

A Neural Network Scheme for Monetary Policy Rate Validation in Nigeria

Oloruntoba Samuel Ogundele^{1,2}, Augustine Ujunwa^{2,3}, Aminu Ado Mohammed^{2,3}

¹Information Technology Unit, West African Monetary Institute, Accra, Ghana

²Monetary Policy Department, Central Bank of Nigeria, Abuja, Nigeria

³Financial Integration/Trade Integration Unit, West African Monetary Institute, Accra, Ghana

Email address:

oogundele@wami-ima.org (O. S. Ogundele), aujunwa@wami-ima.org (A. Ujunwa), amohammed@wami-ima.org (A. A. Mohammed)

To cite this article:

Oloruntoba Samuel Ogundele, Augustine Ujunwa, Aminu Ado Mohammed. A Neural Network Scheme for Monetary Policy Rate Validation in Nigeria. *American Journal of Artificial Intelligence*. Vol. 4, No. 2, 2020, pp. 62-72. doi: 10.11648/j.ajai.20200402.13

Received: October 4, 2020; **Accepted:** November 12, 2020; **Published:** December 11, 2020

Abstract: This research work is an exploratory study that tried to examine the viability of adopting artificial neural network (ANN), an aspect of machine learning in the analysis of monetary data for the design and validation of monetary policy from both optimistic and normative approach. Methodologically, the research is motivated by the work of [33] which used the Greenbook real time data of the U.S. Federal Reserve's in the analysis of monetary policy reaction functions in forecasting performance using ANN. Following the work on the adoption of this technique, we tried to develop a framework based on machine learning for policy rate forecasting by analysing macroeconomic data with the aim of guiding and aiding monetary authority in making monetary policy decisions. From the results, the ANN perform better in predicting the monetary policy rate compared to the linear models and the univariate process. It also revealed the non-linearity in the behavior of the monetary policy rate in Nigeria during the study period. While the work does not mean to advocate that machine will replace human-being in policy rate determination in the monetary policy-making process; we believe that the development and implementation of this system would support building effective prediction system which can be validated. The result from the designed system is expected to enhance credibility, confidence and transparency of central banks in making an independent decision (s) based on objective forecasts and implied analysis in setting policy through a well-structured comparison of results.

Keywords: Monetary Policy, Artificial Neural Network, Taylor Rule, Data

1. Introduction

Monetary policy has been defined as deliberate actions of the monetary authorities to influence the direction and level of economic activities (the amount of production, buying and selling in the economy) [4, 38]. This is achieved through changes to the amount of money in circulation (money supply) by tinkering with the availability of money/credit, the quantity as well as cost in order to achieve the desired macroeconomic objectives of the country and favourable external balance positions. The actions involved the manipulations of the monetary base usually, the money supply, by setting monetary rates with the intention of controlling the quantity of money in the economy. The significance of money in economic life has made policy makers attach importance to the conduct of monetary policy especially in order to achieve macroeconomic

stability. The principal objective of the monetary authority in formulating and implementing monetary policy is to achieve and maintain price stability as well as nurturing and maintaining a stable financial system [38]. In addition to this objective, central banks also promote economic growth and stable employment.

From inception, economics models are usually adopted and formulated to represent the behaviour of humans and how its shape their actions and decisions in the management of resources. These models usually come with some deep cognitive and emotional reasoning which are imprecise and sometimes vague. In order to conduct monetary policy, policy makers monitored and analysed huge volumes of economic data from dissimilar sources and the data normally come with varying range of frequencies and levels of aggregation that are overwhelming. The outcome of the analysis is interpreted to assist in policy formulation aided by inference drawn from

economic theories as postulated in literature and subsequently implemented [7]. For many years, and influenced by the monetarist perspectives and macroeconomic management led by Friedman [3], the objective of monetary policy by central banks have been how to achieve low inflation which usually is a measure of the effectiveness of monetary policies transmission in most economies.

To achieve the objective, usually high volume of data are needed to be analysed to guide decision making with regards to monetary policy formulation. Generally, large volumes of economic dataset are generated daily based on the various business activities that are going on in the economy which are due to various business processes such as finance, accounting, commerce manufacturing, among others [2, 34, 37].

This data comes in various format, that could be seasonal, revised or non-revised, forecast and sometimes with a lag. The cost of analysing this volume of data series by central banks is usually overwhelming and indeed indicates that policy-makers see these activities as critical to their decisions making [1]. Over the years, econometric methodologies have been employed in the analysis of these large volumes of data series and this has significantly improved forecasts, interpretation as well as revealing the behaviour of key macroeconomic variables [10-12]. Notwithstanding the success of the econometric techniques; it has been observed that most empirical analyses of monetary policy were restricted to the frameworks adopted by the central banks [1, 17].

This implies that the analysis will be premised on the exploitation of only a limited amount of data notwithstanding the huge volume that may have been gathered at the onset. For instance, the adoption of Vector Auto regression (VAR) Model in carrying out estimation in an effort to depict the determinants and effects of monetary policy generally comes with some drawbacks. These include the limitation of the analysis to about only eight macroeconomic time series, the atheoretical nature of VAR and the need for the modeler to have a sound understanding of the economic theories. The implication is that the large cost incurred by central banks in gathering and analysing large volume and wide range of data implies waste which amount to inefficiency since only small amount of the data are only used in the long run. Thus, it is inevitable to design a system that can intelligently analyse the whole range and volume of data and generate accurate result.

Another issue is the major divide between the formal models adopted by central banks in analysing the vast volume of data gathered and the real practise of monetary policy by central bank [1]. The formal models are abstract scientific models that involve a lot of academic rigours where solutions are derived by manipulation of various mathematical formulae. The adoption of the formal model also limits the usage of the voluminous data sets that have been gathered in improving forecasting. Therefore, outcome of policy decision maybe vague and subjective which could mislead or mis-represent policy decision. The central bank's practice entails decision making based on personal judgemental view acquired over the years with diverse experience and the combination of macroeconomic,

statistical and heuristic models based on historical trends. This usually comes at a great cost because the formal models usually ignore several key information, given its abstract nature, thus, making the evaluation less informative and inaccurate thereby distorting policy outcome. The complexity thus resulted in the problems of capturing the central banker's approach to data analysis, due to the mixes in the use of large macro-econometric models with other statistical simulations (such as VARs), exploratory and heuristic judgmental analyses, and informal computation of information from varying sources. This may make central banks policies less accurate and informative than it would have ought to be. Central banks are also at risk of losing institutional knowledge and experience, as well as historical facts that would otherwise enrich monetary policy making.

