2016

Artificial neural network model for inflation forecasting in Kenya

Mwangi, C. N. W.
Faculty of Information Technology (FIT)
Strathmore University

Artificial Neural Network Model for Inflation Forecasting in Kenya

Mwangi Carolyn Naomi Wanja

Master of Science in Information Technology

2016

Artificial Neural Network Model for Inflation Forecasting in Kenya

Mwangi Carolyn Naomi Wanja

085082

Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in Information Technology at Strathmore University

Faculty of Information Technology
Strathmore University
Nairobi, Kenya

June, 2016

Declaration

I declare that this work has not been previously submitted and approved for the award of a Degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

...……………...[*Name*]
……………….. [*Signature*]
……………….. [*Date*]

Approval
The thesis of Mwangi Carolyn Wanja was reviewed and approved by the following:

Prof. Ismail Ateya Lukandu
Senior Lecturer, Faculty of Information Technology
Strathmore University
Supervisor

Dr. Joseph Orero
Dean of Faculty of Information Technology
Strathmore University

Professor Ruth Kiraka
Dean, School of Graduate Studies
Strathmore University

Abstract

Forecasts are important in decision making and entail prediction of a future state of a particular subject of interest. These forecasts depend heavily on historical data and the assumption that the past behaviour of forecast inputs will replicate itself in the future. Current linear and macroeconomic theory forecasting models used in Kenya lack reliable accuracy when predictors are futuristic and subject to changes over time. Artificial Neural Network (ANN) allow for the model to be more versatile in incorporating new predictors without altering the structure of the model. They work exceptionally well in environments that are nonlinear and where data is noisy and sometimes unavailable. The structure for the proposed model is a Neural Network with Back Propagation learning algorithm incorporating rainfall and M-Pesa use effects as additional inflation variables. The Backpropagation Neural Network was selected as a useful alternative due to the non-linear data used and to facilitate forecasting of future values. The adaptability of ANNs makes them most suitable for dynamic forecasting and classification problems. The results obtained from the model indicated that the back propagation was an appropriate algorithm that can be implemented in the process of inflation forecasting. The forecasting was done based on inflation variables identified as true inputs to the process of inflation forecasting. The model accuracy performance at 71.4286 % showed that the model is reliable as a tool for inflation forecasting. The study found that the optimum learning rate for the model was 0.5 while the momentum was at 0.9 for the training and 0.7 for the testing and validation data. Total iterations varied between the train, test and validate phases.

Dedication

This research is dedicated to my family: My two lovely and wonderful sisters Elizabeth Faith Nyokabi and Bilha Trista Wamahiga,you are a blessing to me. To my parents Paul Mwangi Njuguna and Rahab Wamucii Githae, thank you for walking this journey with me and encouraging me along the way, and most importantly to Almighty God for who You are to me.

Acknowledgements

I would like to acknowledge the Almighty God for being the source of my inspiration, good health and sanity of mind. I acknowledge my family for their prayers and for always being my cheer leaders. I would also wish to thank my colleagues at Strathmore University class of September 2014, you were a great team especially so Christine Njeri, thank you. I would wish to also appreciate the help, guidance and utmost patience of my supervisor Prof. Ismail Ateya Lukandu throughout the entire study.

Abbreviations/ Acronyms

ANN/NN	**-**	Artificial Neural Network/ Neural Network
BPNN	**-**	Back propagation Neural Network
CBK	**-**	Central Bank of Kenya
CPI	**-**	Consumer Price Index
ECM	**-**	Error Correction Model
EIA	**-**	Energy Information Administration
GDP	**-**	Gross Domestic Product
IMF	**-**	International Monetary Fund
MAE	-	Mean Absolute Error
KSH	-	Kenya Shillings
RMSE	-	Root Mean Squared Error
SARIMA	-	Seasonal Autoregressive Integrated Moving Averages
SSA	**-**	Sub-Saharan Africa
SSE	-	Sum of Squared Errors
SVAR	-	Structural Vector Aggressive
UK	**-**	United Kingdom
USA	**-**	United States of America
VAR	**-**	Vector Autoregressive model
YoY	-	Year on Year
KNBS	-	Kenya National Bureau of Statistics
RME	-	Root Mean Error

Definition of Terms

CPI Consumer Price Index is an indicator of price level changes of a market basket of consumer goods and services as purchased by households (Papadamou & Markopoulos, 2012)

GDP Gross Domestic Product is the monetary value of all services and finished products produced in a country within a given time period. It is the measure of a country's' wealth (Sermpinis, Stasinakis, Theofilatos, & Karathanasopoulos, 2014)

Monetary Policy The measures a monetary authority of a country takes to control money supply for purposes of ensuring a stable currency by regulating money supply (Misati, Nyamongo, & Mwangi, 2013)

Output Gap It is an economic measure of the actual output of a country against its potential or actual output that it could achieve at full capacity (Stella & Stock, 2012).

Table of Contents

List of Figures

List of Tables

List of Equations

Chapter 1: Introduction

1.1 Background of the Study

Inflation is a state in which the cost of goods and services increase reducing the purchasing power of a currency which influences the cost of living and doing business. This humpers a country's Gross Domestic Product (GDP) and the Consumer Price Index (CPI) which determine the rate of economic growth (Misati et al., 2013). The central bank of Kenya is thus mandated with the task of regulating inflation and to keep the output gap at a manageable and efficient level. For such an exercise, central banks generate inflation reports that give periodic analytical inflation targets. The report contributes a lot to the development of many optimal economic policies, measures and planning strategies. This ideally makes inflation forecasting of great importance in maintaining transparency of credible monetary policy implementation (Elliott & Timmermann, 2013).

The African Development Bank, (2011) identifies three categories of inflation inputs or predictors which include, exogenous factors (world oil and food prices), structural factors (domestic production) and Policy variables (monetary policies, fiscal policies and exchange rate policies). In Kenya, inflation is mainly driven by the Money markets and external sector in the long run, while oil prices account for the main inflation driver in the short-run.

While forecasting inflation is a difficult task (Stock & Watson, 2008a), inability to foresee sharp inflation changes negatively affects economic growth and worsens poverty levels (AfDB, 2011). Majority of the inflation reports developed are either created from advanced linear forecasting models or macroeconomic theory models which can be used as a suitable benchmark for new methods of inflation forecasting (Stock & Watson, 2008b; Marcellino, 2008).

Several models have been applied in Kenya for inflation forecasting. The Granger causality and structural vector autoregressive (SVAR) models have been used to formulate a way of estimating factors influencing inflation and the inflation rate (Misati et al., 2013). Another model is the VAR model which has been used as an estimate of inflation in Kenya (Misati, Nyamongo, Njoroge, & Kaminchia, 2012). While Kiptui (2009) employed a Philips curve model approach to inflation. These regression and linear models fail to accurately predict some financial series data due to inherent noise patterns and nonlinear components. Financial data are not generally described well by simple linear structural models for noise or random walks while conventional time series analysis based on stationary processes for inflation forecasting have been found not to perform satisfactorily on financial time series (Li & Ma, 2010).

It is thus clear that keeping inflation under control is a major challenge while accurate inflation forecasting is extremely vital (Durevall & Sjö, 2012).Employing machine learning models is one of the most used techniques to inflation forecasting (McAdam & McNelis, 2005; Esquivel, 2009; Haider & Hanif, 2009; Bahrammirzaee, 2010a; Aamodt, 2010; Stock & Watson, 2011; Nazif Çatik & Karaçuka, 2012 ; Rossi, 2012 ; S, J, & H, 2013 ; Okasha & Yassen, 2013; Hurtado, Luis, Fregoso, & Héctor, 2013; Otto & Voss, 2014 ; Gupta & Kashyap, 2015) These models have been found to effectively include major variants of predictors in their unstable environments, capture linear and nonlinear properties and provide easier ways of predicting future inflation rates in both emerging and established economies (Elliott & Timmermann, 2013; Giacomini & Rossi, 2015).

Predictive, forecasting and classification tasks are a staple of Machine Learning, and artificial neural networks are one of several standard tools used to tackle them. Artificial neural networks have been overwhelmingly used in finance such as portfolio management, credit rating and predicting bankruptcy, forecasting exchange rates, predicting stock values, inflation and cash forecasting and others in order to achieve a reliable decision-making process through scientific approaches (Li & Ma, 2010; Aamodt, 2010).

Artificial Neural Networks (ANNs) are an effective tool in the field of pattern classification and their forecasting accuracy tends to excel over that of well-established linear regression models (Gerstner, 2010). ANNs are a class of computational machine learning algorithms implemented in computer software with the aim of enabling automatic learning and subsequently autonomous problem solving. All approaches to performing classification assume some knowledge of the data. ANNs are inspired by biological neural networks that can be used to generate predictive models from historical data.

An ANN model is made up of interconnected processing units, known as nodes or neurons (Aamodt, 2010). Neurons in a typical ANN model are organized into three layers, an input layer, one or more hidden layer(s) and an output layer. Input layer neurons refer to the independent variables whereas neurons in the output layer represent the dependent variable (Stahl & Jordanov, 2012). The hidden layer(s) replicates interaction effects between the inputs and outputs. These hidden layer(s) are critical for ANN models for identifying the complex patterns in the data. Each neuron in the network is connected to other neurons by an associated numerical value known as a ‘ weight ’ These weights represent the knowledge or information that network has about a specific problem (Philip, Taofiki, & Bidemi, 2011)

ANNs are free from statistical assumptions and more robust to missing and inaccurate data, able to detect and duplicate any complex nonlinear pattern in the data in theory (Gupta & Kashyap, 2015). ANNs easily learn from examples and capture subtle functional relationships among the data even when these relationships are hard to describe or unknown. This modelling approach with ability to learn from experience make ANNs very useful for solving many practical forecasting and prediction problems across different fields (Jan, Tseng, Wang, & Wang,2006; Gosasang,Chandraprakaikul,&Kiattisin,2010; Tsai & Chen, 2010 ; Panchal, Ganatra, Kosta, & Panchal, 2010 ; Youn & Gu,2010; Chiang,Cheng,& Chang,2012; Devi, Reddy, Kumar, Reddy, & Nayak, 2012; Bredart,2014; Tang, Dai, Yu, & Wan, 2015)

1.2 Problem Statement

Current linear forecasting models and macroeconomic theory models for inflation forecasting used in Kenya lack reliable accuracy when predictors are subject to changes over time. These models are not able to accurately identify, capture and represent non-linear aspects of predictors as is often the case in inflation forecasting (Li & Ma, 2010; Choudhary & Haider, 2011 ; AfDB, 2011; Durevall & Sjö, 2012; Misati et al., 2013)

Artificial Neural Network (ANN) allow for the proposed ANN model to be more versatile in incorporating new futuristic predictors without altering the structure of the model. ANNs also have an added advantage of working exceptionally well in environments that are nonlinear and where data is noisy and sometimes unavailable (Choudhary & Haider, 2011; Rossi, 2012; Misati et al., 2013). The adaptability of ANNs thus makes them most suitable for dynamic forecasting and classification problems (Nazif & Karaçuka, 2012; Gupta & Kashyap, 2015)

1.3 Research Objectives

i. To identify variables for inflation forecasting in Kenya
ii. To identify challenges of current inflation forecasting models in Kenya
iii. To examine existing methods of inflation forecasting
iv. To propose the use of an Artificial Neural Network model for inflation forecasting
v. To validate the proposed Artificial Neural Network model

1.4 Research Questions

i. What are the variables for inflation forecasting in Kenya?

ii. What are the challenges facing inflation forecasting in Kenya?

iii. What methods have been used for inflation forecasting in Kenya?

iv. How was the proposed Artificial Neural Network model designed?

v. How was the Artificial Neural Network model validated?

1.5 Justification

Inflation forecasting is a key role of any monetary regulating body. It is vital for an economy to be able to foresee future changes in Inflation rates as it affects the quality of life and the ease of doing business in a country (Andrle, Berg, Morales, Portillo, & Vlcek, 2014). Policy making thus heavily depends on the accuracy of such forecasts for purposes of planning, execution and implementation. This study assisted in exploring the use of a different model that captures the different aspects of inflation and inflation predictors in a way that gives a true reflection of their effects on inflation. Artificial Neural networks (ANNs) have been widely acclaimed to solve many forecasting and decision making challenges by easily modelling parametric and non-parametric processes. ANNs are easy to integrate with information systems, can learn automatically how to perform forecasts and decision making without human intervention. These attributes improve forecasting for variables that are complex, non-linear and noisy. ANNs therefore are practical since they are precise and are able to capture any data movements with a high degree of accuracy making them suitable for use in inflation forecasting (Gupta & Kashyap, 2015).

1.6 Scope

The study mainly focused on designing an Artificial Neural Network model for inflation forecasting inclusive of new predictors as discussed in detail in the study. The model focused on incorporating M-pesa and rainfall shortages in inflation forecasting in Kenya. The data used to train and test the model was collected from secondary data sources. The assumptions made were that the variables chosen for forecasting were ultimate and true inputs to inflation and that the data-split periods were optimum periods.

1.7 Limitations

The limitations of this study were as follows. The study did not explore aspects of forecasting futuristic predictors and only used historical secondary data. The study also did not explore non-economic factors that may contribute to inflation and these were considered to be outside the scope of the study. Due to the complexity of the model design, it was not possible to create a simulation of the entire system nor address the model implementation approach and support services. The model however demonstrated the outputs created from the forecasting process.

Chapter 2: Literature Review

2.1 Introduction

Forecasts are important in decision making and entail prediction of a future state of a particular subject of interest. The process of forecasting involves identifying inputs that are relevant to the prediction process and using them to determine the possible future outcome of the subject under study. These forecasts depend heavily on historical data and the assumption that the past behaviour of forecast inputs will replicate itself in the future. A good forecast model is thus defined by its reliability, ease of use, having an output that is meaningful, compatibility with other systems, timeliness of the forecast and reliable accuracy (Arienda, Asana, & Constantino, 2015).

Forecasts get distorted as the horizon goes further into the future and no single technique can be used entirely on its own. There are certain aspects that direct the selection of a forecasting model. The cost of the forecasting technique must be justified in terms of the importance of the forecasts to critical decision making. The cost of the forecasting technique must be justified in terms of the importance of the forecasts to critical decision making. The accuracy levels of the output given by the predictor have to be optimal and specific to the process of decision making. Availability and accuracy of historical data is critical as different models would require different data types for forecasting (Subramaniam, 2009) and the availability and use of computers. Most forecasting methods require the use of a computer system with specific types of software to run the predictor of choice. The total time needed for data collection and analysis also affects the type of forecast that can be done. Some forecasters would require more time to build and validate as compared to others and the desired forecast horizon, be it short or long run. Different forecasting methods thus perform optimally different as some would be more suited to one forecast horizon than the other (Arienda et al., 2015).

2.2 Attempts at Inflation Forecasting in Kenya

Kenya plays a significant role in the East African region and has experienced some significant economic growth over the past couple of decades. Inflation predictors of particular interest to Kenya is the development of mobile financial services that have revolutionized the money markets specifically M-pesa since its inception in 2007 affecting money velocity (Mbiti & Weil, 2011). This has been attributed to the widespread use of integrated mobile-phone money services contributing to over 60% of the country's GDP (Andrle, Berg, Morales, Portillo, & Vlcek, 2013). Another area of focus would be the adverse effects of rainfall affecting domestic production and supply, and the rising oil prices.

To fully capture the different aspects of inflation, it would be vital to incorporate these factors into the inflation forecasting model (AfDB, 2011). Inflation inputs are subject to change and instabilities which make the task of forecasting very challenging. The most critical aspect of inflation forecasting however, would be to first identify the predictors or combination of predictors that actually have a big impact in determining inflation rates (Stock & Watson, 2010; Giacomini & Rossi, 2015).

Meyer and Pasaogullari (2010) suggest that there are numerous ways to forecast inflation. These vary from sophisticated statistical models with hundreds of variables to simple judgmental hunches based on past experiences. In Kenya, there have been several attempts at inflation forecasting in the recent past using different forecasting approaches.

Durevall and Ndung'u (2001) used a progressive research strategy to model inflation in Kenya from 1974-1996 by developing a parsimonious single-equation error correction and empirically constant model. Their findings indicated that the exchange rate, foreign prices, and terms of trade have long-run effects on inflation, while money supply and interest rate only have short run effects. However they noted that money supply drives inflation as one of the key inflation predictors. The model is represented by the equation 2.1

Equation 2.1 Parsimonious Single Equation Error Correction

$$\Delta p_t = \pi_0 + \sum_{i-1}^{k-1} \pi_{1i} \Delta p_{t-i} + \sum_{i=0}^{k-1} \pi_{2i} \Delta m_{t-i} + \sum_{i=0}^{k-1} \pi_{3i} \Delta y_{t-i} + \sum_{i=0}^{k-1} \pi_{4i} \Delta R_{t-i} + \sum_{i=0}^{k-1} \pi_{5i} \Delta e_{t-i} + \sum_{i=0}^{k-1} \pi_{6i} \Delta p^f{}_{t-i}$$
$$+ \sum_{i=0}^{k-1} \pi_{7i} \Delta pm_{t-i} + \alpha_1 (m - p - \gamma_1 y - \gamma_2 R)_{t-1} + \alpha_2 (p - e - p^f - \tau)_{t-1} + \sum_{i=1}^{3} \pi_{8i} S_{it} + \pi_9 D_t + v_t, \quad (2.1)$$

Where the delta Δ is identified as the first difference operator and the vt is a white noise process. Dt denotes a vector of deterministic variables such as constant, centred seasonal dummies, and impulse dummies, and an interaction dummy for the changes in inflationary inertia. The identifier pm_t is the log of price of maize grain, while *m* is the log of the money stock. The values of *p* is the log of the domestic price level while *y* is the log of real output. Value *R* represents a vector of rates of returns on various assets while *e* denotes log of the exchange rate and p^f is the log of foreign prices. The value τ is identified as the log of the trend in the real exchange rate (Durevall & Ndung'u, 2001).

Kiptui (2009) in his study explored the effects of oil prices on inflation in Kenya. This is extensively discussed in his research on the pass-through effects of world oil prices on the inflation in Kenya. In his findings, he observed that there existed a high correlation between world oil prices and inflation through evaluating a short-run oil price pass-through into inflation. He used a traditional Philips curve to derive estimates for inflation represented by the equation 2.2 (Kiptui , 2009)

Equation 2.2 Traditional Philips Curve Function

$$\Delta CPI_t = \alpha + \sum_{i=1}^{i=m} \beta_i \Delta CPI_{t-i} + \sum_{i=0}^{i=m} \gamma (y_{t-i} - \overline{y_{t-i}}) + \sum_{i=0}^{i=m} \theta_i \, \Delta oilp_{t-i} + \sum_{i=0}^{i=m} \lambda \Delta EXCH_{t-i}$$

(2.2)

Where the delta value **Δ** CPI identifies the change in the logarithm of the annual CPI index and *y* represents the real GDP. In the equation, y is the Hodrick-Precott filtered trend of real output while the delta value **Δ** OILPD is the annual change in the logarithm of the price in US dollars of a barrel of Dubai Petroleum as identified in the study. The delta value **Δ** EXCH captures the change in the logarithm of the annual average nominal exchange rate. In his conclusion, Kiptui (2009) identified that oil prices, exchange rate and aggregate demand conditions had significant effects on inflation with inflation at the short-run being at 0.01 and at 0.64 in the long-run.

