to Beal and Jackson (1990), all inputs to the cell body of the neuron turn up along *dendrites*. Also, dendrites can act as outputs that interconnect inter-neurons. The function of dendrite is mathematically approximated as a summation. On the other hand, *Axons,* with electrical potential are found only on output cells. If excited, past a threshold, it will transmit an electrical signal. Axons terminate at *synapses* that connect it to the dendrite of another neuron. The neuron sends out spikes of electrical activity through a long axon, which splits into thousands of branches. Each branch a structure called a *synapse* at the end which converts the activity from the axon into electrical effects that inhibit or excite activity from the axon into electrical effects that inhibit or excite activity in the connected neurons. As soon as a neuron receives excitatory input that is adequately large compared with its inhibitory input, it sends a spike of electrical action down its axon. By changing the effectiveness of the synapses so that the influence of one neuron on other changes, learning occurs. The human brain contains approximately 10 billion interconnected neurons creating its massively parallel computational capability.

Neuron Model

A neuron may be divided into three parts for analysis purpose. First, the input connections, second the summing and activation functions and lastly, the output connections. As shown in figure 2.



. **Figure 2:** An Artificial Neuron (Adapted from Zhang (2004) Documentation).

A neuron is connected to other neurons in artificial neural network and process the information it receives from them. No limit to the amount of connections a neuron may receive information from. The weights are used to regulate the information that a neuron receives from others. When a neuron receives information from other neurons, each portion of information is multiplied by a weight with a value between -1 and +1, which allows the neuron to judge how important the information it receives from its input neurons is. These weights are essential to the way a network works and is trained: in a reality, training a network means modifying all the weights and regulating information flow to ensure output follows the given criteria, for instance, minimization of root mean squared error or moving average error

Summing and Activation Functions

Summing and activation functions are the second part of a neuron. The information sent to the neuron and multiplied by corresponding weights is added together and used as a parameter within an activation function. A neuron becomes activated when it detects electrical signals from the neurons it is connected to in biological context (Beale & Jackson, 1990). If these signals are sufficient, the neuron will become "activated or excited" - it will send electrical signals to the neurons connected to it. There many activation functions used in ANN literature, but we will discuss the one which we used and that is hyperbolic tangent function: as in equation 1, a continuous function with a domain of (-∞,∞) and a range of (-1, 1):

$$\text{hyperbolic tangent: } f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$ .... Equation 1

It is perfect for predicting whether or not inflation will rise (tanh(x) = 1) or fall (tanh(x) = -1), because it provides a function with a limitless domain and a range of (-1, 1).

Output Connections

Lastly, once the activation function returns a matching value for the summed inputs, these values are sent to the neurons that use the current neuron as an input. This process repeats again, with the current neuron's output being summed with others, and more activation functions accepting the sum of these inputs. It can only be ignored is if the current neuron is an output neuron. After which the summed inputs and normalized sum is sent as an output and not processed again.

Neural Network Architecture: Multilayer Perceptron

This was first introduced by M. Minsky and S. Paper in 1969. It is a special case of perceptron whose first-layer units are replaced by trainable threshold logic units in order to allow it to solve nonlinear separable problem. Minsky and Papert (Zhang, 2004) called multilayer perceptron of one trainable hidden layer. Each layer is fully connected to the next one. Depending on the complexity, performance and implementation point of view, the number of hidden layers may be increased or decreased with, corresponding increase or decrease in the number of hidden units and connections. Both the perceptron and the multilayer perceptron are trained with error-correction learning (Principe *et al*, 2000). But since perceptron does not have an explicit error available, this stopped further work on multilayer perceptron around 1970 until a method to train multilayer perceptron was discovered. The method is called back propagation or the generalized delta rule. With this method, processing is done from the input to the output layer, that is, in the forward direction, following which computed errors are then propagated back in the backward direction, to change the weights to obtain a better result. The algorithm has been rediscovered several times with some variations (Beeker & Le Cun, 1989). The theory about the derivation of back propagation learning rule can be found in the work of Rumelhart (Rumelhart *et al*, 1986) and Drakos (2003). Its structure is shown in figure 3.

The back-propagation (BP) algorithm is a generalization of the delta rule that works for networks with hidden layers. It is by far the most popular and most widely used learning algorithm by ANN researchers. Its popularity is due to its simplicity in design and implementation.