Collecting, integrating and analysing data can be a complex task because the data resides in many internal and external locations and its level of quality may be unknown. Authors in [35] argued that the creation and aggregation of data are approaching the zettabyte/year range within a few years. Another challenge to data usage is the storage and movement data that require research and new paradigms. As highlighted in [6], with the advent of increasing economic activities, large volume of macroeconomic data are created and generated. Formulation of monetary policy and determination of interest rate is a difficult task because of the complex characteristics of data that must be analysed. These characteristics include high volatility (such as exchange rate), inherent noise (stock price), hidden relationship (inflation with varying components) and dependency on other parameters (interest rate and investment), such as employment data, money aggregates, aggregate demand *e.t.c* [1]. However, different researches have been carried out in other fields to solve the difficulties in analysing complex data where ANN has been adjudged to be more efficient than other intelligent forecasting systems.

2. Literature Review

Artificial Neural Networks (ANN) and Machine Learning (ML) are an aspect of cognitive science that evolved from two concepts of pattern recognition and computational learning which are components of Artificial Intelligence (AI) [5, 14, 18, 20]. ML make predictions for the future based on the past information usually through analysis of algorithms that would have been trained. ANN, as a learning process is built on human biological neural networks with some statistical models and its estimate values based on a huge number of inputs. Neurons are interconnects with numeric values and adjusted based on experience. This will allow the inputs to be used in the learning process. ML and data mining processes imitate mathematical optimization procedures to build complex models, where designing and programming explicit and rule-based algorithms are infeasible. The ML process can be supervised or unsupervised depending on the algorithms. ANN is one of the popular supervised learning process methods [18, 20, 21] which can be likened to the

human brain in its process information and can be employed to determine the complex relationship between inputs and outputs of processes. A successfully trained ANN system can predict the output on a set of previously unseen inputs. Several ANN algorithms have been proposed in the literature. Backpropagation (BP) algorithm will be employed in this study [5, 14, 15].

Central bankers' reputations as data fiends is seen in their quest to minimize average forecast errors, and by making and shifting policy objectives, given the uncertainty about the correct model of the economy, and the central bank's political need to demonstrate that it is taking all potentially relevant factors into account. In order to formulate monetary policy, central banks extract large volume of macro-economic data gathered from the both secondary and primary sources in the economy. The practice is the usage of revised rather than real-time data based on empiric which is often not innocuous [1]. For example, [23, 24] has shown that the description of the historical conduct of monetary policy provided by a standard Taylor rule is much less convincing when estimated using real-time data. Usually, these data that are analysed include series on GDP and its components, aggregate price measures, and monetary aggregates and components. In addition, variety of financial indicators (such as stock price indices, interest rates, and exchange rates), as well as developments in international economic; are included in the analysis. It is pertinent to mention that the robustness and the completeness of the databases depend on their breadth of coverage. Unfortunately, most of the data set is necessarily somewhat limited both in the number and scope of the time series included especially given the lag in gathering macroeconomic data which sometimes make the forecasts constructed with this data set to be poor. Thereafter, the evaluated policy reaction functions would be taken as inputs to analysed the target variables (such as inflation and employment figure) to produce the implied but desire policy settings as outputs [33].

The idea of a rules-based monetary policy dates to the work of [8, 9]. A reaction function was proposed where the nominal interest rate depends linearly on the gaps between actual and targeted values of inflation and output. This simple Taylor rule was shown to match Federal Reserve actual interest setting behaviour between 1987 and 1992 very well. Taylor rules explain how central banks set their interest rates in response to inflation and macroeconomic developments [7, 8]. Several researches have adopted the improved versions of ANN in solving prediction problems and these have shown better efficiency in the results. The principled of a rule-based monetary policy was also introduced in [8, 9] with a reaction function where the nominal interest rate is a function of linear gaps between actual and targeted values of inflation and output. He adopted the use of Support Vector Machine (SVM) in Discretion versus policy rules in implementation.

A linear reaction function based on theoretical model was provided by [26] as a solution to the optimization problem of reducing the deviations of output and inflation from the desirable values by central bankers. The result from the

model, however, was affected by the assumption of a quadratic loss function indicating symmetric preferences and a Philips curve that is linear.

The extension of nonlinear model by ANN became imperative given that ANN has the property of being a universal approximator by fitting in-sample data to any degree. Thus, [27, 28] successfully applied ANNs in a macroeconomic time-series forecasting context. The application serves as a valuable tool for time series forecasting notwithstanding the fact that a structural interpretation of the estimated parameters was not provided.

The context of USA monetary policy reaction functions using ANN was considered by [25]. Although, the simulation of the time series forecast seems unrealistic in context given the fact that the periods considered were randomly drawn from subsamples. However, the work revealed that the ANN outperforms a linear Taylor rule and a random walk. This was particularly valid when the data is analyses based on the current value of the federal funds rate, and otherwise when there is time lagged in the value.

Researchers in [33] uses the Greenbook real time data of the U.S. Federal Reserve's in the analysis of monetary policy reaction functions forecasting performance using ANN. From the results, the ANN perform better in predicting the nominal interest rate compared to the linear and nonlinear Taylor rule models and the univariate processes. It also revealed the non-linearity in the behavior of the monetary policy in the USA during the study period. The work of [33] actually inspired the motive behind this work in other to check the behaviour of monetary policy in Nigeria and to validate the decisions of the policy makers over the years.

3. Taylor Rule and Monetary Policy Model in Nigeria

[23] Has emphasized the importance of using real-time data in a Taylor rule framework since inflation and output gap ex post measures might be different due to revision processes yielding misleading reaction functions. [24] Further gave a review of Taylor rules' characteristics relative to alternative monetary policy guideline. The basic structure of the Taylor rule is as follows [8]:

$$i_t = r^n + \pi_t^* + 1.5 (\pi_t - \pi_t^*) + 0.5 \tilde{y}_t \quad (1)$$

where i_t is the central bank's key monetary policy instrument, expressed as a function of r^n the natural rate of interest, π_t^* the inflation target, π_t the inflation rate, and \tilde{y}_t the output gap. In [8], the above parameters were shown to hold for the case of the USA. Numerous studies have ever since tested the simple model described above under different specifications, samples and countries, and econometric techniques.