Misati et al. (2013) observed the inflation in Kenya by capturing effects of food, non-food and non-fuel factors as key contributors to inflation in Kenya. The study implemented a Granger causality function combined with structural vector autoregressive (SVAR) functions to establish the effects of these factors on inflation in Kenya. The Granger model was represented by the equations 2.3 and 2.4 (Misati et al., 2013).

Equation 2.3 Grander Model Function

$$B(L)X_t = \varepsilon_t,$$

(2.3)

While the SVAR function was represented by the below formula estimation:

Equation 2.4 SVAR Function

$$y_t = [CommodityP_t, outputgap_t, M3GDP_t, Interest_t, ER_t, Inflation_t] \tag{2.4}$$

Where ER denotes the exchange rate which is used to capture the effects of exchange rate on inflation and M3GDP is the broad money supply and the endogenous variables identified as X_t are captured as the distributed time lags of the exogenous variables, given by elements of e_t. The final findings of the research were that oil and food price changes had a significant effect on inflation while the oil prices effects are more persistent than food and other non-food factors on inflation.

Andrle et al. (2013) developed a semi-structural new-Keynesian open-economy model, with two separated food and non-food inflation dynamics. The study achieved this by introducing two separate Phillips curves, one for food and one for non–food. The Keynesian model was represented by the equation 2.5

Equation 2.5 New-Keynesian Model

$$\begin{aligned} p_t^{cpi} &= wp_t^f + (1-w)p_t^{nf}, \\ \pi_t^{cpi} &= 4(p_t^{cpi} - p_{t-1}^{cpi}) = w\pi_t^f + (1-w)\pi_t^{nf}, \\ \pi_t^{4,cpi} &= p_t^{cpi} - p_{t-4}^{cpi}, \end{aligned} \tag{2.5}$$

Where the headline price index (p_t^{cpi} , all prices) represented the weighted sum of food (p_t^{f}) and (p_t^{nf}) captures the non–food prices. The value w is the weight of food, while πt identifies the quarterly inflation rate (annualized) and the year–on–year inflation rate represented by π_t^{4} . The Philips curves introduced for food and non-food inflation prediction were represented by the equations 2.6 and 2.7 respectively.

Equation 2.6 Non-Food Inflation Philips Curve Equation

$$\pi_t^{nf} - \overline{\pi}_t^{nf} = \lambda_1(E_t(\pi_{t+1}^{nf}) - \overline{\pi}_t^{nf}) + (1-\lambda_1-\lambda_2)(\pi_{t-1}^{nf} - \overline{\pi}_t^{nf}) + \lambda_2(\pi_{nf,t}^{imp} - \overline{\pi}_t^{imp}) + \lambda_3 rmc_t^{nf} + \varepsilon_t^{\pi nf}, \tag{2.6}$$

Equation 2.7 Food Inflation Philips Curve Equation

$$\pi_t^f - \overline{\pi}_t^f = b_1(E_t(\pi_{t+1}^f) - \overline{\pi}_t^f) + (1-b_1-b_2)(\pi_{t-1}^f - \overline{\pi}_t^f) + b_2(\pi^{imp} - \overline{\pi}_t^{imp}) + b_3 rmc_t^f + \varepsilon_t^{\pi f}, \tag{2.7}$$

Where (rmc_t^{nf}) was the real marginal costs in the non–food sector and $\lambda 2(\pi_{nf,t}^{\ imp} - \pi_t^{-imp})$ represents imported inflation.

Gikungu (2015) modelled the inflation in Kenya by using the SARIMA forecasting model. The model captured inflation data from 1981-2013 provided by KNBS. The model used the equation 2.8 to come up with the inflation forecast over a quarterly time period. In this study RME, MAE and MAPE were used as performance measures over an out of sample forecast for eight quarters.

Equation 2.8 SARIMA Forecasting Function

$$\phi_p(z)\Phi_P(z^s)(1-z)^d(1-z^s)^D Y_t = \theta_q(z)\Theta_Q(z^s)\,\epsilon_t \tag{2.8}$$

Where $\phi_p(z)$, $\Phi_P(z^s)$, $\theta_q(z)$ and $\Theta_Q(z^s)$ represented the characteristic polynomials of orders p, P, q and Q respectively while d and D are the orders of non-seasonal and seasonal differencing respectively.

2.2.1 Inflation Predictors for Kenya

According to Durevall and Sjö (2012), the most important task in the efforts to monitor and control inflation is to econometrically identify and capture inflations' main drivers. It is then the task of the central Bank to keep inflation in check. Most Central banks in advanced and emerging markets employ a variety of models to forecast inflation and keep track of the output gap (Andrle et al., 2013).

Kiptui (2009), African Development Bank -AfDB (2011) and Misati et al. (2013) identified a number of factors discussed as the main predictors of inflation in Kenya. AfDB (2011) identified major drivers of inflation in Kenya as being global oil price changes, money expansion due to growth of money supply in the short-run, world food prices and cereal production. Adverse weather conditions and especially the reduction in rainfall over the recent past have been attributed to affect local food production. The paper recommended, inclusion of rainfall measure to capture local food supply shocks and effects of M-pesa use to identify the effects on money growth and supply.

2.2.1.1 Policy Variables

Policy variables define the monetary, fiscal and exchange rate policies that influence money markets. Previous research has shown that these money markets have a significant influence on inflation in the long-run horizon in Kenya. The policy variable used as a measure of inflation was the monetary exchange rates (AfDB, 2011).

Otuori (2013) noted that a rise in exchange-rate volatilities had negative consequences on the trade sector which captures both exports and import of the local economy. Exchange rates are thus a significant factor affecting inflation because they affect investments, price stability and balance of trade in a country which inevitably affect the inflation rate (Andrle et al., 2013).

2.2.1.2 Food Prices

Food supply shocks affect inflation in Kenya in the short-run horizon according to Durevall and Sjö (2012), however according to AfDB (2011), the effects of food shock on inflation are significant in both the short and long-run horizons.

The production of cereals is identified as the most volatile of food production and the most likely to influence food prices since these cereals make up a huge part of the diet for majority of the people in Kenya (Durevall & Sjö, 2012; Misati et al.,2013) . Kenya imports a lot of it's cereals and thus affected by changes in the prices of imported food seeing that almost 40% of the total cereal consumption, mainly maize, is imported.

Figure 2.1 Comparison of International and Local Food Prices (Andrle et al, 2014).

According to Misati and Munene (2015), food prices are a significant inflation factor since food occupies a weight of 36 percent in the consumer price index contributing a monthly average of over 40 percent to the overall inflation value.

The effects of changes in imported food can be illustrate by making a comparison between the food price index in Kenya against a price index of international food prices as illustrated in the Figure 2.1 (Andrle et al., 2014) .

2.2.1.3 World Oil Prices

Kenya is a net importer of crude oil, thus changes in world oil prices significantly affect inflation, this leads to increased fuel prices and ultimately increased domestic and international food prices. Clearly oil prices are a significant cause of inflation since it affects many other factors directly or indirectly and thus regarded as imported inflation (Misati et al., 2013). Figure 2.2 illustrates the significance of oil price changes on other inflation predictors captured from 2006-2011.

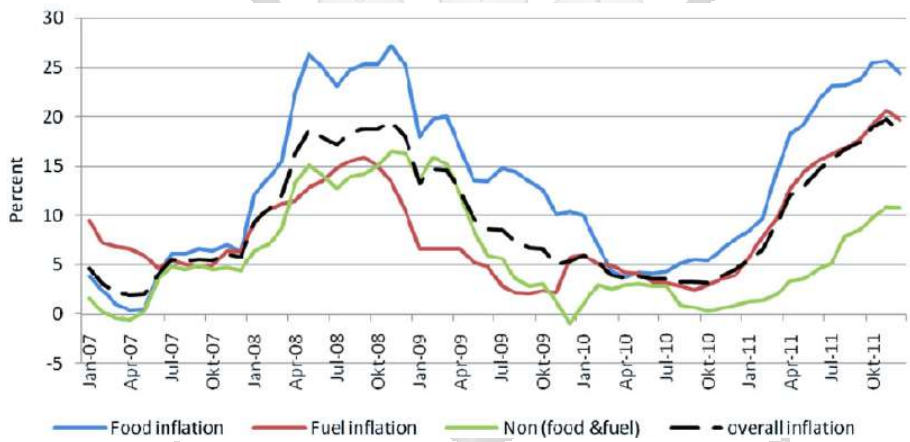

Figure 2.2 World oil price effects on inflation predictors and overall inflation (Misati et al, 2013)

2.2.1.4 Domestic Production/ Output Gap

Output gap defines the disparity between expected output and the actual output of domestic production. The production is mainly of products that are intended for the local market. output gap therefore measures the disparities between that which is produced versus the actual production (Cheng, 2006).

Output gap negatively affects the inflation rates. This is so because a deficit production results in a price increase of consumer goods which makes domestic production a big determinant in the rate of inflation (AfDB, 2011).

2.2.1.5 M-Pesa Use

E-money financial innovations in Kenya especially the use of M-pesa have revolutionized the Kenyan economy in the recent past. This is because M-pesa moves more money locally than Western Union does globally (AfDB, 2011). The use of M-pesa also extends mobile banking capabilities that cater to more than 70% of the population in Kenya (Mbiti & Weil, 2011). Thus the use of M-pesa has led to an increase in money velocity which ultimately affects inflation expectations (AfDB, 2011).

According to Aron and Muellbauer (2015) the spread of mobile money promotes increased velocity of circulation which increases 'effective money' and hence increases inflation. They argue that, in the short run, higher mobile balances might signal plans for impending spending and so proxy a short-term demand increase. Also, the advent of mobile money might transfer spending power to households with a higher propensity to spend, and so reduce saving which affects inflation. Figure 3.5 illustrates the relationship between use of M-pesa and money velocity. Mbiti and Weil (2013) confirm that the use of M-pesa has escalated the money velocity in Kenya since its inception in July 2007 and thus has had a significant effect on the overall money supply which affects the inflation rates.

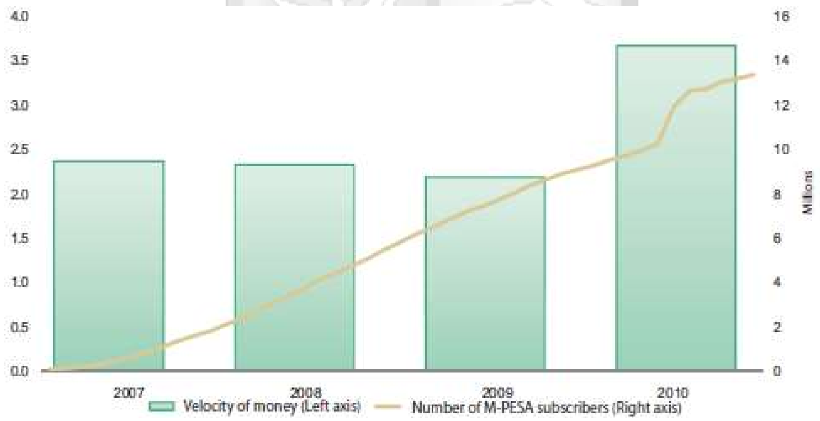

Figure 2.3 Trends in M-Pesa Subscribers (AfDB, 2011)

2.2.1.6 Rainfall

Variations in rainfall often leads to inconsistencies in local production that fuel supply shocks. These supply shocks translate into increased food prices. Rainfall insufficiencies aggravate the production of maize and food security which affects the domestic supply of food (Andrle et al., 2013). It has also been seen that rainfall variations play an even more significant role in influencing inflation by affecting fuel and non-food prices (Aron, Muellbauer, Sebudde, & others, 2015).

McSweeney, New, and Lizcano (2010) indicate the variations in rainfall for Kenya over the past couple of decades have had an effect on the overall food production which in turn affects food prices. Food prices are identified as one of the predictors of inflation, thus an increase in the prices of food due to productivity variations affected by rainfall quantities ultimately affect the inflation rate.

2.2.2 Challenges of Inflation Forecasting In Kenya

Forecasting inflation is key in helping the central bank to adjust its monetary policy to control inflation (Moriyama & Naseer, 2009). However inflation forecasting is a challenging exercise (Stock & Watson, 2008a). According to the International Monetary Fund (2011), one of the challenges identified is the uncertainty of the impact of money growth and money velocity due to increased use of mobile money.

The accuracy levels of a forecasting model have to be optimal and specific to the decision making process. However majority of the linear autoregressive models used for forecasting inflation in Kenya are not very good at forecasting inflation due to their low precision rates (Aron & Muellbauer, 2015). The International Monetary Fund (2011) recommends developing the inflation forecasting capacity by adopting more versatile models.

Gichuki, Oduor, and Kosimbei (2012) note that the Central Bank of Kenya (CBK) has fallen short of its inflation target several times over due to the use of different forecasting tools that have not been efficient in identifying accurate forecast values for inflation prediction and monitoring.

2.3 Inflation Forecasting Models

Inflation forecasting has been done using different models across varying economic environments. A model is a scientific tool used to simulate how a concept works for purposes of understanding and evaluating the different aspects of that concept. This is achieved by defining how the processes and data flow interact in the designed system. (Andic & Ogunc, 2015). Different forecasting models have thus been applied in an effort to capture critical variables of inflation and help in forecasting inflation rates for an economy. Figure 2.4 shows the process of model building and forecasting phases.

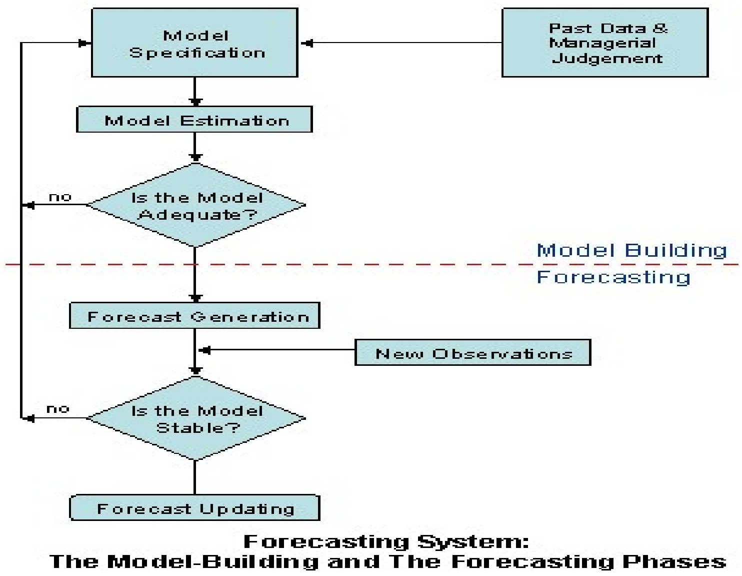

Figure 2.4 Basic Model of a Forecasting System (Arsham, 2015)

2.3.1 Inflation Forecasting Techniques

Forecasting techniques can be identified as causal, time series and smoothing techniques (Hall, 2002a). These forecasting techniques are data intensive and classified as illustration in the Figure 2.5

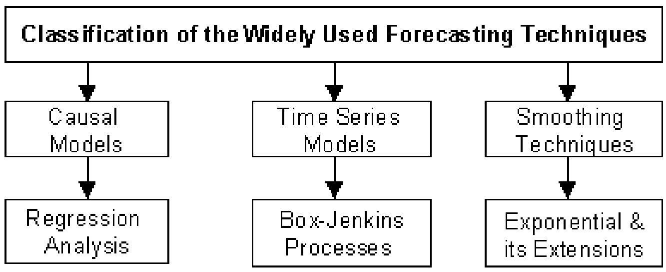

Figure 2.5 Classification of Forecasting Techniques (Arsham, 2015)

Forecasting techniques can also be categorized into judgemental, consumer/ market research, cause-effect and artificial intelligence as per the Figure 2.6

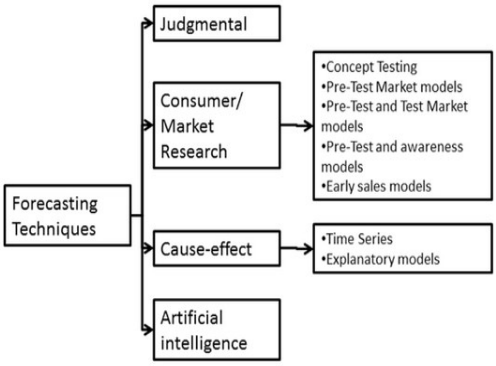

Figure 2. 6 Classification of Forecasting Techniques 2 (Webby & O'Connor, 2014)

2.3.1.1 Causal Models

These models use regression analysis for forecasting. They try to establish cause effect relationships between the forecast and the forecast variables (Benkachcha et al., 2013). The model analyses past trends and identifies major factors that have the most significant effect on the forecast and represent these factors in a mathematical function. According to (Campbell & Campbell, 2008) a simple form of a causal regression model is represented by the function(2.9)

Equation 2.9 Causal Regression Function

$$Y = \beta_0 + \beta_1 X + u \tag{2.9}$$

where β_0 is the intercept parameter and β_1 the slope parameter. Y represents the dependent variable and X is the independent variable. While u is the identifier for the slope variation.

2.3.1.2 Time Series Models

Chatfield (2013) defines time series as a collection of observations made sequentially through time and can be represented as a mathematical function *X(t), t=0,1,2…..* Where t is the time lapsed and *X(t)* is a random variable. The main objective of time series model is to understand the flow or sequence of activities under observation and determine how they was flow into the future. A time series can be univariate consisting of data collection on a single variable or it can be multivariate with multiple variables under study. Time series models are characterised by four main components, the trend, cyclical, seasonal and irregular component.

The trend cycle identifies historical changes in the time series while the cyclical component captures the medium-term changes in the series. The seasonal component is useful in identifying fluctuations and irregular component features the changes in the series caused by unpredictable factors (Adhikari & Agrawal, 2013). The time series models are popular for forecasting in financial and economic business environments. A time series model can be represented by the below function defining it as either multiplicative of additive (Montgomery, Jennings, & Kulahci, 2015).

