INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

Neural Network: An Alternative Statistical Model for Predicting Financial Time Series Data

Mbaga, Yohana Vandi
Assistant Lecturer, Department of Statistics and Operations Research,
Modibbo Adama University of Technology, Yola, Nigeria

Abstract:
This study presents the predictive performance of Neural Network Model. Nigerian Naira (NGN) exchange rates against Ghana Cedi (GHC) using daily exchange rates values were used. The data set were divided into three in the ratio of 3:1:1 for training (parameter estimation), test and validation respectively. Neural Network (NN) Model with back propagation training algorithm using descent gradient minimisation technique and logistic activation function was developed. The architecture for NN Model was determined through Automatic Network Search (ANS). The tuning parameters considered for the training of NN Model are the learning rate and momentum with the values (0.1 and 0.3) and (0.5 and 0.5) for model A and B respectively. The performance metrics considered for the evaluation of the Models is Mean Square Error (MSE) and Mean Absolute Error (MAE). The results show that the performances of NN increase with increase in parameter values, indicating that higher learning rate and momentum values facilitate better convergence.

Keywords: *Foreign exchange rates, neural network, prediction.*
JEL classification: C22, C32, O57

1. Introduction

Various significant structural transformations between 1960 and early 1970s led to the dramatic end of the Breton-woods system of pegged-exchange rates. Numerous efforts to bring back the fixed exchange rate system proved futile and by March 1973, the regime of floating currencies began (Kumar, 2010). Due to the introduction of floating exchange rates and the rapid expansion of the global trading markets over the last few decades, the foreign currency exchange market has experienced unprecedented growth (kamruzzaman and Sarker, 2004). Foreign exchange market is the largest and most liquid of the financial markets with an estimated $1 trillion traded every day (Huang *et al.,* 2004).

These foreign exchange rates provide significant data necessary for currency trading in the international market. Forecasting exchange rate is an important financial problem that has received much attention especially because of it intrinsic difficulty and practical applications (Huang, et al. 2004; and Kodogiannis, 2001). Financial forecasting is an example of signal processing problem which is challenging due to small sample, high noise, non-stationary and non-linearity (Lee *et al.,* 2001). Huang et al. (2004) affirmed that exchange rates series exhibit high volatility, complexity and noise that result from an inclusive market mechanism generating daily observations. Due to the volatility of foreign exchange rates, there is a need to have well-organized power and risk management tool, in order to declare the future trend of exchange rates which could serve as a guide to investors.

Exchange rate prediction though one of the challenging applications of modern time series forecasting, is very important for the success of many businesses and financial institutions (Kamruzzaman et al. 2003). The difficulty of predicting foreign exchange rates due to their high volatility and complexity has long been an imperative concern in international financial market as many econometric models are unable to produce significantly better forecasts than the random walk model (Yu *et al.,* 2005).The superiority acclaimed to random walk is on the basis of the efficient market hypothesis (EMH), which postulates that current prices reflect all relevant knowledge and information about the market and therefore future returns are essentially unpredictable. There is now strong evidence that exchange rate returns are not independent of the past changes (Tenti, 1996). Tenti (1996) further noted that the prevalent view in economic literature that exchange rate follow a random walk has been dismissed by recent empirical works.

The need to obtain a better predictive model for foreign exchange rate due to the vital role it plays in the global market, has been a driving force of several researches in econometrics. Beside, in the recent years, it is not uncommon to be faced with data set involving perhaps millions of observations and hundreds of predictors which will render the conventional forecasting models inefficient (Kutner *et al.,* 2005). Kutner et al., (2005) further noted that the exponential growth in available data has motivated researchers in the field of statistics, artificial intelligence and data mining to develop simple flexible, powerful procedures for data modeling that can be applied to very large set. One of such forecasting techniques that allows for the detection and modeling of non-linear data is neural networks (kodogiannis, 2001).

Neural network (NN) techniques are prime candidates for prediction purpose (Philip et al. 2011). Neural network models are extremely flexible and can be used to represent a wide range of response surface shapes. Artificial neural network is a complex and sophisticated computer program. It is able to adapt to recognize patterns to generalize and to cluster or to organize data (Gould, 2004). The use of Artificial Neural Network (ANN) in exchange rates prediction comes in because they can learn to detect complex pattern in data (Philip et al. 2011).

A fair amount of literature has been generated on the use of ANNs in time series forecasting (Akinwale et al., 2009). Kutsurelis, (1998) also noted that the potential use of NN as a tool for predicting financial markets hasbeen marketed at increasing levels in recent years. Results from these studies indicate a surprise performance of ANN models as compared to traditional statistical models (Nag and Mitra, 2002).