Taylor rule has been augmented in [30, 31] by the inclusion of interest rate smoothing and exchange rates as depicted in equations (2) and (3) to include additional factors affecting interest rates, as follows:

$$i_t = (1 - \rho)i_t^* + \rho i_{t-1} + \varepsilon_t \quad (2)$$

$$i_t^* = r^n + \beta(\pi_{t-1} - \pi_t^*) + \gamma \tilde{y}_{t-1} + \eta \Delta X_t \quad (3)$$

where i_t^* is the target interest rate, i_{t-1} is lagged interest rate to capture smoothing effects, and ΔX_t is the change in the real exchange rate. We use the interbank rate as the CBN's key rate and year-on-year (y-o-y) changes in the CPI to

$$i_t = (1 - \rho)\alpha + (1 - \rho)\beta\pi_{t-k} + (1 - \rho)\gamma\tilde{y}_{t+k} + (1 - \rho)\eta\Delta X_t + \rho i_{t-1} + \varepsilon_t \quad (4)$$

Usually, simple Ordinary Least Square (OLS) is used to examine the determinants of interest rate changes while the equation above is expressed in the regression form.

In a broader form, the monetary policy formulation in Nigeria as an extension of the augmented Taylor rule could be captured using algebra notations (with some assumptions) to depict the Nigeria monetary Policy process.

We assume the economy would always be at a Steady State S defined as

$$S = \{S_{i+1}, S_{i+2}, S_{i+3}, \dots, S_{i+n}\} \quad (5)$$

Such that

$$S(i) = \text{present state of the economy} \quad (6)$$

and could undergo transformation based on the prevailing circumstances that will affect the macroeconomic variables such that

$$S(n_i) = \text{the next state in that sequence} \quad (7)$$

$$\forall 0 \leq i \leq n$$

This implies that at every point in time, the economy will be at a state where data on macroeconomic aggregates would be collected to enable policy makers analyse before making informed policy decisions.

Thus,

We further define other Macroeconomic aggregates (M_a) as a set

$$M_a(t) = \begin{cases} Ma_1 = y_1, \\ Ma_2 = y_2, \\ Ma_3 = y_3, \\ \vdots \\ Ma_n = y_n \end{cases} \quad (8)$$

Where y is the individual variable

It is important to state here that monetary policy formulation in Nigeria considers other macroeconomic aggregates other than Inflation, GDP and Exchange Rate. These include Money Supply, Foreign Reserves position, Stock Exchange variables, Crude Oil price, 90 days Treasury bill rates, e.t.c).

This implies that the summated other macroeconomic aggregates can be captured as the computed union of all the variables given as follow:

calculate the inflation rate. The ρ coefficient in the above equation reflects the interest rate smoothing parameter [29]. Putting equation (3) into (2), we adopt the following specification for the augmented Taylor rule.

$$M_a(t) = \bigcup_{y=1}^n (M_a(n)) \quad (9)$$

We also define the economic threshold for each aggregate as

$$\beta(t)_n = \{\beta_1, \beta_2, \beta_3 \dots \dots \dots \beta_n\} \quad (10)$$

Such that $\{\beta_{min} \leq M_a \leq \beta_{max} \mid \beta_{max} \leq M_a \leq \beta_{min}\}$

And monetary actions as $\alpha = \{\alpha_1, \alpha_2, \alpha_3 \dots \dots \dots \alpha_n\}$

$$\therefore \alpha(n) = \sum \{y_n \Rightarrow M_a(t) \forall y = (1..n) \mid \beta(t)_n \in \beta\} \quad (11)$$

Where $\alpha(n) \in \alpha, n = 1$

Therefore, the set of all possible monetary policy actions that can be taken at a particular state based on the on the summated macroeconomic aggregates and given their various threshold is captured below.

$$\alpha S = \{(S, \alpha) \forall s_i \in S \text{ and } \alpha_i \in \alpha(n)\}; \forall 0 \leq i \leq n \quad (12)$$

Since, the focus of this paper is on policy rate, then,

$$i_t = f(\alpha S) + (1 - \rho)\alpha + (1 - \rho)\beta\pi_{t-k} + (1 - \rho)\gamma\tilde{y}_{t+k} + (1 - \rho)\eta\Delta X_t + \rho i_{t-1} + \varepsilon_t \quad (13)$$

4. Network Architecture

In developing the neural network system, the Multi-Layer Feed Forward (MLFF) neural network model is adopted. The feed-forward neural network is a directed acyclic graph, $G = (V, E)$, and a weight function over the edges, $w: E \rightarrow \mathbb{R}$. The nodes of the graph correspond to neurons. Each single neurons is a scalar function, $\sigma: \mathbb{R} \rightarrow \mathbb{R}$. The neurons are organized in the form of layers which are the set of nodes decomposed into union of disjoint subsets $V = \bigcup_{t=0}^T V_t$ such that every edge E connects some node in V_{t-1} to some node in V_t , for some $t \in [T]$. Thus, it consists of the following layers:

1. The input layer V_0 having N input nodes. It contains $n + 1$ neurons, where n is the dimensionality of the input space.
2. The hidden layers V_1, \dots, V_{t-1} also having N nodes in each layer
3. The output layer V_T with a single neuron which is the output variable. T is the number of layers in the network which does not include V_0 .

The hidden layer of $n_1^{(1)}, n_2^{(2)}, n_3^{(3)}, n_4^{(4)} \dots \dots \dots n_n^{(n)}$ are intermediate variables which interact by means of weight matrices $W_{i,j}^{(o)}$ and $W_{i,j}^{(1)}$ with adjustable weights. Each input is fed in through the input layer nodes as shown below.