Additive: makes the assumption that components are independent and none affects the other

Equation 2.10 Time Series Additive Function

$$Y\ (t)=T(t)+S(t)+C(t)+I(t) \tag{2.10}$$

Multiplicative Model: makes the assumption that components affect each other and thus are interdependent

Equation 2.11Times series multiplicative function

$$Y\ (t)=T\ (t)*S(t)*C(t)*I(t) \tag{2.11}$$

2.3.1.3 Smoothing Techniques

These are techniques related to time series forecasting models and are used to refine the forecast quality of the time series (Montgomery et al., 2015). The most common smoothing techniques are moving averages and exponential smoothing. A simple exponential smoothing model can be represented by a function that captures the smoothing process (Eva Ostertagova, 2011). The function takes the below form.

Equation 2.12 Smoothing Function

$$\hat{y}_{i+1} = \alpha y_i + (1-\alpha)\hat{y}_i \tag{2.12}$$

where y_i represents the actual series value for the time period i, $\hat{}_i y$ represents forecasted value of Y for the time period I while ,$\hat{}y_{i+1}$ is the forecast value for time period i +1 and alpha sign α represents the smoothing constant used.

2.3.1.4 Judgemental Models

These are models that depend on the opinion of experts based on quantitative data which is often historical (Hyndman & Athanasopoulos, 2012). Though these models have been seen as inferior to other forecasting models, they can be very useful when there is little or no sufficient data for statistical analysis techniques. Situations most suitable for judgemental models are instances where the expert has insight into data that may not be captured in the output of a statistical model. Also, the models are used when the environment is suitable and expert has made and learned from previous similar forecasts. Examples of such models include the Delphi method which uses a series of questionnaires given to experts, surveys and scenario building (Kavanagh & Wasiams, 2014).

2.3.1.5 Machine learning Models

These models employ the use of Artificial intelligence in their forecasting process. The models have been proven to produce forecasts that are more accurate and can be used for both classification and prediction purposes (Hall, 2002b). Some of the most commonly used machine learning models include;

Neural networks-these imitate the human cognitive system and are designed to represent non-linear forecasts. They can be used on their own or as an integrated part of the application running the forecast. For example using them to forecast inflation (Elliott & Timmermann, 2013).

Bayesian belief networks- are models structured in a tree format where nodes represent specific variables and the connecting arcs identify the relationships between the variables. These can be used in multiple domains and generate probability forecasts based on some conditional state (van der Maas, 2014).

2.3.1.6 Expert systems

These are knowledge based systems designed around a specific field of interest. These types of systems use rules and formal knowledge representation to emulate a human expert (Medsker, 2012). They can be used in place of people to make predictions based on currents trends like weather prediction (Alsaiari, 2012) and demand forecasting (Vizureanu, 2010).

2.4 Conceptual Architectural Model

Artificial Neural networks (ANNs) have been widely acclaimed to solve many forecasting and decision making challenges by easily modelling parametric and non-parametric processes. In addition, these ANNs are able to capture non-linear and noisy data by transforming the input (Bahrammirzaee, 2010b). ANNs are easy to integrate with information systems, can learn automatically how to perform forecasts and decision making without human intervention. These attributes improve forecasting for variables that are complex.

NN therefore are practical since they are precise and are able to capture any data movements with a high degree of accuracy making them suitable for use in inflation forecasting (Gupta & Kashyap, 2015;Somaratna, Arunatilaka, & Premarathna, 2010)

NNs are also attractive for use because of their robustness and ease of adapting their performance to any changing characteristic of the modelled system, fast processing speed and ease of maintenance (Karlik & Olgac, 2010). The adaptability of NNs thus makes them most suitable for dynamic forecasting (Nazif & Karaçuka, 2012).These capabilities make ANNs one of the most suitable tools for macroeconomic time series forecasts. A new approach to the use of ANNs in macroeconomic forecasting has been the use of arithmetic combinations of established ANNs to maximize on the forecast accuracies achieved from each model of a Neural Network (Choudhary & Haider, 2011) .

ANNs have thus been widely discussed and their implementation in different countries around the world and proved their superiority over their econometric model counterparts in inflation forecasting. Binner, Bissoondeeal, Elger, Gazely, & Mullineux (2005) compares use of NN against linear models and concludes on the superiority of the NN performance over the linear model. Bahrammirzaee (2010) explores the use of NNs as a robust alternative to linear models in financial predictions and planning especially scenarios where data is non-linear. He concludes that the NN models produce better and more accurate outputs.

Multiple variations of the ANNs have been used as per the unique needs and economic structures of individual countries providing more accurate results from both within-sample and out-of-sample forecasts(Choudhary & Haider,2011 : Misati et al.,2013)

2.4.1 Artificial Neural Network

Neural networks were developed from research efforts to create systems with cognitive abilities inspired by the information processing structure of the biological human nervous system. The biological neuron learns by adjusting to the synaptic connections between the millions of neurons (Hyndman & Athanasopoulos, 2012). The number of neurons in the input and output layers is highly dependent on the inputs and outputs of the problem domain. The number of units in the hidden layer on the other hand depends on the problem complexity or the concept to be represented and solved (Gosasang et al., 2010). The weights received from the input layer and the adjusted weights are critical in determining the solutions' final output value. Thus, the higher the absolute value of the weight, the stronger the influence of the associated neuron. The ANN input layer receives input in the form of a data vector. This data vector is then weighted and passed through the connections to the hidden layer.

For each hidden unit/neuron, the incoming products of the weights and the input values are summed up and passed onto an activation/transfer function that generates an output value. This output then becomes input for the next layer and so on depending on the number of hidden layers within the ANN (Stahl & Jordanov, 2012) . The Figure 2.7 illustrates the structure of a biological neuron.

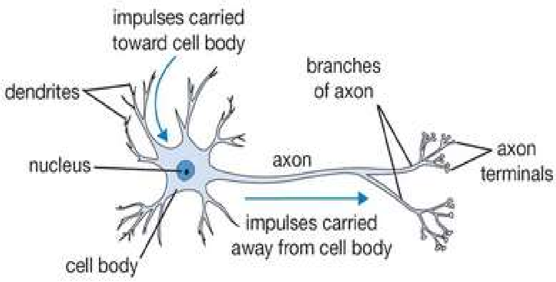

Figure 2.7 Structure of a Biological Neuron ("CS231n Convolutional Neural Networks for Visual Recognition,")

2.4.1.1 Artificial Neuron Structure

The artificial neuron known as a perceptron is modelled after the biological neuron with inputs or an input layer which are the dendrites represented as an approximated mathematical summation. The output axons that connect to other neurons are represented as the output of one neuron as shown in Figure 2.8 (Kashid & Kumar, 2013)

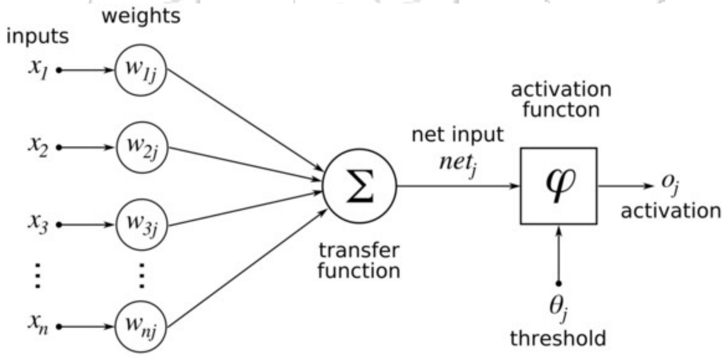

Figure 2.8 Structure of a Perceptron ("Artificial Neural Network/Activation Functions - Wikibooks, open books for an open world," 2015)

2.4.1.2 Artificial Neural Network Architecture

The architecture of the neural network is made up of an input layer that is connected to an intermediary or hidden layer. These neurons or inputs to the hidden layer are then connected to generate the output as depicted in the Figure 2.9 (Gupta & Kashyap, 2015)

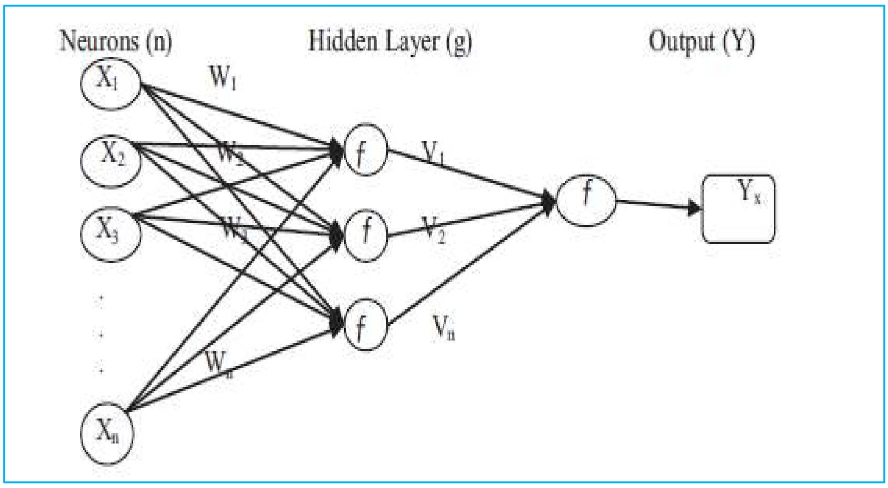

Figure 2.9 Architecture of an Artificial Neural Network (Gosasang et al., 2010)

The architecture refers to the organization of perceptrons and the types of connections permitted (Choudhary & Haider, 2011).The architecture of the neural network can be broadly categorized into;

i. The input layer

Consists of neurons X_0-X_n that are weighted by a predefined W_o -W_n value and connects the input neurons to the hidden layer after the weighted input is calculated by the relationship between the neurons and the weight as X_0W_{o}- X_nW_n. Input weights are used to regulate the input values to the artificial neuron and are specified between -1 and +1. The weighted inputs are constantly adjusted as the NN is trained to line it up with the stated criteria. The inputs can be defined by the below input function (Collova & Capaldo, 2013).

Equation 2.13 ANN Input Function

$$w_0 + w_1 x_1 + \ldots + w_n x_n = \sum_{i=0}^{n} w_i x_i = \vec{w} \cdot \vec{x} \tag{2.13}$$

ii. The summation and activation functions

The weighted inputs are summed up and parameterized and used within an activation function.

Many activation functions exist in NNs and are used as per the design and area of application of the NN such as the sigmoid, Gaussian, hyperbolic tanh,

A simple commonly used sigmoid activation function can be represented by the below equation (Collova & Capaldo, 2013).

Equation 2.14 Sigmoid Activation Function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{1 + e^{-x}} \tag{2.14}$$

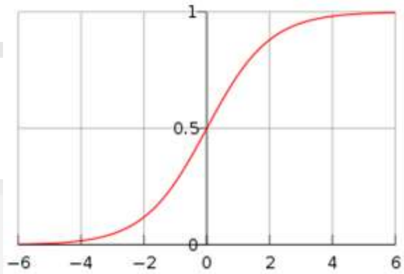

Figure 2.10 Sigmoid Activation Function Graph (Philip, Taofiki, & Bidemi, 2011)

iii. Output Layer

The information in the perceptron moves from inputs through the hidden layers and finally given as a single or multiple Boolean values 1 or 0. The results from the summation and activation layer determine the value of the output (Haider & Hanif, 2009).

2.5 Model Design

The structure for the proposed model is an ANN that was bear a similar structure as that used by Gupta & Kashyap (2015) in their use of Artificial Neural Network for inflation forecasting in the G-7 countries.

2.5.1 Data Specification and Selection of Input Variables

The study's objective was to forecast inflation using the Back Propagation Artificial Neural Network model applying a single hidden layer. Variable selection is at the heart of forecasting and determines how accurate and helpful the output information was be (Andic & Ogunc, 2015). Data variables used were captured from 2002-2015.

The data included effects of M-pesa use by capturing money velocity effects on inflation and rainfall as mentioned in the study. Use of M-pesa data helped capture the effects of increased money velocity, rainfall data captured domestic food supply shocks while World oil prices was substituted for import price index in the forecasting model inputs (African Development Bank [AfDB], 2011). For purposes of simplicity, the assumption made was that there was no statistical bias in the data.

2.5.2 Model Specification

The choice of model is one of the most critical things in forecasting (Andic & Ogunc, 2015)The model was a simple neural network with Back Propagation learning. The activation function selected for the model was the hyperbolic tahn function since a single layered NN performs better with the hyperbolic function (Özkan & Erbek, 2003 ;Karlik & Olgac, 2010). The function was represented by the equation 2.15.

Equation 2.15 Hyperbolic Tanh Activation Function

$$f(x) = \frac{sinx}{cosx} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2.15}$$

The model used a pseudo-out-of sample approach and output of the mode was compared to the actual inflation Figures recorded over the stipulated time period.

2.5.3 Specifying the ANN

Selecting the neural network hidden layer is a big determinant on the optimality of the forecast. Many hidden layers may make the ANN more powerful but could lead to overfitting, while too few layers could limit the capabilities of the model determined by the application of the ANN model (Gupta & Kashyap, 2015). A typical Back Propagation ANN can have at most 15 layers and at the least 1 layer. The model had a single input layer, a single hidden layer to process inputs and a single output layer. The algorithm was represented by the equation 2.16

Equation 2.16 ANN Hidden Layer Function

$$\Delta W_i(j) = -\beta\frac{\partial \lambda}{\partial W_i} + \alpha\Delta W_i(j-1) \tag{2.16}$$

2.5.3.1 Training and Testing the ANN

ANNs learn by experience through training by some supervised training set and involves adjusting to some change in the learning function. The learning process for an ANN comprises of training, validation, and testing phases. The training phase involves the adjustment of the weights from the comparisons done using a training dataset (usually part of the available dataset). The validation data subset is then used to prevent the ANN from overfitting or memorizing instead of generalizing the classifications of the results. This is done in the testing phase where the ANN is checked to see whether and how well it has learned the task. This checks whether the ANN learned the knowledge and can generalize new data which is the ability of a trained ANN to correctly classify new input data that have not been used in the training (Stahl & Jordanov, 2012). The training phase uses several learning rules to determine how the weights are being adjusted.

The ANN learning approaches are divided into two: supervised and unsupervised learning. Supervised learning provides the correct output for each input training data instance beforehand i.e it is known and provided to the ANN. The training input instances are provided to the network and the actual network output is then compared with the target or expected output. Any difference in the values prompts an adjustment of the input weights in a way as to minimize this difference (Kashid & Kumar, 2013). ANNs are a type of supervised learning. On the other hand for unsupervised learning, the output is not known, and the network itself adjusts the weights in order to map inputs to outputs without any additional knowledge training. The back propagation algorithm is the most popular training method for supervised learning (Stahl & Jordanov, 2012; Gupta & Kashyap, 2015)

2.5.3.2 Back Propagation Algorithm

This is a learning algorithm that gives input training values to a model together with the desired output. It works by seeking to minimize the error term between the output of the neural network model and the actual desired outputs. The error measure is then back propagated from the output to the input layer through the hidden layer so as to adjust the synaptic weights, this is a form of supervised learning (Kumar, Ram, & Hanmanthu, 2014). The Figure 2.11 illustrates a perceptron in a back propagated Artificial Neural Network.

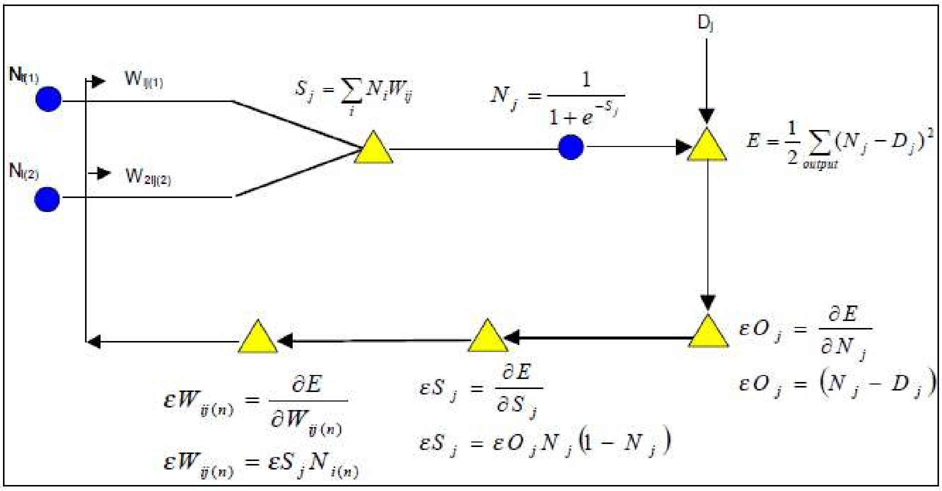

Figure 2.11 Artificial Neuron using Back Propagation Learning (Kumar, Ram, & Hanmanthu, 2014).

The methodology of Back Propagation supervised learning can be described as illustrated in the Figure 2.12

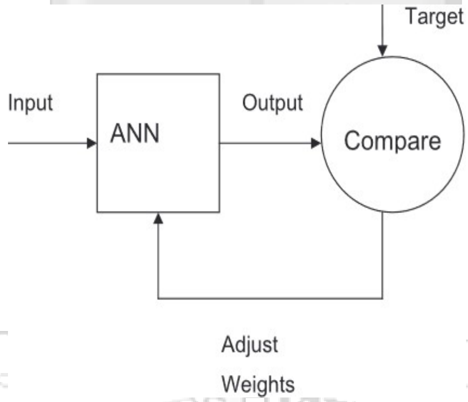

Figure 2.12 Methodology of Supervised Learning (Mohan, Jha, Laha, & Dutta, 2005)

2.5.3.2 Stopping Condition

The training of the neural network is stopped once the error is seen to fall below a lower value than the first. A most suitable stopping condition however would be the use of a validation set. This means that a set of data different from the training set are used to calculate the error of the output after each training cycle. The process of validation ensures that the model does not over train and end up overfitting data. Once the error reaches a minimum, the training stops. When the model is over trained, the validation error starts to raise as illustrated in the Figure 2.13.