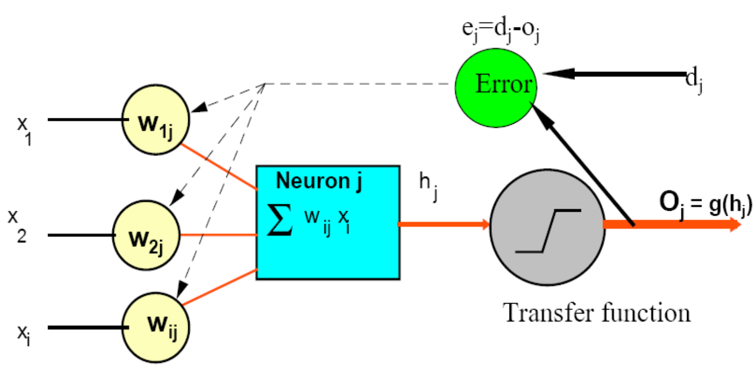


Figure 3: Back-propagation of errors for a single neuron j. (Adapted from (Zhang, 2004).

The algorithm for the back propagation is as follows:

(i) Perform the forward propagation phase for an input pattern and calculate the output error.

(ii) Change all weight values of each weight matrix using the formula
Weight (new) = weight (old) + learning rate * output error * output (neuron i) * output (neuron I + 1) * (1 – output (neurons i+1))

(iii) Go back to step 1

(iv) The algorithm ends, if all output patterns match their target patterns

When backpropagation algorithm is used for weight change, the state of the system is doing gradient descent; moving in the direction opposite to the largest local slope on the performance surface. That is, the weights are being updated in a downward direction. The backpropagation algorithm is general, widely used and not complex, for raining multilayer feedforward networks.

Data and Methodology

Data Used

The main aim of this study is to forecast daily foreign exchange rates for Nigeria. The exchange rate of a US Dollar to Naira for the period of September 2011 to February 2012 using a time delayed artificial neural network model with 3 hidden layers. We used data on daily basis from 26 September, 2011 to 23 February, 2012. Figure 4 represents graphically the data we have used.

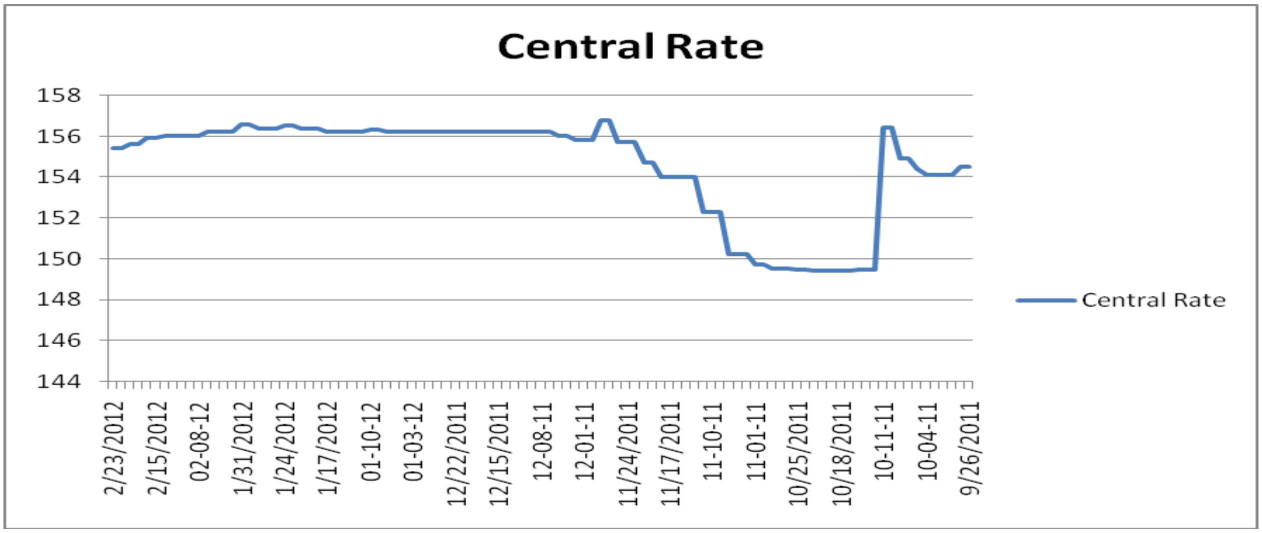


Figure 4: Graphical Representation of the data used

Model Description

We estimate a very time delayed neural network for foreign exchange rate based on *'feedforward with backpropagation'* architecture as in equation 2.