Variability in the exchange rates means huge losses or profits for importers, exporters, and traders in foreign exchange markets (Gujarati 2004). Hence, to maintain and enhance business relationship between two countries, there is a need to build a model that can forecast their exchange rates with optimum accuracy. In view of this, the need to model the exchange rates of Nigerian Naira (NGN), against one of her African major trade-partners (Ghana) becomes imperative.

In this study, we modeled NGN foreign exchange rate using high-frequency data of one thousand, nine hundred and thirty-five (1,935) daily observation obtained online from http://www/oanda.com against Ghana cedi (GHC).

2. Neural Network

In a quest to build an intelligent machine in the hope of achieving human like performance in the field of speech and pattern recognition, natural language processing, decision making in fuzzy situation etc., we have but one natural occurring model: The human brain itself, a highly powerful computing device (Kamruzzaman and Sarker, 2003). Human brain consists of hundred billions of neurons; each neuron is connected to other neurons through ten thousand synapses. The neuron (which is formed by dendrites (input), axon (output) and cell body) is the basic computational unit in the nervous system.

Neural network is an information processing paradigm that is inspired by the way biological nervous systems, such as a brain, process information. The function of an artificial neural network is to produce an output pattern using an input pattern. Inputs enter the neuron and are multiplied by their respective synaptic weights. The weights are then summed and processed by an activation function. The activation function is usually chosen to be sigmoid because it is flexible and can be adapted to a variety of circumstances. These functions are often called "squashing" functions, because they compress an infinite input range into a finite output range. The activation functions transform the incoming signals from the neurons of the previous layer using a mathematical function. The input neurons usually have no activation function rather they use the identity function i.e.; the inputs are not transformed at all. Instead they are combined in a weighted sum and pass on to the neurons in the layer above usually called the hidden layer.

The neural network essentially looks at repeated examples (or input observations) and recall patterns appearing in the input along with each subsequent response. A key in training the network is to find the weights to go along with the activation functions. With an input vector $X_{t-i} = (x_{t-1}, x_{t-2}, ..., x_{t-p})$, each of the input nodes codes the data and fires a signal across the edge to the hidden nodes in the next layer, where they are combined into new values using node specification weights, α_{jk} for node j. The new value computed at node j from an input vector X_{t-i} is

$$H_{k,j} = h_j(\alpha_{o,j} + \sum_{k=1}^{p} \alpha_{j,k} x_{t-i,k}), \text{ for j=1, 2,…,q} \qquad 1$$

These are then transformed by an activation function, which is almost always the logistic function, g(x) = $\frac{1}{1+e^{-x}}$, and then are sent onto each of the nodes in the next layer. At the next layer of the nodes, the outputs from the previous layer are again coined using node specific weights, transformed and sent onto the nodes in the next layer. The outputs from the final layer consist of one or more different weighted combinations of the output from the previous layers, which are then transformed by an output function. In fitting the model, the goal is to seek the values of the unknown parameters often called weights that make the model fit for the training data well.

For time series forecasting, the Neural Network model performs a nonlinear functional mapping from the past observation ($x_{t-1}, x_{t-2}, ..., x_{t-p}$) to the future values y_j, that is,

$$y_j = f(x_{t-1}, x_{t-2}, ..., x_{t-p}; \alpha_1, \alpha_2, ..., \alpha_q; \omega) + \varepsilon_t \qquad 2$$

Where $\alpha_1, \alpha_2, ..., \alpha_q; \omega$ are unknown parameter vector

$x_{t-1}, x_{t-2}, ..., x_{t-p}$, are known vector of past values of X_t

ε_t is the residual at time t

For a series with p input nodes and q hidden layers,

$$y_j = \alpha_o + \sum_{j=1}^{q} w_j (1 + \exp[-\alpha_{oj} - \sum_{k=1}^{p} \alpha_{jk} x_{t-i,k}])^{-1} + \varepsilon_t \qquad 3$$

The weights ω_j , α_{jk} and the bias α_o and α_{oj} are unknown and are usually selected to minimize the prediction residual error.

3. Methodology

Here, we consider apopular learning method capable of handling such large learning problems-the backpropagation algorithm. The back-propagation neural network was chosen for this research because it is capable of solving a wide variety of problems and it is commonly used in time series forecasting. The Backpropagation algorithm seeks to minimize the error term between the outputof the neural net and the actual desired output value. The error term is calculated bycomparing the net output to the desired output and is then fed back through the networkcausing the synaptic weights to be changed in an effort to minimize error. The process isrepeated until the error reaches a minimum value. This method is not only more general than the usual analyticalderivations, which handle only the case of special network topologies, butalso much easier to follow (Rojas, 1996).