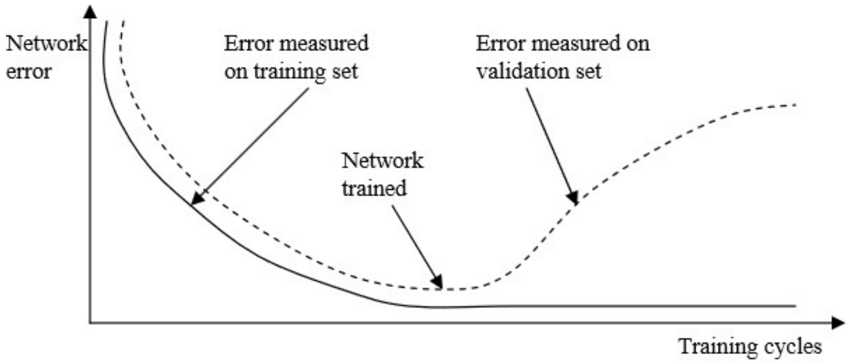

Figure 2.13 Use of Validation Sets (Kumar, Ram, & Hanmanthu, 2014)

The Back Propagation algorithm is normally faced with a challenge of local minima. This is where the algorithm gets stuck on a higher local error gradient than the global minima and cannot reduce the output error. To the weights can be adjusted a random number and the model retrained or by adding a momentum to the weights. The most suitable mitigation option was implementation of momentum identified at the point of training the model (Kumar, Ram, & Hanmanthu, 2014). Figure 2.14 illustrates local minima

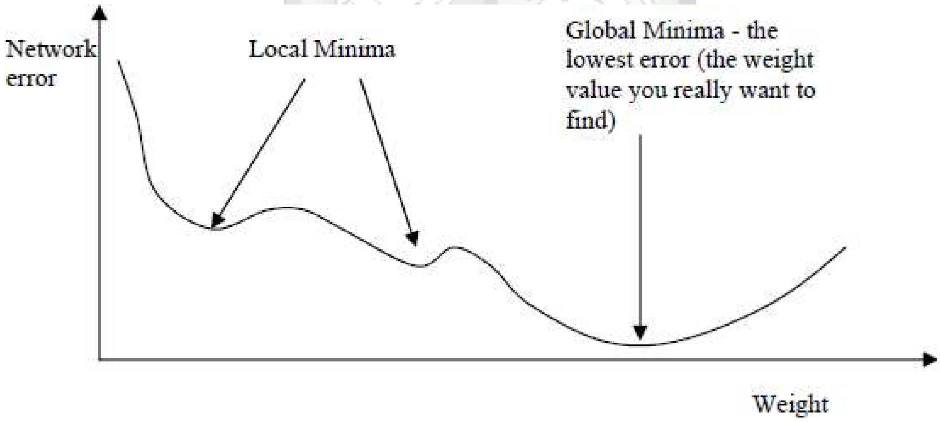

Figure 2.14 Local Minima (Kumar, Ram, & Hanmanthu, 2014).

2.5.4 Forecast Evaluation

The model was evaluated by determining the stated measures of performance, Given that *Yx* is the recorded output and *Yˆx* is the forecasted output while *p* being the number of predictions (Gupta & Kashyap, 2015)

Sum of the squared errors given by the formula;

Equation 2.17 Root Mean Squared Error

$$\text{RMSE} = \sqrt{\frac{(Y_x - \hat{Y}_x)^2}{p}} \tag{2.17}$$

Equation 2.18 Mean Absolute Error

$$\text{MAE} = \frac{1}{p}\sum_{x=1}^{p} |Y_x - \hat{Y}_x| \tag{2.18}$$

2.5.5 System Implementation Design

The proposed model was designed to run as a standalone or integrated part of a working system. This type of implementation allowed the model to be versatile in its use as per the unique needs of user systems. The model picked inputs from a specified input template as per the design of the target system of use. This helped format the data inputs in a way that was simple for the model to run and compute. The Figure 2.15 illustrates the model framework of the conceptual forecasting model to be build.

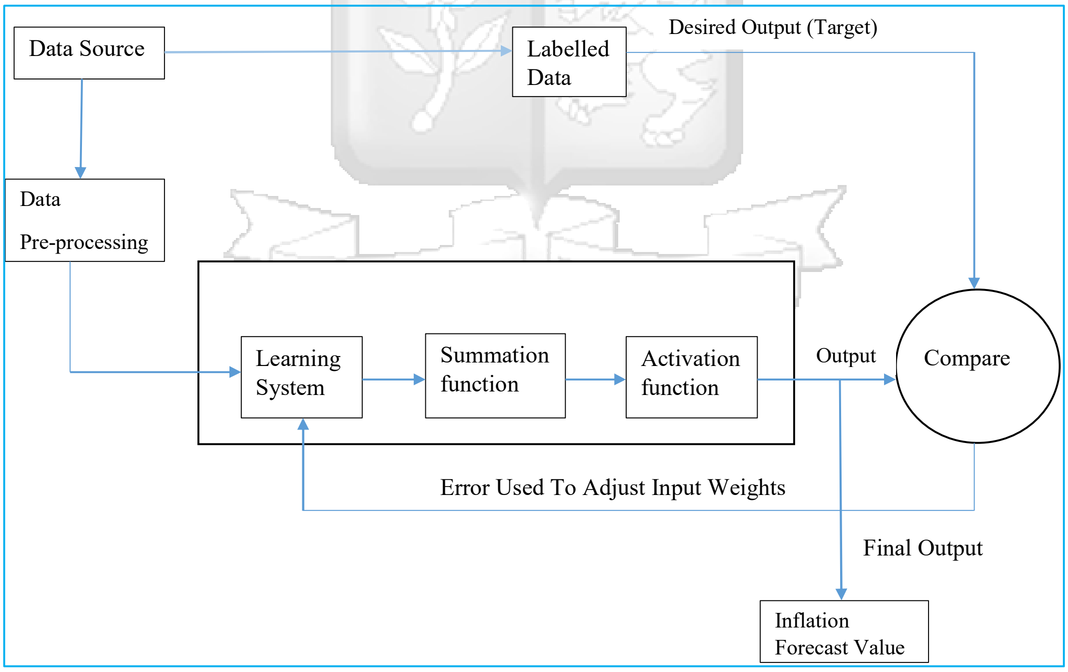

Figure 2.15 Proposed Model Framework

The Artificial Neural Network output of the processed data was represented in user defined result presentation modes such as line graphs.

Chapter 3: Research Methodology

3.1 Introduction

Research is defined as a fact finding activity that involves a scientific investigation or thorough study of a given subject matter of particular interest. A research can be explorative, descriptive or diagnostic in nature and thus qualitative or quantitative approaches are applied as per the research design. Research has been proved to be a vital tool that provides the basis for economic decision making by government institutions and policy makers (Mackey & Gass, 2013) .

The study took an applied exploratory approach to identifying new insights into the possibility of incorporating more diverse predictors for inflation forecasting that would best capture the changing economy structure of Kenya. The model used was intended to facilitate achievement of the stated objectives of the study.

To facilitate the research secondary data was used for analysis. The data was obtained from official publications on data and statistics relevant to the economic sector and specific to the financial industry.

3.2 Research Site

Kenya is the 4th largest economy in Sub Saharan Africa and stands out as a region leader in East Africa due to its service sector, specifically the financial and telecommunications services. However, inflation continues to prove to be a major challenge affecting negatively on the economic performance of the country (Durevall & Sjö, 2012). The Table 3.1 shows the inflation rate trends and its effect on GDP from 2000-2011. Table 3.2 indicates inflation for Kenya in Figures from 2000-2013.

Table 3.1 GDP and Inflation in Percentage from 2000-2011 (Durevall & Sjö, 2012)

GDP Growth an Inflation, in percent													
Country/Year		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Kenya	Real GDP growth	0.6	4.7	0.3	2.8	4.6	6	6.3	7	1.5	2.6	5.6	5
	Inflation	11.8	1.6	4.2	8.3	17.1	4.7	7.3	5.6	15.5	8	4.5	18.6

Table 3.2 Kenya Inflation Figures 2000-2013 (KNBS, 2015)

Inflation Trends 2000-2013		
	Base February 2009=100	
Year	Annual Weighted Average Index	Annual Inflation Rate
2000	49.89	10
2001	52.75	5.8
2002	53.79	2.0
2003	59.06	9.8
2004	66.03	11.8
2005	72.57	9.9
2006	76.95	6.0
2007	80.24	4.3
2008	92.36	15.1
2009	102.09	10.5
2010	106.26	4.1
2011	121.17	14.0
2012	132.53	9.4
2013	140.11	5.7

3.3 Research Design

The research design is a work plan detailing how the research was be undertaken, type of data to be collected, tools and techniques that was be employed to obtain data and the method of data analysis to be used (van Wyk, 2012). The research design adopted for the study was an exploratory-quantitative research approach (Mackey & Gass, 2013).

Exploratory research seeks to help identify new hidden factors that might be found in data and their significance to the study under research. This type of research approach creates a foundation for further future in-depth research from a scientific point of view with the intent of finding new insights and ideas about the topic under study. The exploratory design was adopted to investigate the effects of adding rainfall and M-pesa use in the process of forecasting inflation. Previous exploratory studies done on inflation forecasting have shown possible significant effects of mobile money use and effects of rainfall in the process of inflation forecasting for various economies (Jack & Suri, 2011;Kinda, 2011;Armantier et al., 2013;Aron & Muellbauer, 2015). The study was designed to investigate the effects of these additional variables to the inflation in Kenya. The ANN model was significant to the research design since new variables can be added or removed from it without altering the structure of the model design (Karlik & Olgac, 2010) .

A quantitative design is most suitable for a study whose variables can be measured or quantified. It is a scientific research that involves statistical or programming methods that aid in the process of decision making for an industry (van Wyk, 2012). The model input variables are quantifiable variables capturing the values for each input .The research design took on the below approach.

3.3.1 Acquisition of Data

The data collection method employed was nonexperimental and which included collection of secondary data from relevant publications and the internet. The secondary data collected included the inputs identified as predictors of inflation. These were values for each predictor from the identified time period as determined in the study. The inputs were statistical data that captured the values for each key inflation variable. The values were used as the model inputs for purposes of training, testing and validating the ANN model.

3.3.2 Sample Split

This involved identification of the data to be used in training the model, test data to test model training and the validation data to measure the output error. It is critical to identify the type of data to be put for each data set as identified in the research sampling procedure. The sample taken was representative of the identified population.

3.3.3 Model Training

This process involved giving inputs to the model for processing in order to train the model on the type of input data and the expected output of the training session. The training data was fed into the ANN model via the identified model neurons. The training process was done over a number of iterations between 500-1500 with each iteration targeted at reducing the error rate and adjusting the input weights.

3.3.4 Model Testing and Validation

This process involved the use of a test data set to check whether the system was properly trained by observing the actual model output versus the expected output. By using the validation data set, any disparities in the output captured by error performance measures were used to adjust the weights of the neurons for purposes of fine tuning the model.

3.3.5 Method of Output Presentation

Graphical representations such as graphs and tables were identified to be used to illustrate model outputs and capture the differences with the actual inflation data recorded.

3.4 Period of Study

A population defines all items in a field of study and is used to identify the target population. A target population is a population that is representative of the characteristics of the entire population and can be studied to gain more information about the population (Bordens & Abbott, 2011). Each population is defined within a time period of study. The research study was targeted at the financial sector. The study time period covered inflation variables data between 2002 and 2015. The study took into consideration six inflation variables identified for the time period of 13 years. These were data variables capturing annual values of data from the identified data sources. These were chosen due to the availability of data and its relevant accuracy which were critical to the process of forecasting inflation values.

3.5 Sampling

A sample is a set of items representative of a population selected for study with regards to the topic under research. The study used convenience non-probability sampling. This is a type of sampling where the researcher deliberately selects the items to include in the sample and is often used in explorative research (Battaglia, 2011; Mackey & Gass, 2013).

According to Hansen and Timmermann (2012) “Statistical tests of a model’s forecast performance are commonly conducted by splitting a given data set into an in-sample period, used for initial parameter estimation and model selection, and an out-of-sample period, used to evaluate forecast performance” . ”There are no broadly accepted guidelines for how to select the sample split. Instead, researchers have adopted a variety of practical approaches. One approach is to choose the initial estimation sample to have a minimum length and use the remaining sample for forecast evaluation. Another common approach is to do the reverse and reserve a certain sample length, e.g., 10 or 20 years of observations, for the out-of-sample period”. Thus the sample used for the study was variable data between 2002-2015. This sample was taken from the total identified population of 50 years.

The sample size was split into training and test data. This method of splitting the sample has been effectively used by Pesaran, Pick, and Timmermann, (2011) and Andic and Ogunc (2015).

3.6 Performance Measure

3.6.1 Data Reliability

Data reliability is defined as the ability of a research tool to give consistent results on repeated trials (Heale & Twycross, 2015; Heale & Twycross, 2015). The model output data was tested through repeated trials on the training phase. This was done to ensure that the error rate was reduced to the global minimum for the model in order to provide the most optimal input weights for each variable inputs. The model was evaluated by calculating measures of performance as identified in the study chapter 2. This was done using the root mean squared error and the mean absolute error. This measures were calculated against the actual recorded inflation values in comparison with the inflation values forecasted by the model over the same period of time

3.6.2 Data Validity

Validity is defined as the ability of a research tool to measure what it is supposed to measure (Bordens & Abbott, 2011). The validity of the research can be external, construct, content or internal validity. External validity refers to the ability of the results to be generalised to the entire population. Content validity measures the appropriateness of the content to assess the research objectives (Bordens & Abbott,2011; Heale & Twycross,2015). The research took on the content validity measure for the study.

The variable inputs used for the ANN model were used to observe the inflation rates for the time period identified in the study and the output compared to the actual inflation values for each time period. This ensured that the input variables were valid inflation inputs as per the comparison done with the recorded values. Data validation was also done to ensure that the data used for the study was representative of true inflation inputs and that these inputs were used as model inputs.

3.7 Ethical Issues

Research ethics is critical since it guides the interactions with people, organizations and institutions (Christensen, Johnson, & Turner, 2014). The researcher sought authorization for data collection from the institution and participants by explaining the purpose and importance of study. Privacy and confidentiality was employed to ensure that the data collected from respondents was kept safe, free from interference and protected from unwanted use.

3.8 Model Design, Development and Implementation

The system design approach for developing the model was the CRISP-DM methodology. It incorporates 6 design phases that comprehensively cover the model development process. As used by Cortez (2010), the methodology works well with Artificial Neural Network for predictive purposes and is discussed in detail in chapter 4 and 5. The Figure 3.1 illustrates the development cycle of the methodology

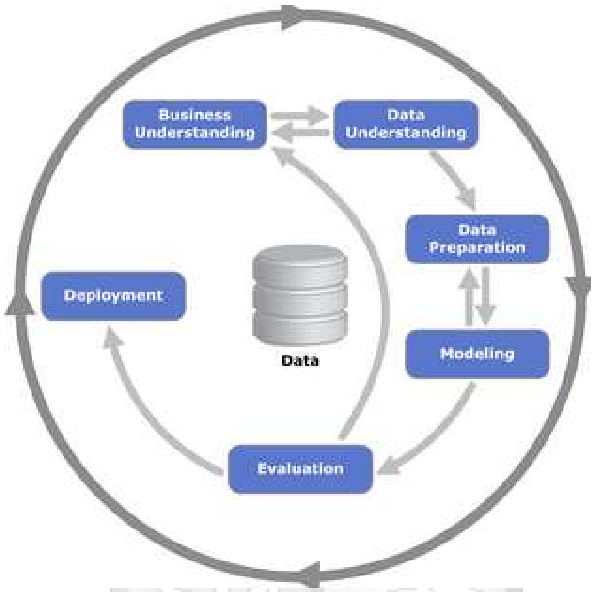

Figure 3.1 CRISP-DM Methodology (Wirth & Hipp, 2000)

Chapter 4: Model Analysis and Design

4.1 Introduction

The process of model analysis and design encompasses defining the model parameters, data understanding and pre-processing for use by the model algorithm. This is captured by defining the model requirements analysis. It is a process that describes the process of studying and developing the business and the user needs to arrive at a definition of the problem domain and model requirements. It is the most critical aspect of the study and determines the goals and functions of the developed model (Dennis, Wixom, & Roth, 2012).

The model was designed to provide an alternative way for forecasting inflation using a non-linear tool ANN in order to fit the nonlinear nature of financial markets (Adhikari & Agrawal, 2013). The model requirements captured are centred on the use of the backpropagation algorithm to facilitate forecasting of inflation values for the stated time period.

4.2 Proposed Model

The model design describes how the model takes in input variables and transforms these independent variables by manipulating them using the model functions into dependent variables being the inflation forecast values.

4.2.1 Defining Model Parameters

This describes the input data used for the development of the proposed model. These are discussed as below.

4.2.1.1 Data Source

Describes the environment from where the data into the model comes from. The input data was received from the financial market reports. These are reports generated and published by the Kenya National Bureau of Statistics (KNBS) and other identified official documentations as stated in the study.

4.2.1.2 Data

Describes the raw facts and Figures captured by the model from the data source for purposes of processing. The numerical data captured included policy variables (exchange rate), food prices, oil prices, M-pesa usage, rainfall and domestic production data for the sample period covered.

4.2.1.3 Algorithm

Describes the pre-defined set rules and procedures that result in an action or a decision by the model. The algorithm involved the steps as described below:

Step 1: Set the initial value of all weights to small real numbers. Input the learning rate and momentum.

Step 2: Pick and load an example from the training data set and specify the desired target output.

Step 3: Run the network normally with the example set.

Step 4: Calculate the error between the resulting output of the ANN and the actual expected solution.

Step 5: Propagate the error value backwards through the network by calculating the weight adjustments.

Step 6: Calculate the new weights and change the weights of all the connections in the direction (positive or negative) that minimizes the error of the output.

Step 7: Repeat the procedure with all training sets until global minimum error is reached.

Step 8: If no remaining training set data, exit else go back to step 3 and repeat the loop.

4.2.2 Model Training Parameters

When training the ANN, the learning rate and training momentum parameters were set. The learning rate defines how much each training example set affects the connection weights. The non-zero training momentum set for the model defines how much each training example affects the connection weights. The non-zero momentum in the connection weight meant that the changes made on the weights on one iteration depended slightly on not only that example, but also on how much it changed in the previous iteration. This way, the delta of change in the connection weights are somewhat smoothened over time (Aamodt, 2010; Rossi, 2012).

4.3 Data Understanding

4.3.1 Model Inputs

Data inputs for the model incorporated multiple variables. This made the ANN a multivariate model for processing. Multivariate models have been found to be more credible as long-term forecasting tools. This is because they take in a more complete picture of the financial environment into account unlike univariate models that consider only one single variable (Aamodt, 2010).

Multivariate approaches fare better than univariate models in differences, and this is especially likely to hold for emerging market countries like Kenya (Aron & Muellbauer, 2015). Dependent variables in the model were defined by the model outputs at each model level. These were reliant on the initial model inputs received by the model algorithm.

4.3.2 Assumptions

The assumptions made were that the variables chosen for forecasting were ultimate and true inputs to inflation and that the data-split periods were optimum periods. Another assumption made was that the external environment like political stability of the country would not affect the outcome of the forecast.

4.4 Data Preparation

This step involved deciding what a good set of training data would be and how to quantify and present the data to the model. More recent data was used for training purposes which was seen to produce better forecasting performance than older training data sets. As Aamodt (2010) explains, this is mainly because the predictive powers of the variables learned through observing a given time period decline over time since the underlying forces driving these variables undergo significant changes over time. The figure 4.1 illustrates the process of data preparation.