$$F_{t+j} = \sum_{k=1}^{n} \phi_k \text{Tanh}\,(w_k x_{t-1} + b_k) \qquad \text{...... Equation 2}$$

Where: Ft+j is the neural network foreign exchange rate forecast *j* days ahead, xt-1 is a vector of lagged inflation variables [Ft-1, Ft-2] tanh which is the hyperbolic tangent function used as transformation process. ф's are layer weights, wi are input weights and b's are biases. The implementation is as follows; A Time-delayed neural network model was used to forecast the foreign exchange rate of US-Dollar to naira. This model is built to capture the relationship between the historical exchange rates and next week's exchange rate. In this model, the normalized exchange rates of the previous periods are fed to a neural network so as to forecast the next period exchange rate. The formula for normalizing the raw data is given in equation 1. The inputs to the neural network are FX_{i-4}, FX_{i-3}, FX_{i-2}, FX_{i-1}, and FX_i, while the output of the neural network is FX_{i+1}, the next day's exchange rate where *FXi* stands for the current day's exchange rate. The FX_{i-1} means the current day exchange rate minus the previous week exchange rate, FX_{i-2} stand for the previous week exchange rate minus the previous two week exchange rate, as so on see equation 3.

$$Nm = \frac{2 * Y - (Max + Min)}{Max - Min} \qquad \text{........... Equation 3}$$

We used equation 1 to achieved the original scaling of data within the range [-1 , +1]. Results are shown here in figure 5.

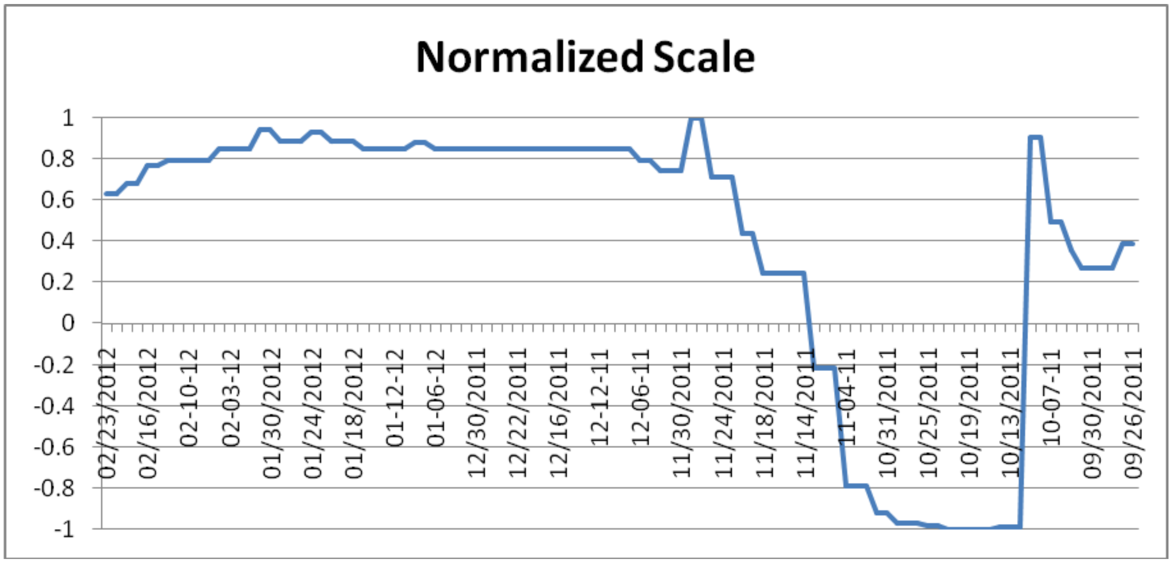


Figure 5: Foreign exchange rate in normalized scale [+1, -1]

Network Architecture

The architecture of this neural network is 5-3-1 where 5 represents the number of inputs to the network, 3 represent the number of hidden layers and 1 represent the number of output as shown in Figure 6. We use NMSE to measure the performance of this network.