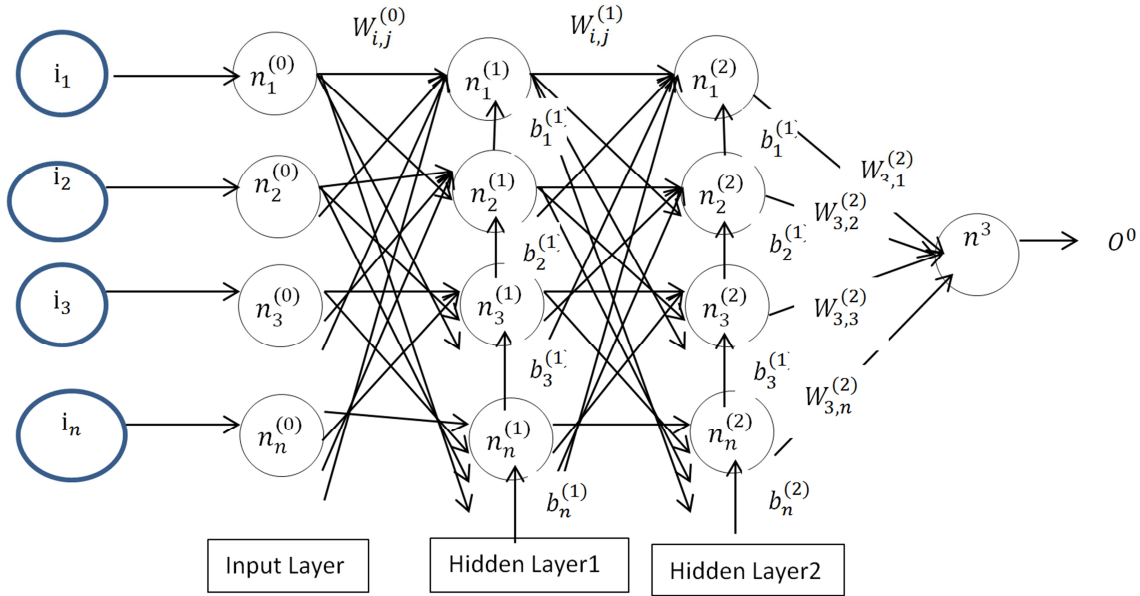


Figure 1. Architecture of the Neural Network design for Policy rate Prediction.

5. Network Training

In the model input layer units distribute input signals to the network. Connection weights modify the signals that pass through it. Hidden layers and output layer contain a vector of processing elements with an activation function. Usually, the Sigmoid function is used as the activation function. [32]. Every unit $n^{(1)}$ computes its new activation $n^{(2)}$ as a function of the weighted sum of the inputs to unit $n^{(2)}$ from directly connected cells. The input data to the network comes through the input units. The network consists of an input layer with n neurons. The connections coming out of an input unit have weights $W_{i,j}^{(0)}$ associated with them.

Each hidden node calculates the weighted sum of its inputs and applies a threshold function to determine the output of the hidden node. The weighted sum of the inputs for the hidden node $n^{(1)}$ is calculated as:

$$S_i^{(1)} = \sum_{j=0}^n (n_i^{(0)} \times W_{i,j}^{(0)}) + b_i^{(1)} \quad (i = 1, \dots, n; j = 1, \dots, n) \quad (14)$$

At the hidden node, transfer functions calculated a layer's output from its net input and passed through a nonlinear mathematical function called a sigmoid transfer function

$a = \text{tansig}(n) = \frac{2}{(1 + e^{-2n})}$ where n is the input node and a is the node output, is applied to the weighted sum of the inputs to the hidden node to obtain the output of hidden node $n_i^{(1)}$ given by:

$$n_i^{(1)} = f(S_i^{(1)}) \quad (i = 1, \dots, n; j = 1, \dots, n) \quad (15)$$

The process is repeated for the second hidden layer to obtain its output $n_i^{(2)}$

A similar computation will be done for the output nodes $S_i^{(3)}$. Finally, a set of outputs is produced for the network. During the forward pass the synaptic weights of the network are all fixed. The forward phase finishes with the computation of an error signal

$$e_i = d_i - y_i \quad (16)$$

where d_i is the desired response and y_i is the actual output produced by the network in response to the input x_i [35].

The algorithm involves the weight update and neuron error gradient. The weights in the neural network are updated to give the desired output. This forms the basis of training the neural network. The back-propagation is used for the weight updates. The input is fed in, the errors are calculated and sent through the network making changes to the weights to minimise the error [36].

The changes to the weight are calculated by using the gradient descent method. The weight update is determined as follows

$$W_{i,j} = W_{i,j} + \Delta W_{i,j} \quad (17)$$

$$W_{j,k} = W_{j,k} + \Delta W_{j,k} \quad (18)$$

Where $\Delta W_{i,j}(t) = \alpha \cdot \text{InputNeuron}_j \cdot \Delta \delta_j$ and $\Delta W_{j,k}(t) = \alpha \cdot \text{InputNeuron}_j \cdot \Delta \delta_k$

$\alpha = \text{learning rate which usually is between 0 and 1}$

Error Gradient $\delta_k = y_k (1 - y_k)(d_k - y_k)$

Where y_k is the value at output neuron K and d_k is the desired value at output neuron K . This is a difference between the error gradients at the output and hidden layers. The hidden layers' error gradient is based on the output layers' error gradient (back propagation) so for the hidden layer the error gradient for each hidden neuron is the gradient of the activation function multiplied by the weighted sum of the errors at the output layer originating from that neuron captured as follow

$$\delta_j = y_j (1 - y_j) \sum_{k=1}^n W_{j,k} \delta_k$$

It is pertinent to mention that [22] highlighted the universal approximator property of ANN models where any unknown function H (under mild regularity assumptions) can

be approximately arbitrarily close by a linear combination of activation functions G , i.e

$\left| H \left(z_t - \sum_{j=1}^q \gamma_j G(\beta_j' z_t) \right) \right| < \delta$ with finite q and $z_t \delta \in \mathbb{R}_{>0}$. Therefore, the fact that the specification is data driven implies that there is no need to specify a specific functional form which is a great advantage of ANN.

6. Model Simulation and Experimental Setup

The ANN will be carried out in the following stages: Data Collection, Data Pre-processing, Network Creation, Network Training and Network Validation and result generation. The source of the data used for the research are macroeconomic variables downloaded from the Statistical Database of the central bank of Nigeria available at <http://statistics.cbn.gov.ng/cbn-onlinestats/>.

These data are analysed in preparation of each meeting of the Monetary Policy Committee. The use of these real-time data sets ensures that only information is used in the forecasts that was actually available by the central bank at the time they set the interest rate. It circumvents the potential problem of estimating misleading reaction functions due to the use of revised data as pointed out in [23]. The study employed monthly data from January 2012 – December 2019 and the ending period is due to fact that the data is published only after a lag. Estimation is done in MATLAB_R2020a by the Levenberg-Marquardt algorithm (LMA) based on similar model selection strategy adopted by [33]. The model is a non-linear least squares solver that combines the Gauss-Newton algorithm with the gradient descent method, together with an early stopping procedure that ensures that training stops if the network performance fails to improve or remains the same for n consecutive epochs. The variables considered for inputs include: *inflation, gross domestic product, Money supply, Non-Performing loan, Reserve position, Exchange rate, capital flows, oil price, All share index, values of share, market capitalization, debt, value of Federal Allocated Account shared, stock turnover while monetary policy rate is the output variable.*

It is pertinent to mention that [22] highlighted the universal approximator property of ANN models where any unknown function H (under mild regularity assumptions) can be approximately arbitrarily close by a linear combination of activation functions G , i.e.