The backpropagation algorithm looks for the minimum of the error functionin weight space using the method of gradient descentwhere error metrics are usually squared to obtain squared error or mean squared error. Hence, in this study, we used supervised learning paradigm using gradient descent to minimize the learning cost-mean squared error, which minimize the average squared error between the network's output and the target value over all the pairs. A unit in the output layer determines its activity by following a two (2) steps procedure. First, the total weighted input $X_{t-i,k}$ through:

$$X_j = \sum_{k-1}^{p} x_{t-i,k} \alpha_{jk} \qquad 4$$

Second, the unit calculates y_j using some function (sigmoid function) of the total weighted output x_j

$$y_j = \left[1 + e^{-x_j}\right]^{-1} \qquad 5$$

Once the activity of the output layer is determined, the networks compute the error

$$E = \frac{1}{2} \sum_{j=1}^{q} \left(y_j - \hat{y}_j\right)^2 \qquad 6$$

If there is an indication that forecast errors are larger, the estimates of the model's parameters will be adjusted in the following five (5) steps.

First, compute the rate of changes of error (E) as activity of an output unit is changed.

Second, compute the rate of changes of (E) as the total input received by an output unit is changed.

Third, compute the rate of changes of error (E) as a weight on the connection into an output unit is changed.

Fourth, compute the rate of changes of error (E) as activity of unit in the previous layer is changed. As the activity of the unit in the previous layer changes, it affects the activities of all the output units to which it is connected. Hence we add together all these separate effects on output units to compute overall effect on the error.

The last step performs the actual gradient descent by adjusting the weights.

The algorithm trains the neural network incrementally, meaning that each instance is considered independently, and after considering a training instance, the weights and biases are updated before considering another one. The update is done in a way that performs a gradient descent minimization of the error with respect to the weights and biases.

4. Data Analysis and Results

Although, there is no consensus on how to split the data for neural network applications, the general practice is to allocate more data for model building and selection. In this study however, the data has been apportioned into 2:1:1 ratio. We took the daily data from 18th April, 2007 to Sept. 3rd, 2012 as in-sample data set with 1162 observations for training, 23rd July, 2010 to 13thAugust, 2011 and 14th July, 2011 to 3rd September, 2012 making 387 observations each for validation and test sample respectively. To determine the NN architecture that best describe the series, the series were first subjected to Automatic Network Search (ANS). This is done by running the series without specifying the number of hidden layers; rather, we assigned a range of 2-12 hidden layers, without specific number of iterations, learning rates and momentum parameter and network randomization. The error function (sum of squares) was however indicated. The number of network to train in each case is 40 while the best is automatically retained for further training. Having obtained the network structure; the next thing is the training of the NN. At this stage, the number of iteration required, learning rate, momentum parameter and the network randomization were specified.

The result of the ANS is an NN 1-4-1. The performance of the models is shown in table 2on the basis of two performance metrics. Figure 1 and figure 2 are the NN residual plot for learning rates and momentum parameter (0.1 and 0.3) and (0.5 and 0.5) respectively.

Model	MSE	MAE
A	0.8090	0.6778
NN 1-4-1 B	0.7988	0.6533

Table 1: NGN/GHC test sample error performance.

5. Discussion

Based on our findings both the MSE and MAE for NN 1-4-1 is smaller when both the learning rate and momentum parameter are 0.5 each compared 0.1 and 0.3 for learning rate and momentum parameters respectively. This disparity lends supports to the findings of other researchers that increase in these parameters leads to faster convergence (Gould, 2004, Azad and Mashin, 2011, and Dunne, 2007). Figure 1 and 2are residual plots for model A and B respectively.