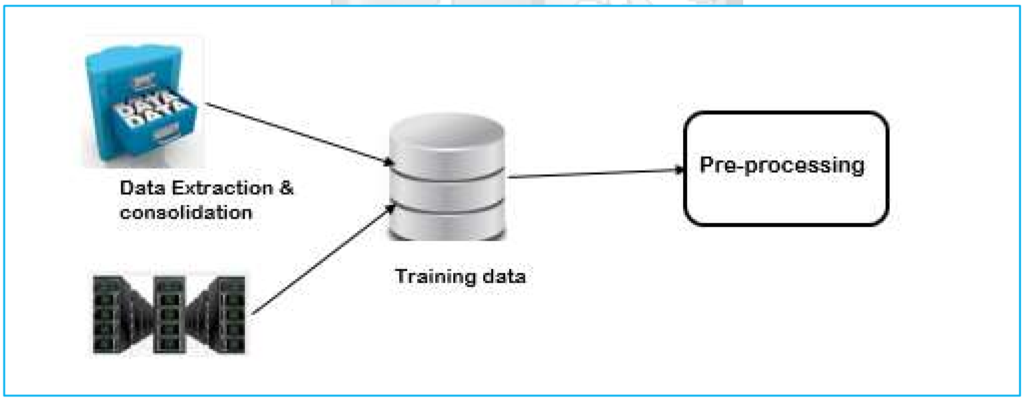

Figure 4.1 Data Preparation

Data preparation also involved data partitioning. This was the process of segmenting the sample data into three sets for purposes of training, testing and validating the model. The training set constituted the largest portion of the sample data and was used by the ANN to learn the patterns present in data. The selection of the training set was done using the 60/20/20 rule where the training data set takes up 60 of the entire sample data set while the remaining 20% is reserved for the testing and validation data sets(Jayalakshmi & Santhakumaran, 2011).

The test data set 20% of the training set, is used to evaluate the generalization ability of the trained network. The sample data was partitioned as follows:-

1. Training set= 60% of total sample size
2. Testing data set= 20% of training set data
3. Validation data set= 20% of total sample size

4.4.1 Data Pre-Processing

This described the process of transforming the input variables from their raw form by normalizing it. Data transformation is crucial in ANNs for achieving a good prediction performance by removing the bias and correlations between the inputs and making them statistically independent (Oancea & Ciucu, 2014).

Data transformation through normalization has been seen to also speed up training time by starting the training process for each feature within the same scale. It is especially useful for modelling application where the inputs are generally on widely different scales (Jayalakshmi & Santhakumaran, 2011).

This process of normalization involved scaling the input data variables for the model between the upper and lower bonds of the activation function (-1, 1). The data normalization was done using the Min-Max normalization represented by the Equation 4.1 (Chittineni & Bhogapathi, 2012). Latha and Thangasamy (2011) suggest that the Min-Max normalization is best suited for where the bounds (maximum and minimum values) of the output values are known beforehand in this case the output value bounds are -1 , 1.

Equation 4.1 Min Max Normalization

$$x' = (x_{\max} - x_{\min}) \times \frac{(x_i - x_{\min})}{(x_{\max} - x_{\min})} + x_{\min} \tag{4.1}$$

The Min-Max normalization has the advantage of preserving exactly all relationships in the data. Where the value of $x_{\max}$-$x_{\min}$ =0 indicates a constant for the input variable which meant that the input value does not provide any information to the ANN and is removed (Jayalakshmi & Santhakumaran, 2011).

4.4 Model Design

The model design describes the organization of the ANN and defined the number of input neurons. The input neurons were used to capture each independent variable. Model design also involved determining the number of ANN hidden layers, the number of neurons in the hidden layer, number of output neurons and the ANN activation function. The proposed model design of the ANN is illustrated in the figure 4.2.

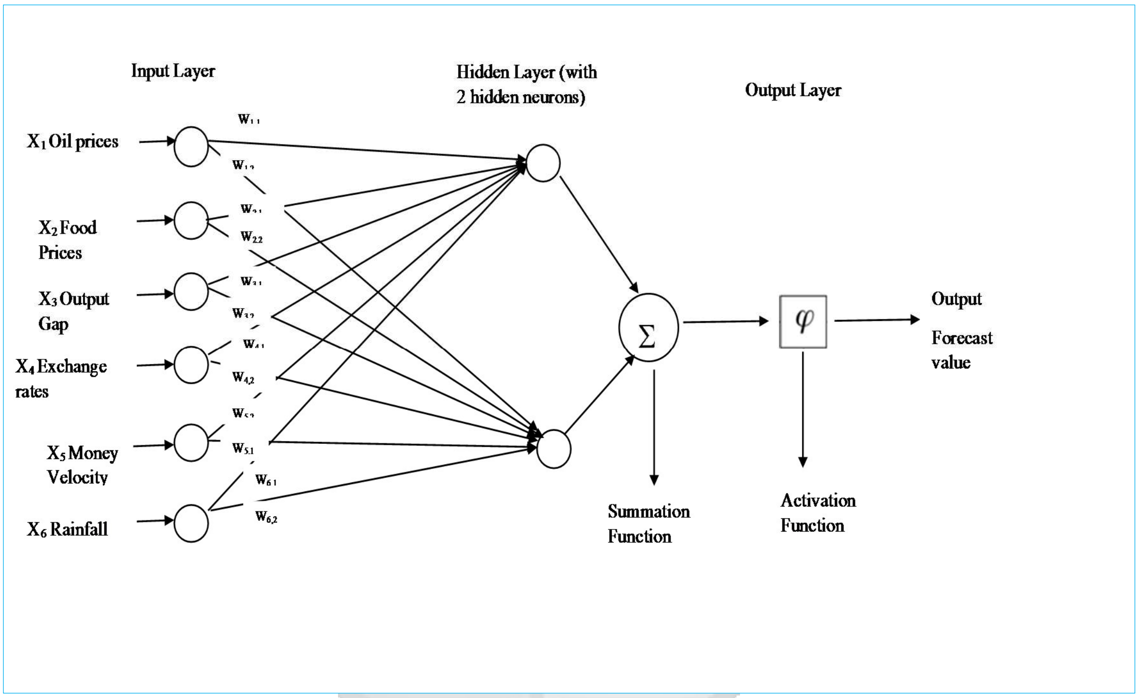

Figure 4.2 Proposed Artificial Neural Network

4.4.1 Number of Hidden Layers

The hidden layers provide the ANN with the capability to learn and generalize. ANN with 1 or 2 hidden layers have been widely used and proven to perform very well. Increasing the number of hidden layers can make an ANN more complex and susceptible to overfitting (Aamodt, 2010). Thus the model design used a single hidden layer for the ANN.

4.4.2 Number of Hidden Neurons

The process of selecting the hidden neurons for the ANN hidden layer is a critical process which is mostly done through a process of trial and error by many researchers (Shin-Ike, 2010).

Sheela and Deepa (2013) suggest that there are multiple ways to define these hidden neurons adopted from an extension of the rule of the thumb which consider different features as per the ANN design. Karsoliya (2012) suggests that the number of hidden layer neurons should be less than twice of the number of neurons in input layer as one of the rules in determining the number of hidden neurons for a BPNN. Thus the study applied this rule to all data with 2, 4, 6 and 8 neurons in the hidden layer as to describe the best structure performance for the model.

4.4.3 Number of Output Neurons

ANNs with multiple output neurons have been found to produce inferior results as compared to ANNs with a single output layer (Sheela & Deepa, 2013). For this reason, the model ANN had a single output neuron layer structure.

4.4.4 Transfer Function

In this study the transfer function used for the model ANN is the hyperbolic tahn activation function. This function is a nonlinear function as is the nature of financial data.

4.4.5 Diagrammatical Representation of the Model

UML diagrams were used to capture the model components and their internal interactions as defined by the specific diagrams. These diagrams help capture and illustrate the process of designing and implementing the model.

4.4.5.1 Use Case Diagram

The use case diagram for the model is illustrated in Figure 4.3.

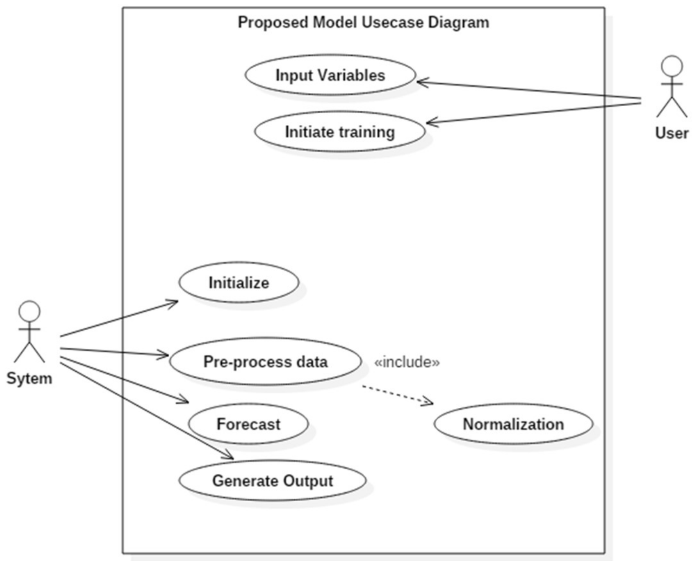

Figure 4.3 Proposed Model Use Case Diagram

The use case diagram captures interactions of the actors with the model within the domain of choice. The user to provide input variables and initiate the process of training the model.

4.4.5.2 Flow Chart Diagram

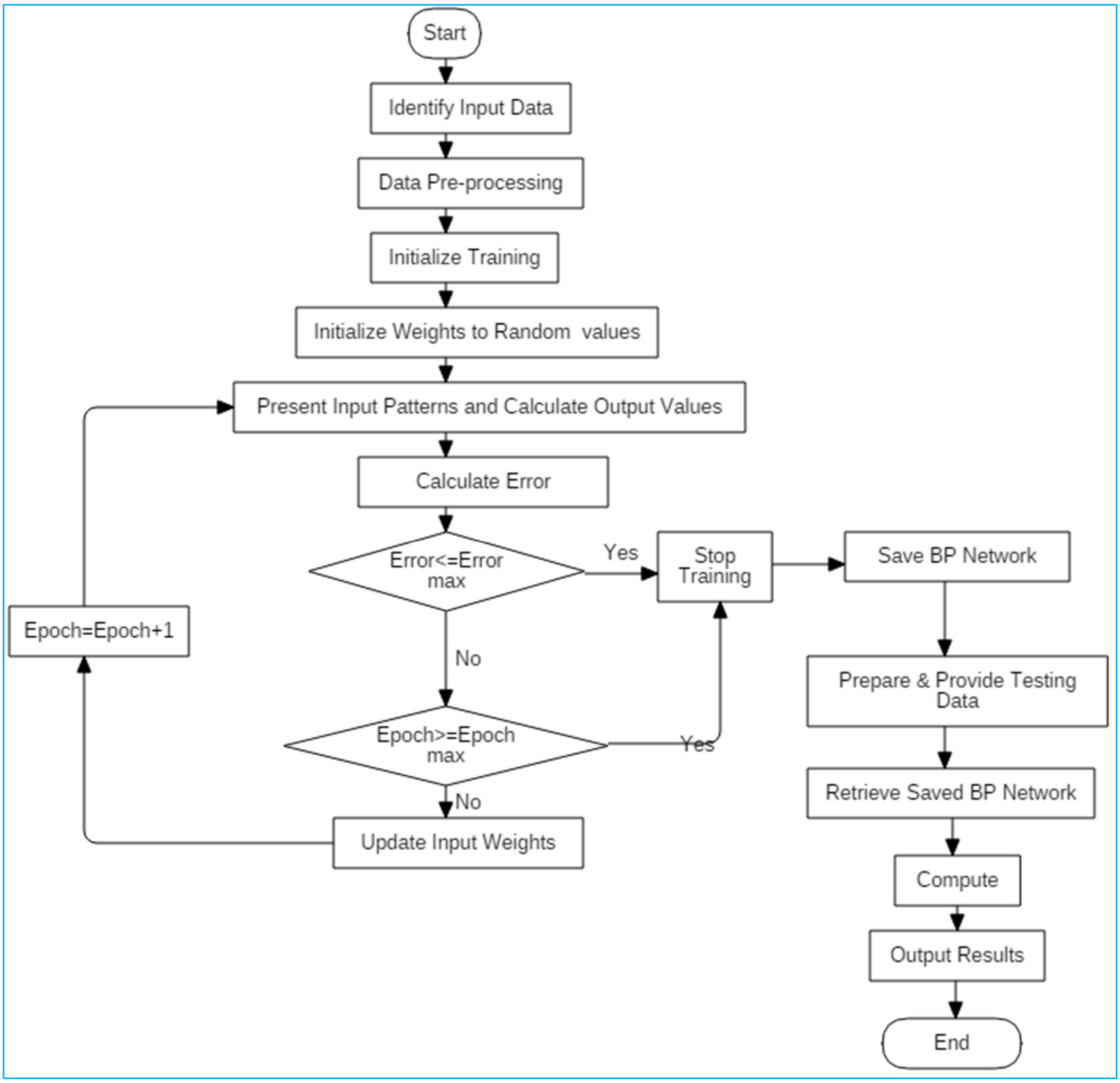

Figure 4.4 Proposed Model Flow Chart

The proposed model flow chart is illustrated by Figure 4.2 and shows the flow of the processes within the ANN. The flow chart diagram illustrates the sequence of steps and decisions taken by the algorithm in evaluating the forecast values.

4.4.5.3 Sequence Diagram

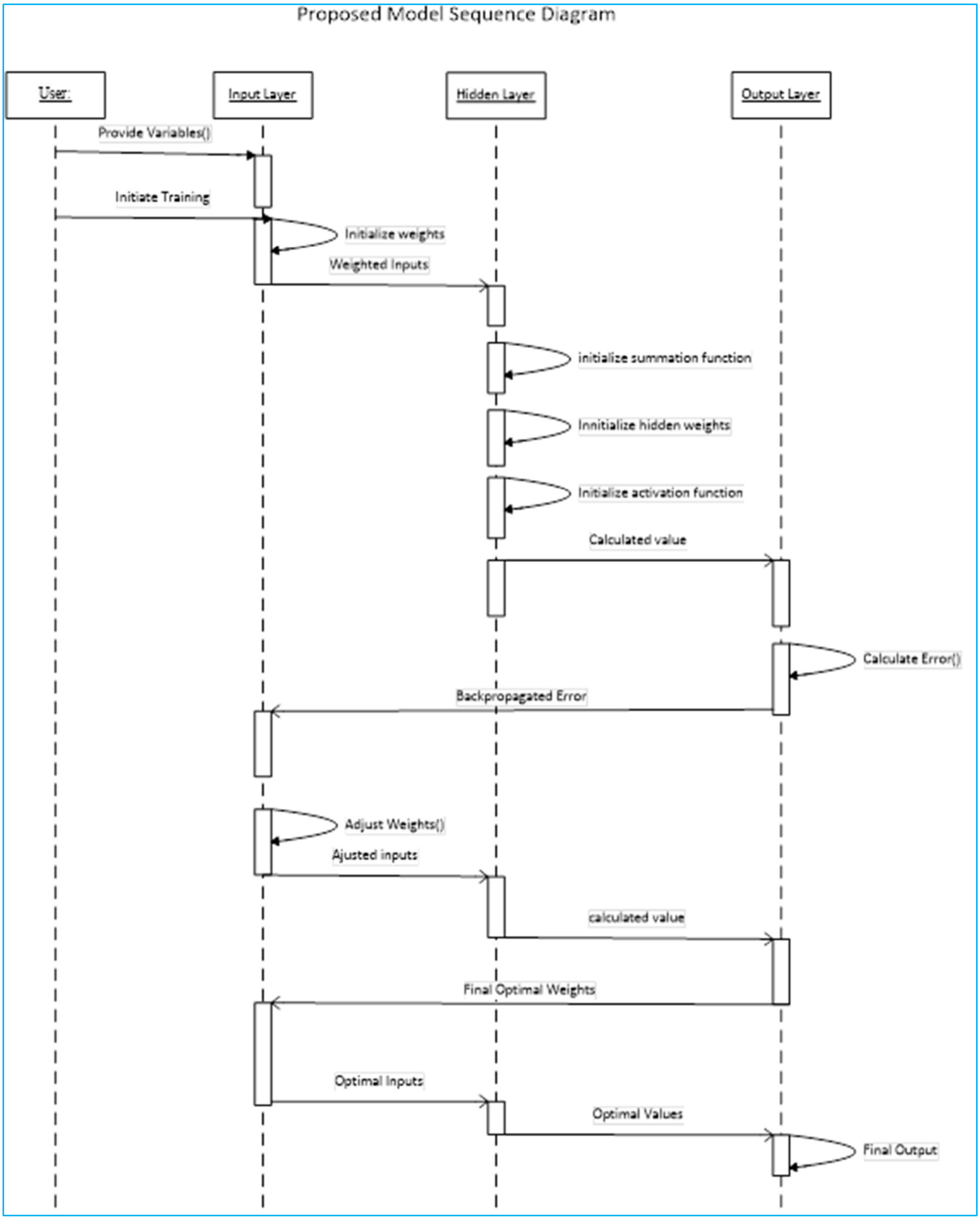

Figure 4.5 Sequence Diagram for Proposed Model

Input processing generates the sum total value of each processing element. This value is passed onto the activation function which compares it to the predetermined threshold for an action which becomes the model output.

4.4.5.4 Class Diagram

The mode class diagram captures a static view of the system indicating objects within the model, their attributes, methods and the interactions between them. The Figure 4.4 illustrates the class diagram for the proposed model.