$\left| H \left(z_t - \sum_{j=1}^q \gamma_j G(\beta_j' z_t) \right) \right| < \delta$ with finite q and $z_t \delta \in \mathbb{R}_{>0}$. Therefore, the fact that the specification is data driven

implies that there is no need to specify a specific functional form which is a great advantage of ANN.

7. Forecasting Models

For the forecasting, four different types of models were used for the experiment: Non-linear models and one regression model that was used to test and compare the performance efficiency of the ANN models. The Non-Linear Models adopted are stated below while the performance of the test statistics are shown in Table 1.

Non Linear Autoregressive with External (Exogenous) Input (NARX)

$$y(t) = f(x(t-1)), \dots, x(t-d), y(t-1), \dots, y(t-d) \quad (19)$$

This predict series $y(t)$ given d past values of $y(t)$ and another series $x(t)$

Non Linear Autoregressive (NAR)

$$y(t) = f(y(t-1)), \dots, y(t-d) \quad (20)$$

This predict series $y(t)$ given d past values of $y(t)$

Non Linear Input-Output (NIO)

$$y(t) = f(x(t-1)), \dots, x(t-d) \quad (21)$$

This predict series $y(t)$ given d past values of $x(t)$

8. Results

The data was partition into three kind of Target Timesteps (70%, 15%, and 15%). The Training set are presented to the network during training, and the network is adjusted according to its error. (70%, 84 Target Timesteps). The validation set are used to measure network generalization and help the network to determine endogenously the mean squared error (MSE) in the set and the number of hidden neurons with their initial weights. It also set the criteria to stop the estimation by the Levenberg-Marquardt algorithm (LMA) once the optimal point is attained and generalization stops improving. (15%, 18 Target Timesteps). Testing set have no effect on training and so provide an independent measure of network performance during and after training. (15%, 18 Target Timesteps).

It is important to state that there is no specific rule to follow in choosing the data splitting percentages. Table 1 gives an overview of the MSE outputs of the network from the different models.

Table 1. The Forecasting Models and their Statistics.

TargetTimesteps	Training		Validation		Testing	
Target Values	84		18		18	
Statistics	MSE	R	MSE	R	MSE	R
NARX	1.96828e-3	9.99889e-1	2.30216e-2	9.98204e-1	9.54704e-1	9.33413e-1
NAR	1.13904e-1	9.89879e-1	1.12748e-1	9.99133e-1	1.63124e-1	9.85646e-1
NARX	5.84628e-1	9.99961e-1	1.25631e-1	9.91327e-1	5.04940e-1	9.54167e-1

The results from the ANN based on the different models are shown in figures 2-5. The output from NARX performs best out of the three models. The MSE is relatively low compared to the output from NAR and NIO. Overall, the results reveal that ANN, irrespective of the non-linear model adopted can predict the monetary policy rate in Nigeria with very high accuracy.

The findings are in line with the work of [33] where ANN was shown to perform better in predicting the nominal

interest rate compared to the linear and nonlinear Taylor rule models and the univariate processes. Further to this, the outputs from our ANN models also validate the non-linear structure of monetary policy rate in Nigeria over the period of study which is in agreement with the findings of [25] where ANN in the context of monetary policy reaction functions outperforms a linear Taylor rule and a random walk using federal funds rate as well as with the work of [33].

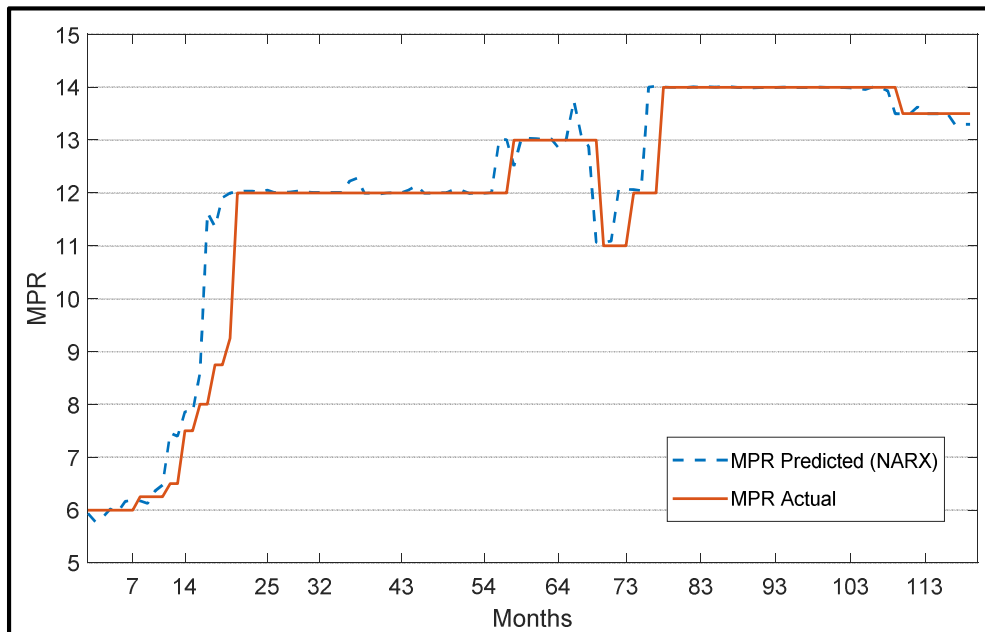


Figure 2. Actual and Predicted Forecast with Non-Linear Autoregressive with External (Exogenous) Input (NARX).