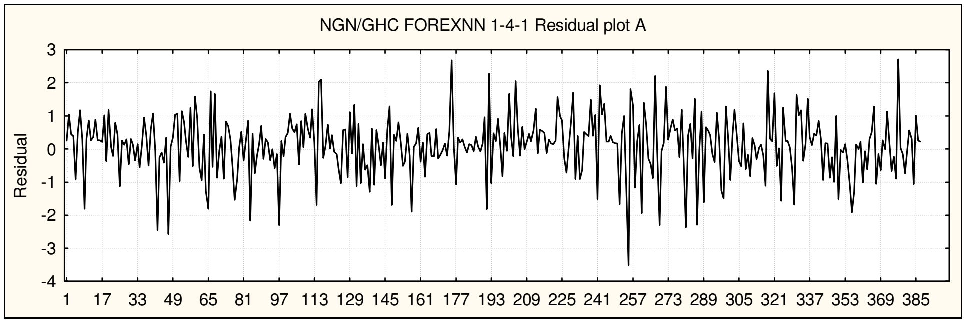

Figure 1: NGN/GHC NN Residual plots

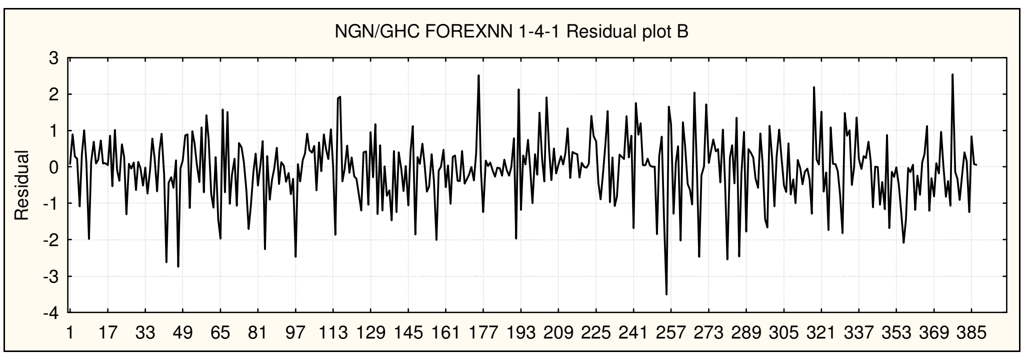

Figure 2: NGN/GHC NN Residual plots

6. References

i. Akinwale, A. T., Arogundade, O.T. and Adekoya A. F. (2009).“Translated Nigeria stock market prices using artificial neural network for effective prediction”.Journal of Theoretical and Applied Information Technology.

ii. Al-Zeaud, H. A. 2011. “Modelling and forecasting volatility using ARIMA model”. European Journal of Economics, Finance and Administrative Sciences ISSN 1450-2275 Issue 35

iii. Azad, A. K. and Mahsin, M. (2011).“Forecasting exchange rates of Bangladesh using ANN and ARIMA models: A comparative study”. International Journal of Advanced Engineering Sciences and Technology 10.1: 031 – 036.

iv. Gould, J. N. (2004). “Forex prediction using an artificial intelligence system”. Thesis.Science, graduate college.Oklahama State University. 1-58

v. Gujarati, D. N. (2004).Basic Econometrics.4thed. New York: McGraw-Hill.

vi. http://www.oanda.com/curency/historical-rates/

vii. Huang, W., Lai, K. K., Nakamori, Y.and Wang, S. (2004).“Forecasting foreign exchange rates with artificial neural networks: a review”. International Journal of information Technology and Decision Making 3.1:145 – 165.ISSN: 1927-5307

viii. Kamruzzaman, J and Sarker, R. A. (2004). “ANN-Based Forecasting of Foreign Currency Exchange Rates”. Neural Information Processing - Letters and Reviews Vol. 3, No. 2,

ix. Kamruzzaman, J. and Sarker, R. A. (2003).“Comparing ANN based models with ARIMA for prediction of forex rates”.ASOR Bulletin. 22.2:2-11

x. Kamruzzaman, J. and Sarker, R. A. and Ahmad, I. (2004). SVC based models for predicting foreign currency exchange rates. ICDM: 557 – 560.

xi. Kodogiannis, V. and Lolis, A. (2001).“Prediction of foreign exchange rates by neural network and fuzzy system based technique”. European Symposium on Artificial Neural Networks. ESANN 2001 proceedings.

xii. Kumar, M. (2010).“A time varying parameter vector autoregression model for forecasting emerging market exchange rates”. International Journal of Economic Sciences and Applied Research 3.2:21-39.

xiii. Kutner, M. H., Nachtsheim, C. J., Neer, J. and Li, W. (2005).Applied linear statistical models.5th ed. McGraw-Hill Ivwin.

xiv. Kutsurelis, J. E. (1998). Forecasting financial markets using neural networks: analysis of methods and accuracy. Thesis. Management, Naval Postgraduate School, Monteren, California: iv-66

xv. Lee, C. G., Steve, L. and Ah, C. T. (2001). “Noisy time series prediction using recurrent neural networks and grammatical inference”.Machine Learning 44.1:161-183.

xvi. Nag, A. K. and Mitra, D. (2002). “Forecasting daily exchange rates using genetically optimized neural networks”.Journal of Forecasting 21:501 – 511.wileyinterscience.

xvii. Philip, A. A., Taofiki, A.A. and Bidemi, A.A. (2011). “Artificial neural network model for forecasting foreign exchange rate”. World of Computer Science and Information Technology Journal (WCSIT) ISSN: 2221-0741 1.3:110-118

xviii. Rojas, R. (1996). Neural Networks. Springer-Verlag, Berlin.

xix. Tenti, P. (1996). Forecasting foreign exchange rate using recurrent neural networks, applied artificial intelligence, 10: 567 - 581. Taylor and Fancis.

xx. Yu, L., Wang, S. and Lai, K. K. (2005). Adaptive smoothing neural networks in foreign exchange rate forecast. Eds. V. S. Sunderam et al. LNCS 3516:523 – 530.