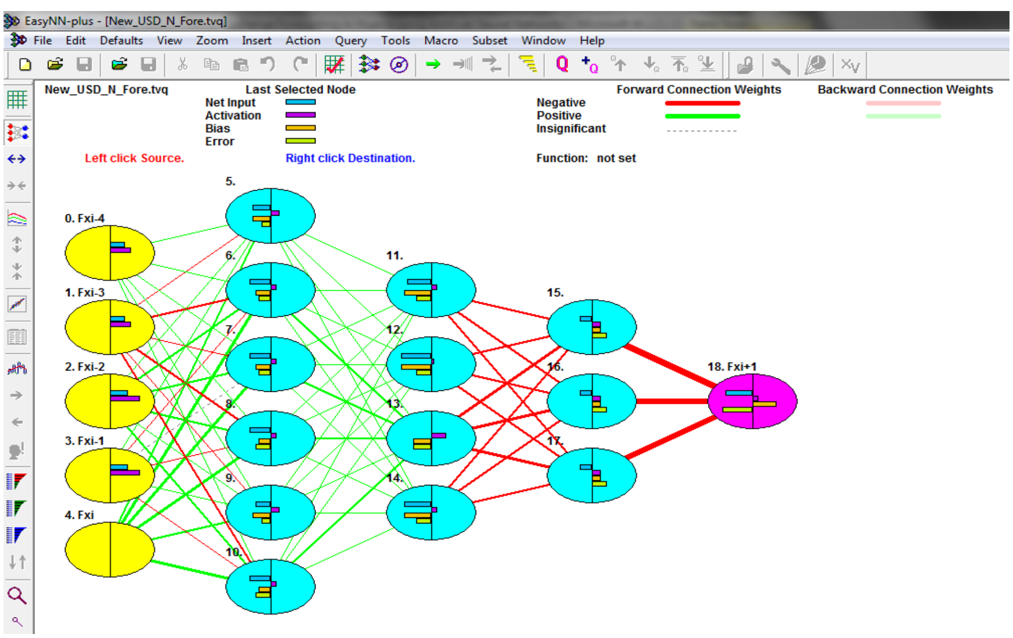


Figure 6: Architecture of the Artificial Neural Networks

Forecasting Evaluation

We calculate the root mean of squared errors (RMSE) to evaluate the forecast with the equation 4. The training algorithm is run on the training set until the RMSE starts to increase on the validation set.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{it} - Y_{it})^2}{n}} \quad \text{.. . . . . Equation 4.}$$

Our neural network was trained using the Easy Neural Networks. Before training, this model requires some default values, which are given in table 1.

Table 1: Neural network parameters used

S/NO	PARAMETER	VALUE
1	Hidden Layers	3
2	Training rows	101
3	Input Columns	5
4	Input Nodes connected	5
5	Learning Rate	0.6000
6	Learning Cycles	401
7	Target Error	0.0100
8	Momentum	0.8000

Comparing Forecast Performance of Ann and Ses and Arima Based Models

Compared the performance of the ANN with Single Exponential Smoothening (SES) and Autoregressive-Integrated-Moving-Average (ARIMA) using the same set of data. Results based on ANN methodology as well as both SES and ARIMA methodologies are presented in table 2. We evaluated forecasting performance on the basis of RMSE criteria. We observed that RMSE of ANN based forecasts is less than the RMSE of forecasts based on SES and ARIMA models. At least by this criterion forecast based on ANN are more precise.

Table 2: Comparison of the forecasting tools

TOOL	MIN. ERROR	RMSE
TNN	0.0000	0.6995
ARIMA(1,1,1)	0.2460	0.7880
SES	0.7880	0.9890

Conclusion

In this paper, a time-delayed artificial neural network model applied to forecast daily foreign exchange rate of a US Dollar to Naira for Nigeria by using ANN methodology on the basis of daily data for September 2011 to February 2012. The main reason of this work is to find reliable out-of-sample forecast based on RMSE minimization criteria, in which error instability is minimized after training network with 3 hidden layers. The leaning rate of our model is 0.6. Feedforward with backpropagation methodology is used as model simulation; this requires an activation function which used generalized delta rule. From our forecast result, foreign exchange rate of a US Dollar to Naira for the end of next few days in February 2012 is on average high as compared with past days.In the end, we compared ANN with SES and ARIMA models, the ANN forecasting tool proved to be more accurate than the SES and ARIMA as it had a smaller root mean squared error of 0.6995 as compared to the root mean squared error of the SES which was 0.9890 and ARIMA which was 0.7880. More research work can be carried out by comparing ANN with other available forecasting tools.