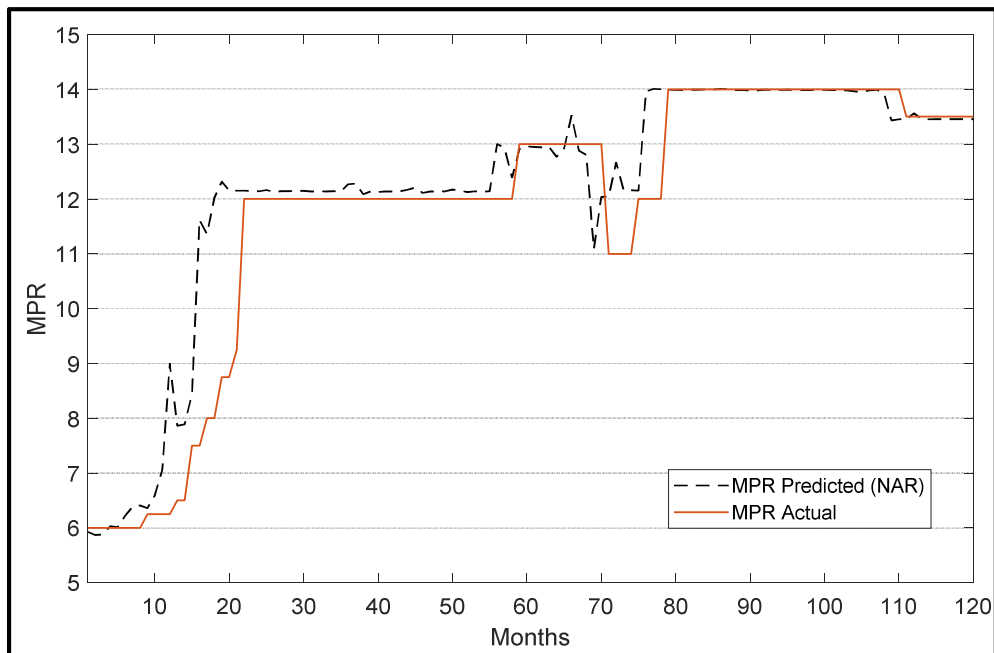


Figure 3. Actual and Predicted Forecast with Non-Linear Autoregressive (NAR).

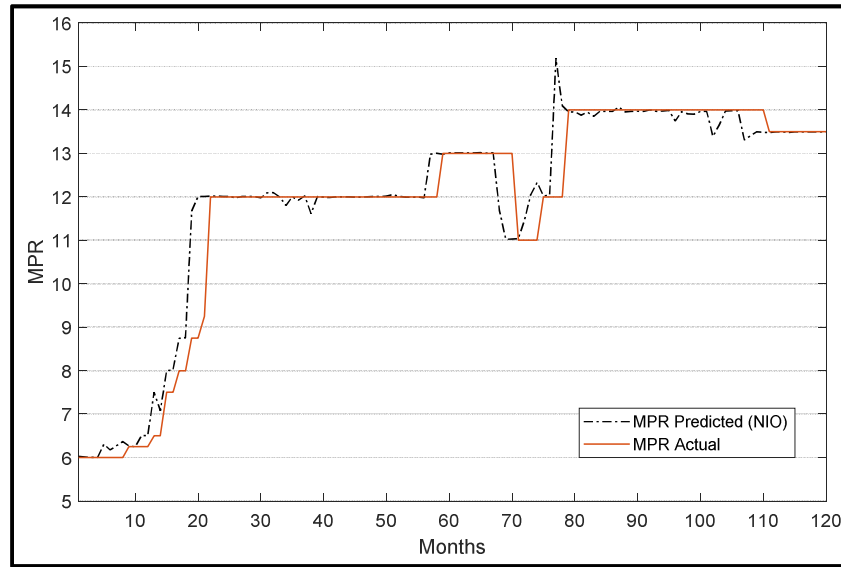


Figure 4. Actual and Predicted Forecast with Non-Linear Input-Output.

A similar analysis was undertaken by performing a linear regression analysis to predict MPR using the same data. This was done to further evaluate the performance of the ANN model. The regression is of the form:

$$Y_i = \beta_0 + \beta_i X_i + \varepsilon_i \quad (22)$$

The result as shown in Figure 5 compared to the ANN results in Figures 2-4 indicated that the ANN for monetary policy rate prediction performed better and produced improved forecasts than the linear regression model which is linear and thus, any univariate processes.

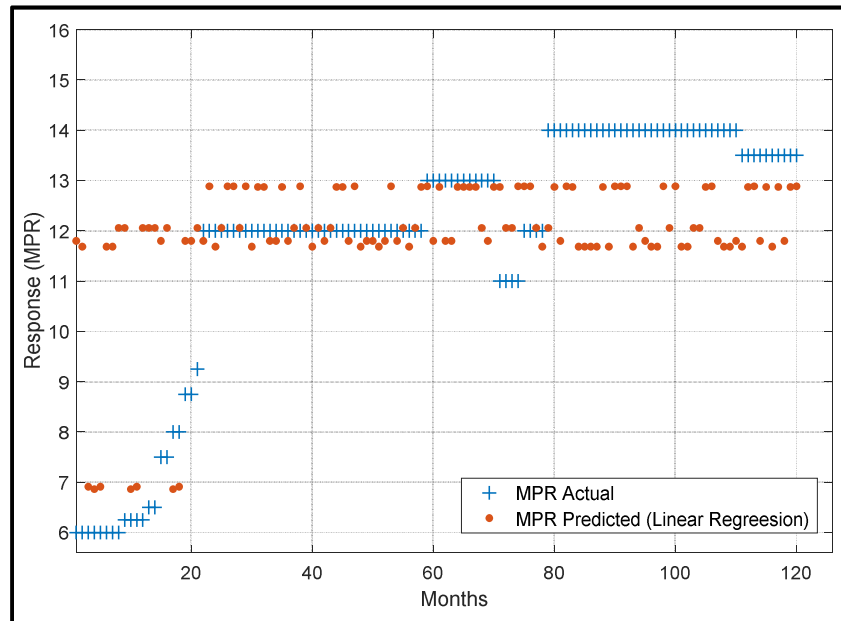


Figure 5. Actual and Predicted Forecast using linear regression analysis.

9. Conclusion

This work has shown artificial neural networks as a potential forecasting and validating tool for monetary policy formulation in Nigeria. This work does not mean to advocate that machine will replace human-being in the monetary policy-making process; but the development and implementation of this system would support the building of effective validation system which would have several advantages. The result from

the designed system would enhance integrity, confidence and transparency of central banks in making an independent decision (s) based on objective forecasts and implied analysis in setting policy though a well-structured comparison of results. Market watchers and public analysts, as well as research institutes, could also use the output of the system to make an informed judgemental comparison which could also form the basis of determining the credibility and transparency of the central bank of Nigeria in setting monetary policy rate. However, the model being only locally adapted and the lack of

real life economic interpretation of the variables may be a downside to the work. The real task of predicting and validating monetary policy in real life may requires deeper structural interpretations of the Nigeria economic behaviour which might involve human emotions and some qualitative social science behaviours. In view of this, it is pertinent to state

that a practical machine learning system would require a deeper and intensive design environment especially to implement the framework demonstrated in this paper. The need to build more intelligence into the system by incorporating human emotional filter based on appropriate algorithm would further enhance the research work.