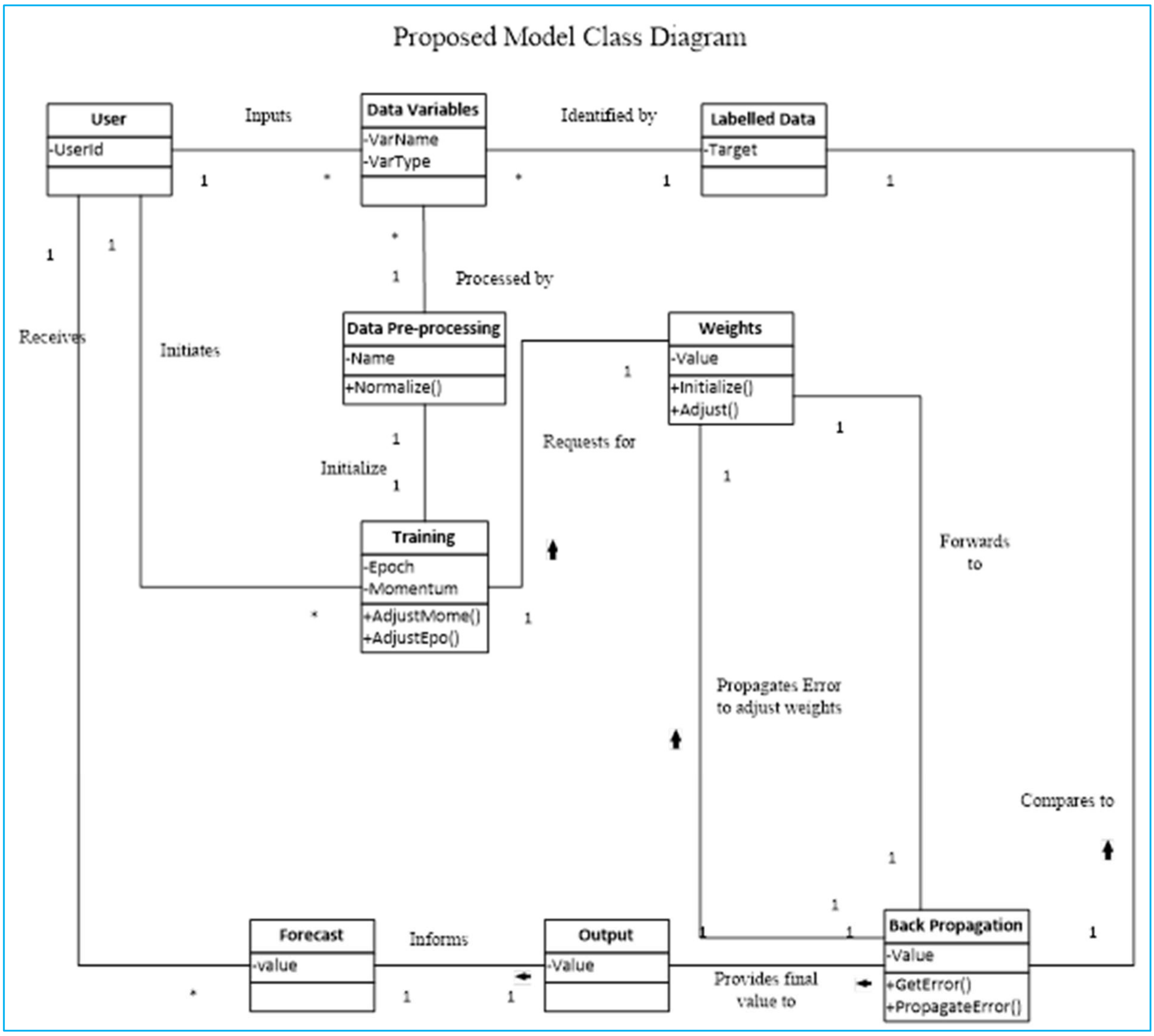

Figure 4.6 Class diagram of the proposed model

Chapter 5: Model Implementation and Validation

5.1 Introduction

The algorithm implementation entails execution of the algorithm and the steps involved in backpropagation neural network. In a BPNN, learning is initiated with the presentation of a training set to the network. The learning algorithm utilized has two stages. The first of these stages is when the training input pattern is presented to the network input layer. The network propagates the input pattern from layer to layer until the output layer results are generated (Somaratna et al., 2010). Then, if the results differ from the expected, an error is calculated, and then transmitted backwards through the network to the input layer. It is during this process that the values for the weights are adjusted to reduce the error encountered. This mechanism is repeated until a terminating condition is achieved (Swain, 2012). For purposes of the study, the model implementation was done as standalone.

5.2 Proposed Model Components

5.2.1 Summation Function

This is the model summation component. The function computes the weighted sum of all the inputs and their corresponding weights. The total input signal sum is the addition of the dot function of each input and the associated weight. The resulting output from this component was a single valued number/vector.

5.2.2 Transfer Function

This component takes in the resulting output from the summation function. The output is then transformed to a working output which is compared to a given threshold. If the sum is greater than the threshold value, the function fires or generates a signal (Aamodt, 2010). If the sum of the input and weight products is less than the threshold, the function generates no signal. Both types of response are significant since they determine the action taken by the model on the inputs received for purposes of identifying the resulting model forecast output values (Suykens, Vandewalle, & Moor, 2012).

5.2.3 Model Outputs

The model outputs describe the output function, error function and the backpropagated value. The output function captures the single output of each processing element. In this case, the output value is equivalent to the activation function's result.

The error function calculates the difference in the values of the output and the desired output. The error value is then propagated back to the learning algorithm as the backpropagated input value. The backpropagated value is used to adjust the input weights before the next model learning cycle.

5.2.4 Learning Function

The Learning function is an important component which modifies the model input weights for each processing element according to the backpropagated value. The value is important as it helps adjust input weights in order to reduce the error rate as close to zero as possible.

5.3 Model Implementation

5.3.1 Initialisation

The weights and threshold values for the network are assigned values that are uniformly distributed over a small range eg: determined using the Haykin approach identified by 5.1

Equation 19 Haykin Approach

$$\left(-\frac{2.4}{F_i},+\frac{2.4}{F_i}\right) \tag{5.1}$$

Where Fi is equivalent to the number of inputs to a neuron i. The initialisation of the weights mentioned is performed on each neuron within the network individually.

5.3.2 Activation

It is at this point that input values from a training set are presented to the networks input layer neurons, and the expected output values that are declared within the set qualified. The networks hidden layer neurons then calculate their outputs.

5.3.3 Update Weights

This is the step in which the weights of the BPNN were updated through the process of propagating backwards the errors related to the output neuron results.

5.3.4 Iteration

Increment the value of P by 1, and return to the second step of the process. This iterative process is conditional upon a terminating condition, if the terminating condition is realised, the training is complete, and the algorithm terminates.

5.4 Training the ANN

The process of training the ANN to learn patterns in the input data involves iteratively presenting the model with labelled data that describes the desired output. The goal of the training process was to find the set of weights that define the global minimum of the error function (Sheela & Deepa, 2013) . The BPNN model used the gradient descent training algorithm.

5.4.1 Number of Training Iterations

Many studies that mention the number of training iterations report convergence from 85 to 5000 iterations (Moustafa, 2011). As Maciel and Ballini (2008) and Sheela & Deepa (2013) suggest that training is affected by many parameters such as the choice of learning rate and momentum values, proprietary improvements to the backpropagation algorithm, among others, which differ between studies and so it is difficult to determine a general value for the maximum number of runs. The training iterations for the model were varied between 500 and 1500 to find the best performance for the model.

5.4.2 Learning Rate

Moustafa (2011) suggests that a small learning rate requires a long training time and is tedious while a high learning rate makes the weights and error function diverge and no learning takes place. Thus the common practice is to start the learning/training rate at a high value and slowly reduce it as the model performance is evaluated. The model learning rate was defined within the NN class and was varied between 0.1 to 0.9 to find the best learning rate for the model. However Many network programs automatically decrease the learning rate as convergence is reached (Sheela & Deepa, 2013).

5.5 Model Validation

Model validation is defined as the process of determining the degree to which the model is a representation of the intended uses in order to quantify confidence in the capability of the model by comparing with experimental data (Thacker, Anderson, Senseny, & Rodriguez, 2005).

The model validation was done to ensure that the model implementation met the model requirements in terms of the algorithm employed and the results obtained to address the right problem of accurate inflation forecasting. The model validation incorporated a review of the model requirements, the performance measure of the model as discussed in chapter 6 of the study and the data used as inputs to the model ANN. An information overview of the process of model validation and verification can be illustrated by the Figure 5.1

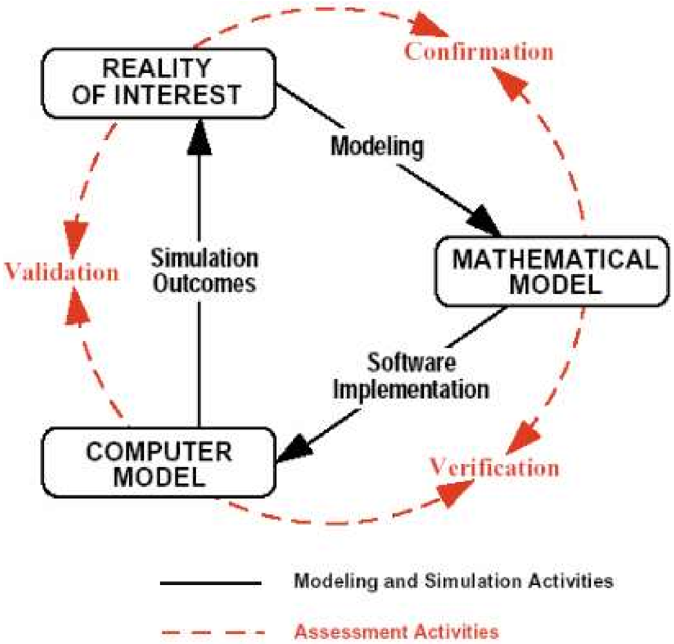

Figure 5.1 Overview of Model Validation (Macal, 2005)

Petty (2014) identifies that the model type, variable data type , processing resources needed and the skills required to execute the model determine the types of model validations that can be done.

5.5.1 Requirements Validation

The validation of the model requirements involved the process of ensuring that the requirements were in line with the expectations of the stakeholders which was to forecast annual inflation values for the time period under study. The Table 5.1 illustrates the requirements validations done for the collected user requirements.

Table 5.1 Requirements Validation Features

Validation Check	Priority	Comments
Does the model capture inflation prediction requirements?	High	All identified requirements captured
Do the requirements conflict?	High	No conflict

5.5.2 Data Validation

Data validation is defined as the process of inspecting the data input variables for accuracy, completeness, removing erroneous data and identifying the missing data (Petty, 2014). To further the validation process of the model data, data specific to inflation was used as discussed in the study captured as true predictors of inflation. The Table 5.2 represents the validation checks done on the data variables. Data validity and data reliability were also used as part of the validation process with the later measured by RMSE and the MAE while the former employed content validity.

Table 5.2 Data Validation Features

Validation Check	Priority	Results
Is the data captured representative of true inflation inputs	High	Captured model variables are true inflation inputs
Are all data values available for the specified time period	Medium	Not all values are available. Missing values replaced with a zero (0)
Does the model use the variables as inputs to forecasting inflation	High	Model utilizes all inputs received for forecasting
Does the model give correct outputs with normalized data	High	The model provides expected outputs

5.5.3 Module Validation

Module validation involved the process of ensuring that the modules within the model work together and do what they are expected to do. It also involved ensuring that the model assumptions made were in line with the real-life environment within which the model was to be operational. The Table 5.3 illustrates the validation checks done on the model modules.

Table 5.3 Module Validation Features

Validation Check	Priority	Results
Does the summation function calculate correct weighted output from received inputs	High	Function calculates the correct values from received model variables
Does the activation function calculate the correct output values?	High	The function calculates the expected output values
Does the back propagation algorithm calculate the correct error values and propagate to input layer?	High	The module makes the expected error calculations and back propagates to input layer
Does the learning function correctly adjust the input weights with back propagated error values?	High	The function updates weights as per the propagated error values.
Does the model logic accurately represent the process of inflation forecasting	High	The mathematical computations of the model capture the process of forecasting inflation values.

Chapter 6: Discussions

6.1 Introduction

This chapter reviews the data variables used for the model and thus discusses the results and findings obtained from the study. The chapter summary discussions were based on the study objectives, data variables and the relevant literature review. These formed the basis for this chapter result interpretation and discussion of results.

The results obtained from the model indicated that the back propagation was an appropriate algorithm that can be implemented in the process of inflation forecasting. The forecasting was done based on inflation variables identified as true inputs to the process of inflation forecasting. The model accuracy performance at 71.4286 % showed that the model is reliable as a tool for inflation forecasting. The simplicity of the implementation also made for easy incorporation and calculation of input integer values easily captured from a simple input format (excel file) used to facilitate the inflation prediction. The study found that the optimum learning rate for the model was 0.5 while the momentum was at 0.9 for the training and 0.7 for the testing and validation data. The optimal number of hidden neurons for the model was found to be 2 neurons implemented for the training, testing and validation. Total iterations varied between the train, test and validate phases.

6.2 The Model Output Data

The data set used to fit the study model consisted of annual changes in the study variables as identified in the Figures following. The data normalization process was done using the Min-Max normalization as specified in chapter 4 to fit within the activation function bounds (-1,1) and to eliminate data bias.

6.2.1 World Oil Prices

The variable data was obtained from the Energy Information Administration Website in dollars per barrel from 2002-2015 (Energy Information Administration, 2016). The data was then normalized without converting it into Ksh since the crude oil is purchased with the dollar. The Figure 6.1 captures the normalized data for oil prices (see Appendix A Table A.3).

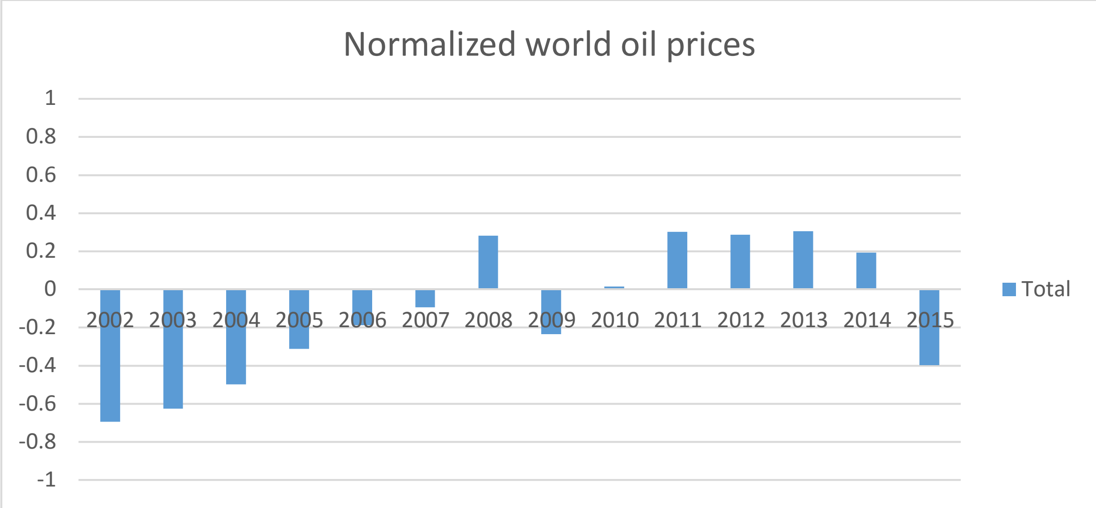

Figure 6.16 Normalized World Oil Prices

6.2.2 Food Prices (Maize Grains)

The variable data was obtained from the Kenya National Bureau of statistics(KNBS) capturing the average consumer price per kg of maize for each year from 2002-2015 (KNBS, 2015). The year 2003 saw the lowest prices for maize at Ksh.17.96 while data for 2015 was missing hence the value of 0. The Figure 6.2 illustrates the normalized maize prices in Kshs (see Appendix A Table A.2)

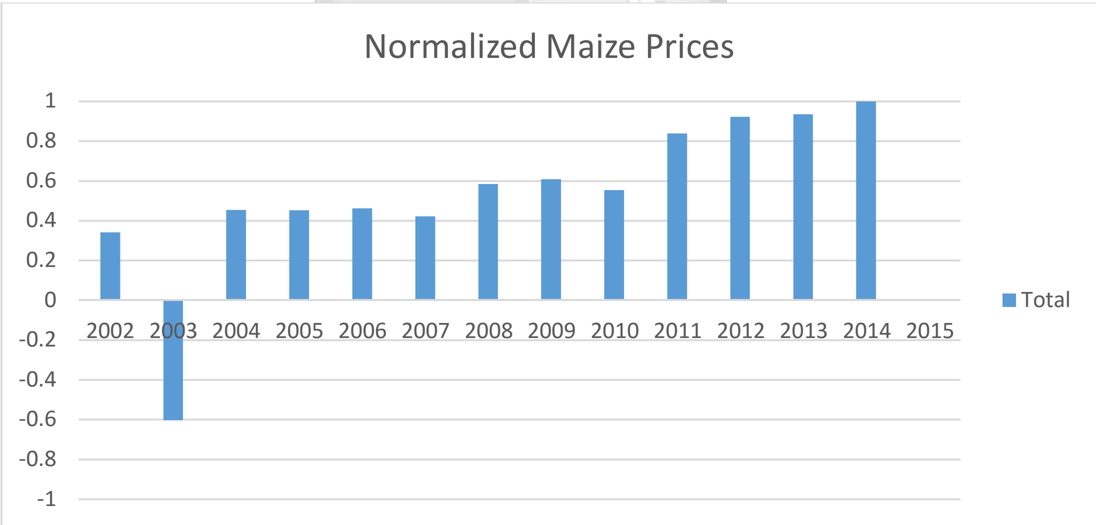

Figure 6.2 Normalized Maize Prices

6.2.3 Output Gap

The variable data was obtained from KNBS, the values however for the period 2012-2015 were missing while 2007 and 2012 had no deficit on the expected and the actual production output hence the values 0 for each year 2015 (KNBS, 2015) . The rest of the years recorded a negative output gap with actual production falling much lower than the expected output. The Figure 6.3 captures the normalized output gap values (see Appendix A Table A.4).

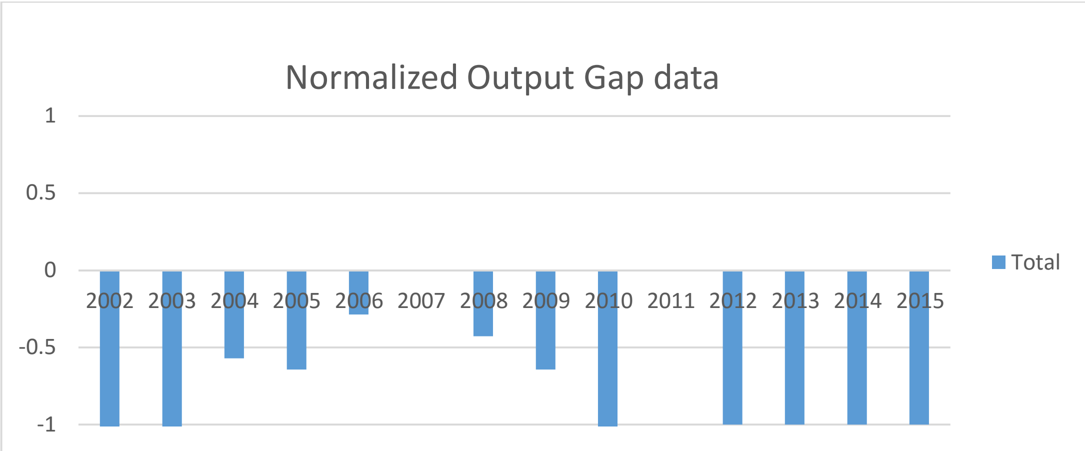

Figure 6.3 Normalized Output Gap Data

6.2.4 Policy Variables (Exchange Rates)

The exchange rates data was obtained from the CBK website for the time period 2002-2015. The exchange rate captured was Ksh against the US dollar as the currency of trade for the country (Central Bank of Kenya, 2016).