References

Alhassan, J. K & Sanjay, M. (2011). Using a Weightless Neural Network to Forecast Stock Prices: A case study of Nigerian Stock Exchange. *Scientific Research and Essays*, 6(14) P.2934-2940 Online: http://www.academicjournals.org/SRE

Beale, R., & Jackson, T., (1990). *Neural computing: An introduction, adam hilger, bristol england.* Cited on 4th June, 2012 from Online at http://mpra.ub.uni-muenchen.de/14645/ MPRA Paper No. 14645, posted 14. April 2009 / 10:39

Bose, N. K & Liang, P. (1996). *Neural networks fundamentals with graphs, algorithm and applications*. McGraw-Hill Inc.Pp30-36.

Beeker, S. & Le Cun, Y. (1989). *Improving the convergence of the backpropagation learning with second-order methods.* In Hinton, G., Touretzky, D and Sejnowski, T. editors, Proc. 1988 Connectionist Models Summer School, Pittsburgh, pp.29-37. Morgan Kaufmann, Sun Mateo, CA.

Dayhoff J. & DeLeo J.(2001). Artificial neural networks opening the black box. *Cancer Supplement, 91(8), 1615-1635.*

Drakos, N. (2003). *Backpropagation.* http://ebl.leeds.ac.uk/nikos/pail/intml/subsection3.11.4.html. 20th March, 2003

Greg, T., Sarah H., (1999). *Forecasting GDP growth using artificial neural networks, working paper 99-3*, Bank of Canada. Cited on 4th June, 2012 from Online at http://mpra.ub.uni-muenchen.de/14645/ MPRA Paper No. 14645, posted 14. April 2009 / 10:39

Haykin, S., (1994). *Neural networks: A comprehensive foundation, Macmillian College Publishing Company,* New York. Cited on 4th June, 2012 from Online at http://mpra.ub.uni-muenchen.de/14645/ MPRA Paper No. 14645, posted 14. April 2009 / 10:39

Huang, D. S. & Ma, S. D. (1999). Linear and nonlinear feedforward neural network classifiers: A comprehensive understanding. *J. Intelligent Syst.*, 9,1–38.

Hwang, H. B. (2001). Insights into neural network forecasting of time series corresponding to ARIMA (p,q) Structures. *Omega, 29, 273-289.*

Haider, A. & Adnan, H. (2007). *Inflation forecasting in pakistan using artificial neural networks.* MPRA Paper. Cited on 4th June, 2012 from Online at http://mpra.ub.uni-muenchen.de/14645/ MPRA Paper No. 14645, posted 14. April 2009 / 10:39

Marek, H., Michael, K. & Josef, C. (2005). *The application of structured feedfrward neural networks to the modeling of daily series of currency in circulation.* Working paper series 11, Czech National Bank. Cited on 4th June, 2012 from Online at http://mpra.ub.uni muenchen.de/14645/ MPRA Paper No. 14645, posted 14. April 2009 / 10:39

Medeiros, M. C. & Pedreira, C. E. (2001). What are the effects of forecasting linear time series with neural networks? *Engineering Intelligent Systems*, 4, 237-424.

Nakamura, E., (2006). *Inflation forecasting using a neural network.* Economics Letter, 86 (3), pp 373-378.

O'Sullivan, A. Steven M. S. (2003). *Economics: Principles in action.* Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. pp. 458. ISBN 0-13-063085-3

Perantonis, S. J., Ampazis, N., Varoufakis, S. & Antoniou, G. (1998). Constrained learning in neural networks: Application to stable factorization of 2-D polynomials. *Neural Processing Lett.*, *vol. 7, pp. 5–14.*

Piuri V. & Scotti, F. (2004). Morphological classification of blood leucocytes by microscope images. In *2004 IEEE Int. Conf. Computational Intelligence for Measurement Systems and Applications*, pp. 103-108.

Principe, J. C., Eliano, N. R, & Lefebvre, W. C. (2000). *Neural networks and adaptive systems: Fundamentals through simulation.* John Wiley and Sons, Inc.Pp40-50.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the microstructure of Cognition edited by Rumelhart, McClelland and the PDP Research Groups Vol.1*, MIT Press, Cambridge Mass., USA. pp. 216-271.

Serju, P. (2002). *Monetary conditions & core inflation: An application of neural networks.* Working Paper, Research Services Department Research and Economic Programming Division Bank of Jamaica. 1-23. Downloaded on 6/6/2012 from http://www.cemla.org/old/pdf/red/RED_VII_JAMAICA-Prudence-Serju.pdf

Zhang, G. P. (2001). An investigation of neural networks for linear time-series forecasting. *Computer & Operation Research, 28, 1183-1202.*

Zhang, G. P. (2004). *Artificial neural networks for business forecasting.* IRM Press, Hershey PA, USA. P3.