Appendix

Appendix I: Predicted data and Target response with the difference (Error) for Non Linear Autoregressive with External (Exogenous) Input

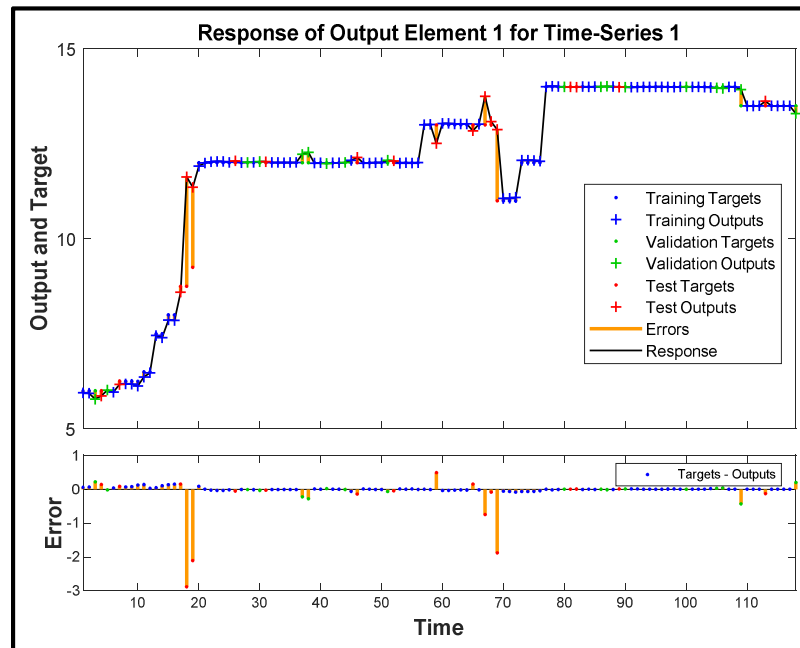


Figure 6. Output and Target's Response (NARX).

Appendix II: Predicted data and Target response with the difference (Error) for Non Linear Autoregressive Input

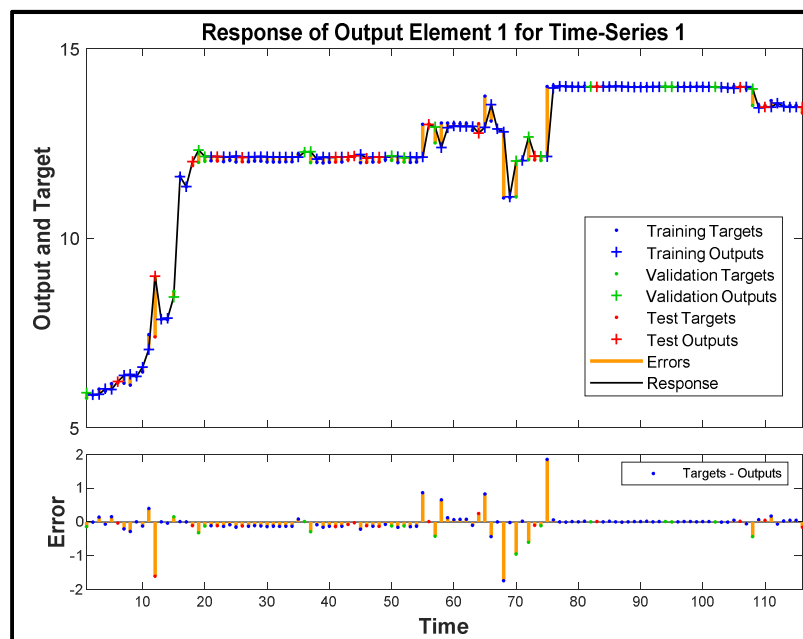
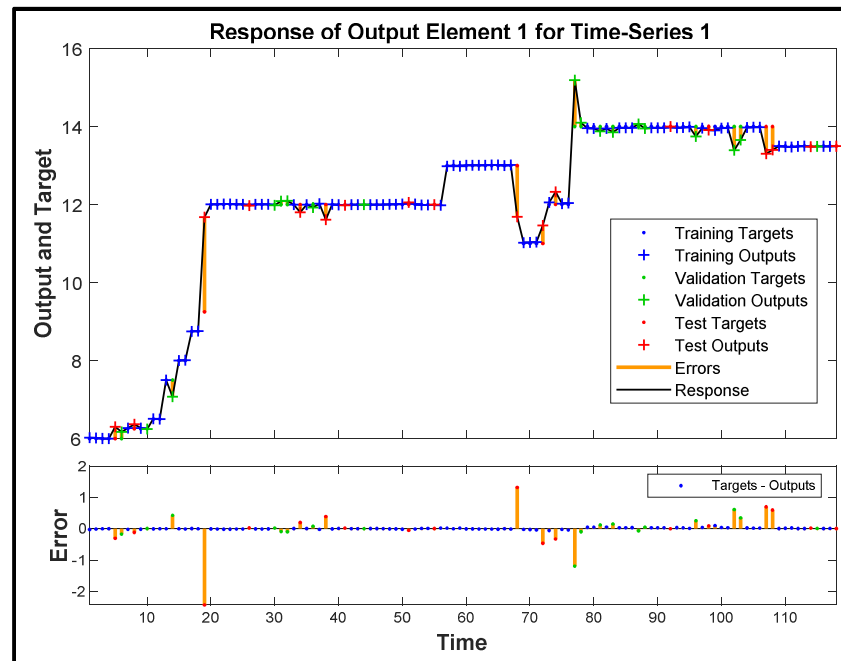


Figure 7. Output and Target's Response (NAR).

Appendix III: Predicted data and Target response with the difference (Error) for Non Linear Input-Output**Figure 8.** Output and Target's Response (Non-Linear Input-Output).**References**