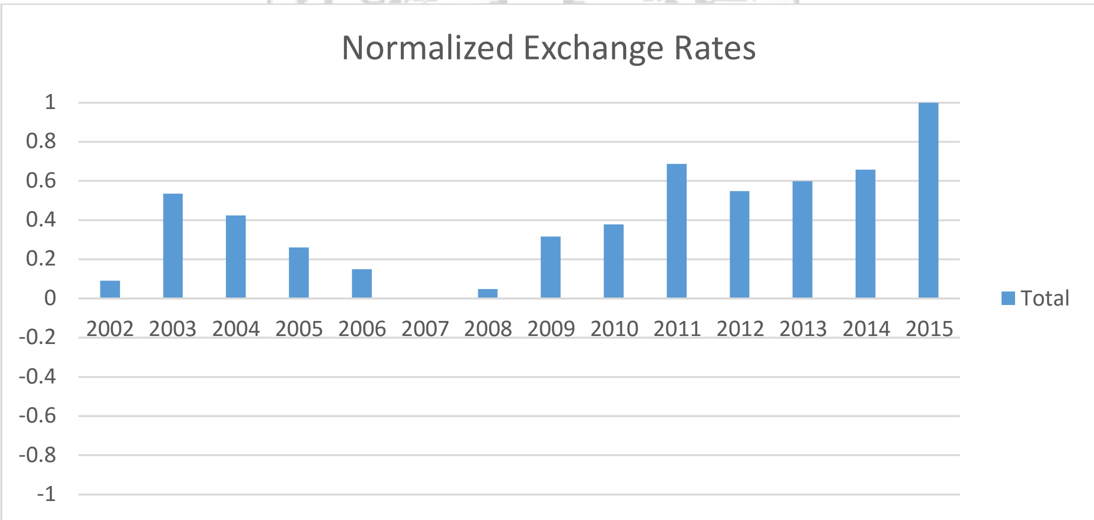

Figure 6.4 Normalized Exchange Rates

This data was averaged per year and normalized the mean value for each time period. The year 2007 had a null value as the min and actual value were the same. The Figure 6.4 captures normalized data for the exchange rates (see Appendix A Table A.1).

6.2.5 Money Velocity (Effects of M-Pesa Use)

The money velocity data was obtained from the KNBS trending money velocity captured as the ratio of GDP to money supply over 3 main indicators which were averaged for each year identified in the study period 2015 (KNBS, 2015). Data for the time period 2013-2015 was missing and thus replaced with a zero (0) for each subsequent value. The Figure 6.4 illustrates the normalized data (see Appendix A Table A.5).

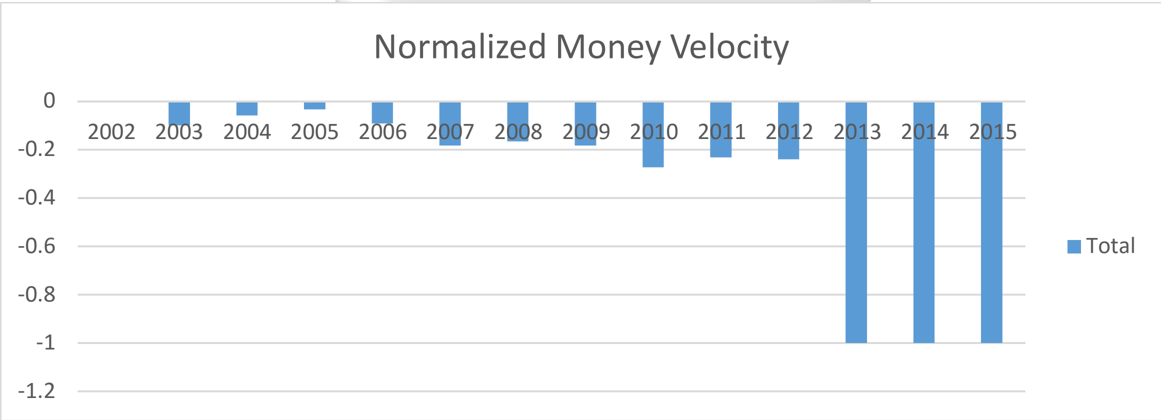

Figure 6.5 Normalized Money Velocity Data

6.2.6 Average Rainfall Data

The rainfall data was obtained from the weather Online website for the period spanning 2002-2015(Weather Online, 2016) .

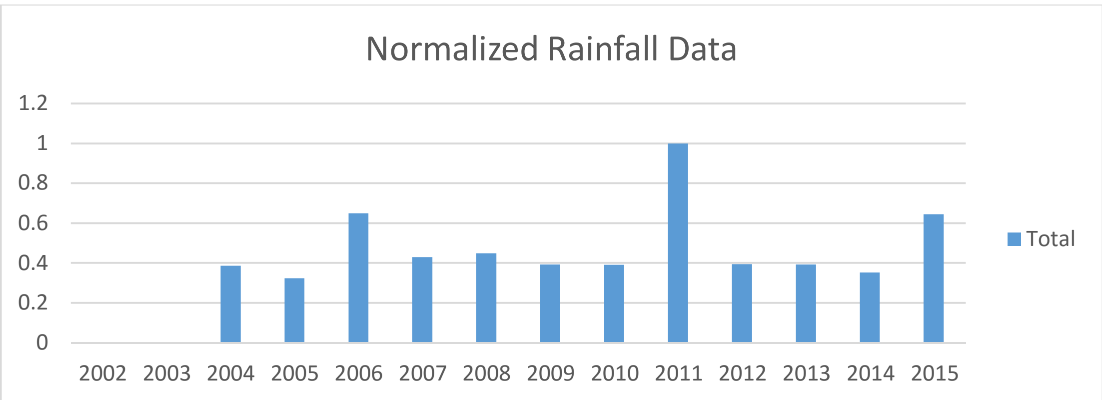

Figure 6.6 Normalized Rainfall Data

The captured rainfall data represented averaged monthly data for each year. The rainfall data for the year 2002 and 2003 was replaced by a null value since actual rainfall data was not available. The Figure 6.6 captures the Normalized data (see Appendix A Table A.6).

6.3 Model Implementation Outputs

6.3.1 Training

The training data outputs were captured as indicated in the Table 6.1. This were obtained after 1500 iterations when the algorithm converged at a Mean Square Error of 0.286 with a learning rate of 0.5 and a momentum rate of 0.9. This was achieved on 10 years of input data as the training set. Figure 6.7 illustrates the train data output

Table 6.1 Training Output Data

Output: 0.9881	Desired output: 1	Error: -0.0119	Iterations
Output: 0.9881	Desired output: 0.6489	Error: 0.3392	
Output: 0.9881	Desired output: 0.1603	Error: 0.8277	
Output: 0.9881	Desired output: 0.916	Error: 0.072	
Output: 0.9881	Desired output: 0.5649	Error: 0.4232	
Output: 1	Desired output: 0.2824	Error: 0.7176	
Output: 0.9881	Desired output: 0.374	Error: 0.614	
Output: 0.9881	Desired output: 0.3435	Error: 0.6445	
Total Mean Square Error: 0.28649323150690287			1500

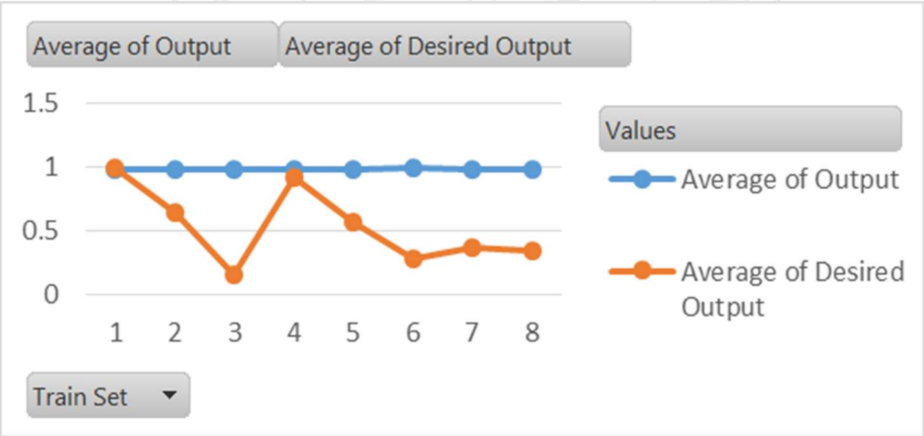

Figure 6.7 Train Data Output

6.3.2 Test Outputs

The test data outputs were captured as indicated in the Table 6.2. This were obtained after 914 iterations when the algorithm converged at a Mean Square Error of 0.007 with a learning rate of 0.5 and a momentum rate of 0.7. This was an improvement on the iteration rates needed to achieve model convergence. These results show that the model learned well and was able to converge easily on fewer iterations. This was achieved on 4 years of input data as the testing set. Figure 6.8 illustrates the test data output

Table 6.2 Testing Data Output

Output: 0.3892	Desired output: 0.4483	Error: -0.0591	Iterations
Output: 0.3524	Desired output: 0.3053	Error: 0.0471	
Output: 0.283	Desired output: 0.1756	Error: 0.1074	
Output: 0.8884	Desired output: 1	Error: -0.1116	
Total Mean Square Error: 0.007426916082736606			914

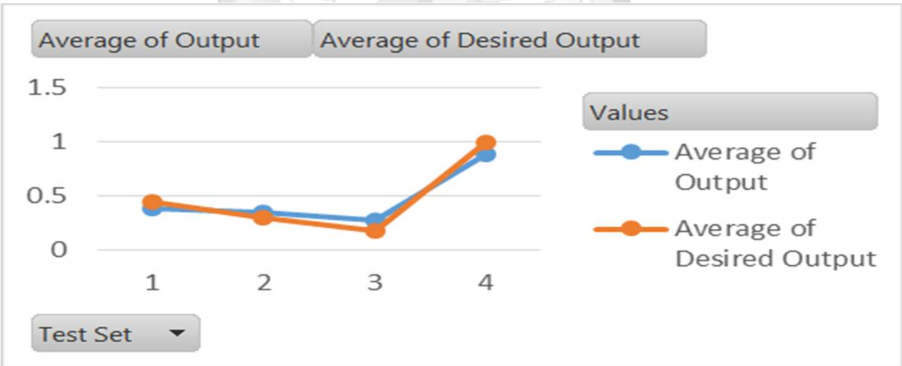

Figure 6.8 Test Data Output

6.3.3 Validation Output

The validation output was captured as indicated in the Table 6.3. This were obtained after 52 iterations when the algorithm converged at a Mean Square Error of 0.13 with a learning rate of 0.5 and a momentum rate of 0.7. This was achieved on 2-3 years of input data as the validation data set. Though the validation error rate was higher than the testing rate, the model converged at fewer iterations with output values much closer to the desired output. Figure 6.9 illustrates the validation data output.

Table 6.3 Validation Output Data

Output: 0.6428	Desired output: 0	Error: 0.6428	Iterations
Output: 0.7722	Desired output: 0.7481	Error: 0.0241	
Output: 0.6336	Desired output: 0.5954	Error: 0.0382	
Total Mean Square Error: 0.13842361323305066			52

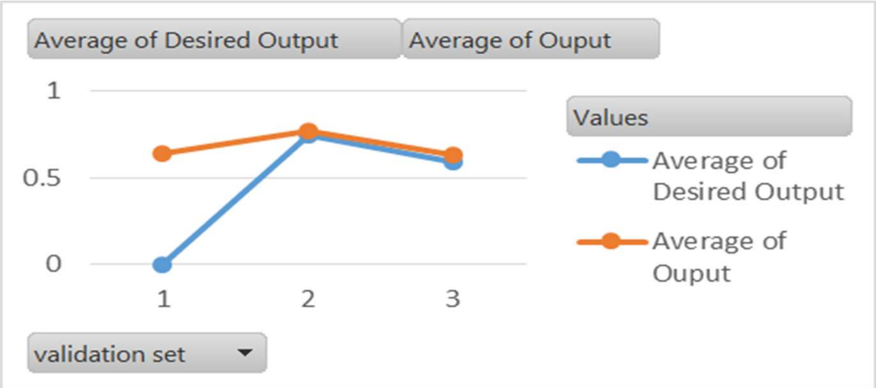

Figure 6.9 Validation Data Output

6.4 Performance Evaluation

6.4.1 Performance Measure Values

The performance measure is used to evaluate the output values for the model in terms of the model accuracy. Performance measures identified for the study were Mean Absolute Error and the Root Mean Squared Error as captured on Table 6.4. The accuracy for the model was 71.4286 % with 10 correctly forecasted outputs and 4 incorrectly forecasted values out of a total sample set of 14 value entries.

Table 6.4 Performance Measure Values

Name	**Value**	**Percentage**
Correctly forecasted Instances	10	71.4286 %
Incorrectly forecasted Instances	4	28.5714 %
Mean absolute error (MAE)	0.4426	
Root mean squared error (RMSE)	0.4694	
Total Number of Instances	14	

Mean Absolute Error (MAE) - Describes the difference in the forecasted value as compared to the actual outputs recorded. The MAE thus informs on the error that can be expected from the forecast model for any predicted value. The MAE 0.4426 represents the error between each forecasted value and the actual recorded inflation value which the study found to be an acceptable value. The lower the MAE the better the prediction accuracy of the model.

Root Mean Squared Error (RMSE) - This performance measure provides the standard deviation of the model output error. The model RMSE was 0.4694 which was acceptable to the study indicating that there were large error values in the model output which were minimized to give reliable results.

6.4.2 Confusion Matrix

Table 6.5 Model Output Confusion Matrix

a b c	<-- classified as
0 0 4 | a	low
1 5 0 | b	medium
0 0 4 | c	high

The confusion matrix Table 6.5 identifies the number of correctly forecasted values as per the different values rated as low high or medium inflation values for purposes of prediction. The first category of low inflation values had a total of 4 instances which were correctly identified. The second medium category had a total of 6 instances with 5 correctly predicted values and 1 incorrectly classified value. The last category had all 4 high inflation predictions correctly predicted.

6.5 Challenges

Availability of data was a main challenge to the process of developing the model. This meant that some missing data items had to be replaced by a null. This was not a true representation of the variables since actual data existed at the time in question but unavailable at the point of carrying out the research. Time constraints also affected the full development of the model since some external variables could not be incorporated into it.

Chapter 7: Conclusions and Recommendations

7.1 Conclusions

Misati and Munene (2015) notes that the government has often failed to predict accurate inflation Figures over the years. This often leads to improper policy development and implementation which negatively impact the economy and as a result impacting the quality of life, investments and ease of doing business in the country. This implies that it is critical to achieve accurate inflation forecasts that was facilitate timely policy implementations and inflation control. According the study Food prices and especially so maize prices have a significant effect on the inflation rate and thus should be considered an essential inflation variable which should be inclusive in any effort made by the CBK to forecast inflation. This ties in with rainfall information which affects food production and thus the final food prices for cereals especially so maize production.

The research findings point out the possibility of achieving accuracy in determining inflation forecasts. This can be done by analysing input variables via the ANN model whose flexibility allows for the variables to be altered as per the changes in inflation data without altering the structure of the model. The study shows that the best way to achieve this is by first pre-processing the input thus transforming the variables from their raw form by normalizing it in order to achieve a good prediction performance by removing the bias and correlations between the inputs and making them statistically independent.

From the study, it was observed that the model was able to forecast inflation values with minimum error and the accuracy level of the output enhanced by varying the number of hidden neurons. This indicated that the backpropagation algorithm used for the Artificial Neural Network forecasting model is an efficient tool in creating a model suitable for forecasting inflation values.

It is critical for the central bank to identify more versatile models of implementing inflation forecasting as is their mandate to keep the inflation rate at an acceptable level (Andrle et al., 2013). An important area of focus for the central bank would be the use of emerging machine learning tools to enhance the process of inflation forecasting by implementing ICT capabilities within the process of inflation targeting. This can help in improving the resulting policy making decisions based on accurate feedback from the predicted values for purposes of planning and intervention measures.

7.2 Recommendations

The following are the recommendations done from the findings of the study

i. The study recommends adoption of the model to forecast inflation for both the short and long term periods.
ii. Inflation forecasting should be viewed as a vital part of the process of developing feasible policy measures for the purposes of curbing inflation rates
iii. More considerations need to be made on the applicability of the model within the larger financial sector and not only restricted to inflation forecasting.

7.3 Suggestions for Future Research

Other research projects can focus on including an expert system within the model to help interpret the resulting outputs from the ANN model. One of the demerits of ANNs is that their outputs can be difficult to interpret by a human user. This would extend the capabilities of the model to also provide worst and best case scenario predictions for the desired time period.This was clear during the result interpretation as some of the resulting outputs were hard to explain the model rationale on its output decision.

The study also recommends that other data variables like political stability and security threats can be included to enhance the predictive capabilities of the model. This would place it in a more real-world setting which would improve on the model accuracy. This is based on the fact that external environments to the model do affect inflation directly or indirectly by influencing the conditions positively/negatively.

The study also suggests linking the model with relevant data sources so that it picks data variables automatically from the model environment for forecasting. This was enhance performance of the model by eliminating human intervention in uploading data variables which would be prone to errors. The automation would help the model do periodic forecasting based on the data received from the linked data sources.

From the research findings the study suggests adopting the same model to other sectors within the financial sector to predict the Consumer price index (CPI) and the Gross domestic product (GDP). This was help in making more use of the model within the same environment since ANNs are versatile and can be used with different variables without altering the structure of the model.

Lastly, the study suggests use of different types of artificial neural networks in order to explore possibilities of finding the most optimal ANN model. Neural Networks like the neural fuzzy networks and self-organizing maps are recommended for future studies.

References

Aamodt, R. (2010). Using artificial neural networks to forecast financial time series. Retrieved from http://brage.bibsys.no/xmlui/handle/11250/252166

Adhikari, R., & Agrawal, R. K. (2013). An Introductory Study on Time Series Modeling and Forecasting. *arXiv Preprint arXiv:1302.6613*. Retrieved from http://arxiv.org/abs/1302.6613

African Development Bank [AfDB]. (2011). Inflation Dynamics in selected East African countries: Ethiopia, Kenya, Tanzania and Uganda.