- [1] Ben S. Bernanke, Jean Boivin, "Monetary policy in a data-rich environment", *Journal of Monetary Economics* 50 (2003) 525–546, doi: 10.1016/S0304-3932(03)00024-2.
- [2] Attention, shoppers: Store is tracking your cell, *New York Times*. URL: <http://www.nytimes.com/2013/07/15/business/attention-shopper-storesare-tracking-your-cell.html>.
- [3] Friedman M Benjamin, *A Program for Monetary Stability*. 1959. New York: Macmillan.
- [4] Friedman M Benjamin, *Monetary Policy, Fiscal Policy, and the Efficiency of Our Financial System: Lessons from the Financial Crisis*, *International Journal of Central Banking*, Vol. 8 No. S1, 2012, pp 301-309, <https://www.ijcb.org/journal/ijcb12q0a20.pdf>.
- [5] M. Naimur Rahman, et al, *Machine Learning with Big Data An Efficient Electricity Generation Forecasting System*, *Big Data Research*, (2016), <http://dx.doi.org/10.1016/j.bdr.2016.02.002>.
- [6] P. S. Yu, On mining big data, in: J. Wang, H. Xiong, Y. Ishikawa, J. Xu, J. Zhou (Eds.), *Web-Age Information Management*, in: *Lecture Notes in Computer Science*, vol. 7923, Springer-Verlag, Berlin, Heidelberg, 2013, p. XIV.
- [7] Taylor, J. B. 1979. "Estimation and control of a macroeconomic model with rational expectations." *Econometrica*, 47 (5): 1267–86.
- [8] Taylor, J. B. 1993. "Discretion versus policy rules in practice". *Carnegie Rochester Series on Public Policy*, 39: 195–214, North Holland.
- [9] Taylor, J. B. 1999. "A historical analysis of monetary policy rules". In J. Taylor, ed., *Monetary Policy Rules*, *Business Cycle Series*, 31: 319–348. University of Chicago Press, Chicago.
- [10] Stock, J. H., Watson, M. W., 1989. *New indexes of coincident and leading economic indicators*. NBER Macroeconomics Annual 4. MIT Press, Cambridge, MA.
- [11] Stock, J. H., Watson, M. W., 1999. *Forecasting Inflation*. *Journal of Monetary Economics* 44 (2), 293–335.
- [12] Stock, J. H., Watson, M. W., 2002. *Macroeconomic forecasting using diffusion indexes*. *Journal of Business & Economic Statistics* 20 (2), 147–162.
- [13] P. F. Pai, W. C. Hong, *Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms*, *Electr. Power Syst. Res.* 74 (3) (2005) 417–425.
- [14] P. C. Chang, Y. W. Wang, C. H. Liu, *Fuzzy Delhi and Backpropagation model for sales forecasting in PCB industry*, *Expert Syst. Appl.* 30 (4) (2006) 715–726.
- [15] Wikipedia contributor (2014, Jan 10), *Backpropagation* (version ID: 634418284) [online], available: <http://en.wikipedia.org/wiki/Backpropagation>.
- [16] Wikipedia contributor (2014, April 20), *Artificial Neural Network* (version ID: 174595685) [online], available: http://en.wikipedia.org/wiki/Artificial_neural_network.
- [17] Christiano, Lawrence J. & Eichenbaum, Martin & Evans, Charles L., 1999. "Monetary policy shocks: What have we learned and to what end?," *Handbook of Macroeconomics*, in: J. B. Taylor & M. Woodford (ed.), *Handbook of Macroeconomics*, edition 1, volume 1, chapter 2, pages 65-148, Elsevier.
- [18] Wikipedia contributor (2013, Nov 23), *Machine Learning* (version ID: 175534373) [online], available: http://en.wikipedia.org/wiki/Machine_learning, 2013, Web-12 Jan, 2014.

- [19] Wikipedia contributor (2016, Jan 16) [online], Cognitive Science (version ID: 699869928) [online], available: https://en.wikipedia.org/wiki/Cognitive_science.
- [20] Wikipedia contributor (2015, May 10), List of Machine Learning con-cepts (version ID: 685871451)[online], available: http://en.wikipedia.org/wiki/List_of_machine_learning_concepts.
- [21] Wikipedia contributor (2015, April 12), Data Science (version ID: 7000450339) [online], available: http://en.wikipedia.org/wiki/Data_science.
- [22] Hornik, K, Stinchcombe, M, and White, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2 (5): 359-366, 1989.
- [23] Orphanides Athanasios, "Monetary Policy Rules Based on Real-Time Data", *American Economic Review* vol. 91, no. 4, September 2001 (pp. 964-985).
- [24] Orphanides, Athanasios, Taylor Rules: FEDS Working Paper No. 2007-18. Available at SSRN: <https://ssrn.com/abstract=999563> or <http://dx.doi.org/10.2139/ssrn.999563>.
- [25] Malliaris A. G and M. Malliaris. Modeling federal funds rates: A comparison of four methodologies. *Neural Computing and Applications*, 18 (1): 37-44, 2009.
- [26] Svensson, L. Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review*, 41 (6): 1111-1146, 1997.
- [27] Teräsvirta T, Van Dijk, D and M. C. Medeiros, M. C.. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21 (4): 755774, 2005.
- [28] Gonzalez, S. Neural networks for macroeconomic forecasting: A complementary approach to linear regression models. Working Paper, Department of Finance Canada, 2000.
- [29] Moura Marcelo and Carvalho de Alexandre, What can Taylor rules say about monetary policy in Latin America? *Journal of Macroeconomics*, 2010, vol. 32, issue 1, 392-404.
- [30] Hammond, Gill, Kanbur, Ravi, and Prasad, Eswar S. Monetary Policy Challenges for Emerging Market Economies (August 4, 2009). *Brookings Global Economy and Development Working Paper No. 36*. Available at SSRN: <https://ssrn.com/abstract=1492191> or <http://dx.doi.org/10.2139/ssrn.1492191>.
- [31] Mohanty, M. S. & Klau Marc, Monetary policy rules in emerging market economies: issues and evidence, *BIS Working Papers No. 149*, 2004. Bank for International Settlements.
- [32] Egeli Birgul, Ozturan Meltem, Badur Bertan, Stock Market Prediction Using Artificial Neural Networks <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.134.1114>.
- [33] Hinterlang, Natascha, Predicting Monetary Policy Using Artificial Neural Networks, Contributions to the Annual Meeting of the Social Policy Association 2019: 30 years Fall of the Wall - Democracy and Market Economy - Session: Econometrics - Forecasting I, No. B07-V3, ZBW - Leibniz Information Center for Economic Affairs, Kiel, Hamburg, <https://www.econstor.eu/bitstream/10419/203503/1/VfS-2019-pid-25781.pdf>.
- [34] B. Franks, Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, first ed., in: *Wiley and SAS Business Series*, Wiley, 2012.
- [35] Haykin, S., *Neural Networks: A Comprehensive Foundation*. 2nd Edition, 1999. Prentice- Hall, Englewood Cliffs, NJ.
- [36] Unlocking Game-Changing Wireless Capabilities: Cisco and SITA help Copenhagen Airport Develop New Services for Transforming the Passenger Experience, Customer case study, CISCO (2012). http://www.cisco.com/en/US/prod/collateral/wireless/c36_696714_00_copenhagen_airport_cs.pdf.
- [37] Friedman M Benjamin, Monetary Policy, National Bureau of Economic Research (NBER) Working Paper 8057, October 2000, <https://www.nber.org/papers/w8057.pdf>.
- [38] Friedman M, The Role of Monetary Policy, *American Economic Review*, 58: 1-17.