Alsaiari, N. O. (2012). An expert system for weather prediction based on animal behaviour. In *The International Conference on Informatics and Applications (ICIA2012)* (pp. 77–84). The Society of Digital Information and Wireless Communication. Retrieved from http://sdiwc.net/digital-library/an-expert-system-for-weather-prediction-based-on-animal-behaviour

Andic, S. B., & Ogunc, F. (2015). Variable Selection for Inflation: A Pseudo Out-of-sample Approach. *Research and Monetary Policy Department. ·Istiklal Cad. No: 10, 06100 Ulus, Ankara, Turkey.*

Andrle, M., Berg, A., Morales, R. A., Portillo, R., & Vlcek, J. (2013). Forecasting and Monetary Policy Analysis in Low-Income Countries: Food and non–Food Inflation in Kenya. *IMF Working Paper WP/13/61*.

Andrle, M., Berg, A., Morales, R. A., Portillo, R., & Vlcek, J. (2014). On the Sources of Inflation in Kenya: A Model-Based Approach: South African Journal of Economics. *South African Journal of Economics*, n/a-n/a. http://doi.org/10.1111/saje.12072

Arienda, A., Asana, J. L., & Constantino, T. (2015, February). *Operations management forecasting*. Retrieved from http://www.slideshare.net/Euwinx/operations-management-forecasting

Armantier, O., Bruine de Bruin, W., Potter, S., Topa, G., van der Klaauw, W., & Zafar, B. (2013). Measuring Inflation Expectations. *Annual Review of Economics*, *5*(1), 273–301. http://doi.org/10.1146/annurev-economics-081512-141510

Aron, J., & Muellbauer, J. (2015). Does Mobile Money Cause Inflation? Evidence From Inflation Models For Uganda. *Centre for the Study of African Economies, Oxford*.

Aron, J., Muellbauer, J., Sebudde, R., & others. (2015). *Inflation forecasting models for Uganda: is mobile money relevant?* Centre for the Study of African Economies, University of Oxford. Retrieved from http://www.sbs.ox.ac.uk/sites/default/files/research-projects/mobile-money/inflation-forecasting-16April.pdf

Bahrammirzaee, A. (2010a). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, *19*(8), 1165–1195. http://doi.org/10.1007/s00521-010-0362-z

Bahrammirzaee, A. (2010b). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, *19*(8), 1165–1195. http://doi.org/10.1007/s00521-010-0362-z

Battaglia, M. (2011). Encyclopedia of Survey Research Methods. *SAGE Publications*. Retrieved from https://edge.sagepub.com/system/files/Ch5NonprobabilitySampling.pdf

Benkachcha, S., Benhra, J., & El Hassani, H. (2013). Causal Method and Time Series Forecasting model based on Artificial Neural Network. *International Journal of Computer Applications (0975 – 8887)*, *Volume 75-No. 7*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.402.7437&rep=rep1&type=pdf

Binner, J., Bissoondeeal, R., Elger, T., Gazely, A., & Mullineux, A. (2005). A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro

Divisia. *Applied Economics*, *37*(6), 665–680. http://doi.org/10.1080/0003684052000343679

Bordens, K. S., & Abbott, B. B. (2011). *Research design and methods: a process approach* (8th ed). New York: McGraw-Hill.

Bredart, X. (2014). A “User Friendly” Bankruptcy Prediction Model Using Neural Networks. *Accounting and Finance Research*, *3*(2). http://doi.org/10.5430/afr.v3n2p124

Campbell, D., & Campbell, S. (2008). Introduction to Regression and Data Analysis. *StatLab Workshop Series 2008*.

Chatfield, C. (2013). *The Analysis of Time Series: An Introduction, Sixth Edition*. CRC Press.

Cheng, M. K. C. (2006). *A VAR Analysis of Kenya’s Monetary Policy Transmission Mechanism: How Does the Central Bank’s REPO Rate Affect the Economy?* International Monetary Fund.

Chiang, Y.-M., Cheng, W.-G., & Chang, F.-J. (2012). A hybrid artificial neural network-based agri-economic model for predicting typhoon-induced losses. *Natural Hazards*, *63*(2), 769–787. http://doi.org/10.1007/s11069-012-0188-3

Chittineni, S., & Bhogapathi, R. B. (2012). A Study on the Behavior of a Neural Network for Grouping the Data. *arXiv Preprint arXiv:1203.3838*. Retrieved from http://arxiv.org/abs/1203.3838

Choudhary, M. A., & Haider, A. (2011). Neural network models for inflation forecasting: an appraisal. *Applied Economics, 44:20, 2631-2635,*. http://doi.org/10.1080/00036846.2011.566190

Christensen, L. B., Johnson, B., & Turner, L. A. (2014). *Research methods, design, and analysis* (Twelfth Edition). Boston: Pearson.

Collova, F., & Capaldo, R. (2013, December). *Machine Learning: Introduction to Neural Networks*. Retrieved from http://www.slideshare.net/fcollova/introduction-to-neural-network

Cortez, P. (2010). Data Mining with Neural Networks and Support Vector Machines Using the R/rminer Tool. In P. Perner (Ed.), *Advances in Data Mining. Applications and Theoretical Aspects* (pp. 572–583). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-14400-4_44

Dennis, A., Wixom, B. H., & Roth, R. M. (2012). *Systems analysis and design* (5th ed). Hoboken, NJ: John Wiley.

Devi, C. J., Reddy, B. S. P., Kumar, K. V., Reddy, B. M., & Nayak, N. R. (2012). ANN approach for weather prediction using back propagation. *International Journal of Engineering Trends and Technology*, *3*(1). Retrieved from http://ijettjournal.org/volume-3/issue-1/IJETT-V3I1P204.pdf

Durevall, D., & Ndung'u, N. S. (2001). A dynamic model of inflation of Kenya, 1974-96. *Journal of African Economies*, *10*(1), 92–125.

Durevall, D., & Sjö, B. (2012). *The Dynamics of Inflation in Ethiopia and Kenya*. African Development Bank Group. Retrieved from http://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Working%20Paper%20151%20-%20The%20Dynamics%20of%20Inflation%20in%20Ethiopia%20and%20Kenya.pdf

Elliott, G., & Timmermann, A. (2013). *Handbook of Economic Forecasting*. Newnes.

Energy Information Administration, eia. (2016). U.S. Crude Oil First Purchase Price (Dollars per Barrel). Retrieved March 25, 2016, from https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=F000000__3&f=M

Esquivel, M. (2009). Performance of Artificial Neural Networks in Forecasting Costa Rican Inflation. *Central Bank Researchers Network of the American Continent*.

Eva Ostertagova, O. O. (2011). The simple exponential smoothing model.

Gerstner, W. (2010). Supervised Learning in Neural Networks.

Giacomini, R., & Rossi, B. (2015). Forecasting in Nonstationary Environments: What Works and What Doesn't in Reduced-Form and Structural Models. *Annual Review of Economics*, *7*(1), 207–229. http://doi.org/10.1146/annurev-economics-080614-115338

Gichuki, J., Oduor, J., & Kosimbei, G. K. (2012). The choice of optimal monetary policy instrument for Kenya. Retrieved from http://etd-library.ku.ac.ke/handle/123456789/6752

Gikungu, S. W. (2015). Forecasting Inflation Rate in Kenya Using SARIMA Model. *American Journal of Theoretical and Applied Statistics*, *4*(1), 15. http://doi.org/10.11648/j.ajtas.20150401.13

Gosasang, V., Chandraprakaikul, W., & Kiattisin, S. (2010). An Application of Neural Networks for Forecasting Container Throughput at Bangkok Port. *World Congress on Engineering*, *Vol 1*.

Gupta, S., & Kashyap, S. (2015). Forecasting inflation in G-7 countries: an application of artificial neural network. *Foresight*, *17*(1), 63–73. http://doi.org/10.1108/FS-09-2013-0045

Haider, A., & Hanif, M. N. (2009). Inflation Forecasting In Pakistan Using Artificial Neural Networks. *Pakistan Economic and Social Review*, *Vol. 47*, 123–138.

Hall, O. P. (2002a). Artificial Intelligence Techniques Enhance Business Forecasts - Graziadio Business Review | Graziadio Business Review | Graziadio School of Business and Management | Pepperdine University. Retrieved from https://gbr.pepperdine.edu/2010/08/artificial-intelligence-techniques-enhance-business-forecasts/

Hall, O. P. (2002b). Artificial Intelligence Techniques Enhance Business Forecasts - Graziadio Business Review | Graziadio Business Review | Graziadio School of Business and Management | Pepperdine University. Retrieved from https://gbr.pepperdine.edu/2010/08/artificial-intelligence-techniques-enhance-business-forecasts/

Hansen, P. R., & Timmermann, A. (2012). Choice of sample split in out-of-sample forecast evaluation. Retrieved from http://cadmus.eui.eu/handle/1814/21454

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing*, ebnurs-2015-102129. http://doi.org/10.1136/eb-2015-102129

Hurtado, C., Luis, J., Fregoso, C., & Héctor, J. (2013). Forecasting Mexican inflation using neural networks. In *2013 International Conference on Electronics, Communications and Computing (CONIELECOMP)* (pp. 32–35). http://doi.org/10.1109/CONIELECOMP.2013.6525753

Hyndman, R., & Athanasopoulos, G. (2012). Forecasting: Principles and practice. an online textbook. In *Forecasting: Principles and practice*.

International Monetary Fund, I. (2011). *Kenya: Request for a Three-Year Arrangement Under the Extended Credit Facility-Staff Report; Press Release on the Executive Board Discussion; and Statement by the Executive Director for Kenya*. International Monetary Fund.

Jack, W., & Suri, T. (2011). *Mobile money: the economics of M-PESA*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w16721

Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical Normalization and Back Propagationfor Classification. *International Journal of Computer Theory and Engineering*, *3*(1), 89.

Karlik, B., & Olgac, A. V. (2010). Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks. *Bekir Karlik and A. Vehbi Olgac International Journal of Artificial Intelligence And Expert Systems (IJAE)*, *Volume (1)*(Issue (4)).

Karsoliya, S. (2012). Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International Journal of Engineering Trends and Technology*, *3*(6), 713–717.

Kashid, S., & Kumar, S. (2013). Applications of Artificial Neural Network to Sheet Metal Work - A Review. *American Journal of Intelligent Systems*, *2*(7), 168–176. http://doi.org/10.5923/j.ajis.20120207.03

Kavanagh, S., & Williams, D. (2014). Making the Best Use of Judgmental Forecasting. *Government Finance Review*. Retrieved from http://www.gfoa.org/sites/default/files/GFR61508.pdf

Kiptui, M. (2009). Oil Price Pass-Through into Inflation in Kenya. In *African Econometric Society (AES) Conference Proceedings*. Retrieved from http://africametrics.org/documents/conference09/papers/Kiptui.pdf

KNBS. (2015). Kenya National Bureau of Statistics. Retrieved October 14, 2015, from http://www.knbs.or.ke/

Kumar, P., Ram, B. R., & Hanmanthu, B. (2014). A Machine Learning Approach to Riddle Abused Scraps from OSN User Space. *International Journal of Computer Applications*, *106*(12). Retrieved from http://search.proquest.com/openview/16d8180861eb7b7b98bb2a0ac312568a/1?pq-origsite=gscholar

Latha, L., & Thangasamy, S. (2011). Efficient approach to normalization of multimodal biometric scores. *International Journal of Computer Applications*, *32*(10), 57–64.

Li, Y., & Ma, W. (2010). Applications of Artificial Neural Networks in Financial Economics: A Survey. In *2010 International Symposium on Computational Intelligence and Design (ISCID)* (Vol. 1, pp. 211–214). http://doi.org/10.1109/ISCID.2010.70

Macal, C. M. (2005). Model verification and validation. In *Workshop on" Threat Anticipation: Social Science Methods and Models*. Retrieved from http://jtac.uchicago.edu/conferences/05/resources/V%26V_macal_pres.pdf

Maciel, L. S., & Ballini, R. (2008). Design a neural network for time series financial forecasting: Accuracy and robustness analysis. *Instituto de Economía, Universidade Estadual de*

Campinas, Sao Paulo-Brasil. Retrieved from http://www.cse.unr.edu/~harryt/CS773C/Project/895-1697-1-PB.pdf

Mackey, A., & Gass, S. M. (2013). *Second Language Research: Methodology and Design*. Routledge.

Marcellino, M. (2008). A Linear Benchmark for Forecasting GDP Growth and Inflation? *Journal of Forecasting J. Forecast. 27, 305–340 (2008)*.

Mbiti, I., & Weil, D. N. (2011). *Mobile banking: The impact of M-Pesa in Kenya*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17129

Mbiti, I., & Weil, D. N. (2013). The Home Economics of E-Money: Velocity, Cash Management, and Discount Rates of M-Pesa Users. *The American Economic Review*, *103*(3), 369.

McAdam, P., & McNelis, P. (2005). Forecasting inflation with thick models and neural networks. *Economic Modelling*, *22*(5), 848–867. http://doi.org/10.1016/j.econmod.2005.06.002

McSweeney, C., New, M., & Lizcano, M. (2010). UNDP Climate Change Country Profiles-Kenya. Retrieved from https://open.uct.ac.za/handle/11427/2413

Medsker, L. R. (2012). *Hybrid Neural Network and Expert Systems*. Springer Science & Business Media.

Meyer, B. H., & Pasaogullari, M. (2010). Simple ways to forecast inflation: what works best? *Economic Commentary*, *17*. Retrieved from http://core.ac.uk/download/pdf/6453488.pdf

Misati, R. N., & Munene, O. (2015). Second Round Effects And Pass-Through Of Food Prices To Inflation In Kenya. *International Journal of Food and Agricultural Economics*, *3*(3), 75.

Misati, R. N., Nyamongo, E. M., & Mwangi, I. (2013). Commodity price shocks and inflation in a net oil-importing economy. *OPEC Energy Review*.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.

Moriyama, K., & Naseer, A. (2009). *Forecasting Inflation in Sudan*. International Monetary Fund.

Moustafa, A. A. (2011). Performance Evaluation of Artificial Neural Networks for Spatial Data Analysis. *Contemporary Engineering Sciences, 4*, 149–163.

Nazif, A. Ç., & Karaçuka, M. (2012). A comparative analysis of alternative univariate time series models in forecasting Turkish inflation. *Journal of Business Economics and Management*, *13*(2), 275–293. http://doi.org/10.3846/16111699.2011.620135

Oancea, B., & Ciucu, Şt. C. (2014). Time series forecasting using neural networks. *arXiv Preprint arXiv:1401.1333*. Retrieved from http://arxiv.org/abs/1401.1333

Okasha, M. K., & Yassen, A. A. (2013). The Application of Artificial Neural Networks In Forecasting Economic Time Series. *The International Journal of Statistics and Analysis (IJSA)*, *3*(2), 123–143.

Otuori, O. H. (2013). Influence Of Exchange Rate Determinants On The Performance Of Commercial Bank In Kenya. *European Journal of Management Sciences and Economics*, *Vol.1*(Issue 2).

Özkan, C., & Erbek, F. S. (2003). The Comparison of Activation Functions for Multispectral Landsat TM Image Classification. *Photogrammetric Engineering & Remote Sensing*, *69*(11), 1225–1234. http://doi.org/10.14358/PERS.69.11.1225

Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2010). Forecasting Employee Retention Probability Using Back Propagation Neural Network Algorithm. In *2010 Second International Conference on Machine Learning and Computing (ICMLC)* (pp. 248–251). http://doi.org/10.1109/ICMLC.2010.35

Papadamou, S., & Markopoulos, T. (2012). The Monetary Approach to the Exchange Rate Determination for a "Petrocurrency": The Case of Norwegian Krone. *International Advances in Economic Research*, *18*(3), 299–314. http://doi.org/10.1007/s11294-012-9360-5

Pesaran, M. H., Pick, A., & Timmermann, A. (2011). Variable selection, estimation and inference for multi-period forecasting problems. *Journal of Econometrics*, *164*(1), 173–187. http://doi.org/10.1016/j.jeconom.2011.02.018

Petty, M. D. (2014). Model Verification and Validation Methods. Retrieved from http://staging.uah.edu/images/research/cmsa/pdf/Downloads/AlaSim_2014_Tutorial_47_VV_Methods_v2.pdf

Philip, A. A., Taofiki, A. A., & Bidemi, A. A. (2011). Artificial neural network model for forecasting foreign exchange rate. *World of Computer Science and Information Technology Journal*, *1*(3), 110–118.

Roseline Nyakerario Misati, Esman Morekwa Nyamongo, Lucas Kamau Njoroge, & Sheila Kaminchia. (2012). Feasibility of inflation targeting in an emerging market: evidence from Kenya. *Journal of Financial Economic Policy*, *4*(2), 146–159. http://doi.org/10.1108/17576381211228998

Rossi, B. (2012). Advances in Forecasting Under Instability. *Social Science Research Network Electronic Paper Collection*. Retrieved from http://ssrn.com/abstract=1934348

Sermpinis, G., Stasinakis, C., Theofilatos, K., & Karathanasopoulos, A. (2014). Inflation and Unemployment Forecasting with Genetic Support Vector Regression. *Journal of Forecasting*, *33*(6), 471–487. http://doi.org/10.1002/for.2296

Sheela, K. G., & Deepa, S. N. (2013). Review on Methods to Fix Number of Hidden Neurons in Neural Networks, Review on Methods to Fix Number of Hidden Neurons in Neural Networks. *Mathematical Problems in Engineering, Mathematical Problems in Engineering*, *2013*, *2013*, e425740. http://doi.org/10.1155/2013/425740, 10.1155/2013/425740

Shin-Ike, K. (2010). A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network. In *SICE Annual Conference 2010, Proceedings of* (pp.

238–242). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5603258

Somaratna, P. E., Arunatilaka, S., & Premarathna, L. (2010). Which is better for inflation forecasting? Neural networks or data mining. In *2010 International Conference on Information Society (i-Society)* (pp. 116–121).

Stahl, F., & Jordanov, I. (2012). An overview of the use of neural networks for data mining tasks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(3), 193–208. http://doi.org/10.1002/widm.1052

Stella, A., & Stock, J. H. (2012). A state-dependent model for inflation forecasting. *FRB International Finance Discussion Paper*, (1062). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2184946

Stock, J. H., & Watson, M. W. (2008a). Forecasting inflation. *Journal of Monetary Economics*, *44*(2), 293–335. http://doi.org/10.1016/S0304-3932(99)00027-6

Stock, J. H., & Watson, M. W. (2008b). *Phillips curve inflation forecasts*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w14322

Stock, J. H., & Watson, M. W. (2010). *Modeling inflation after the crisis*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w16488

Stock, J. H., & Watson, M. W. (2011). Dynamic factor models. *Oxford Handbook of Economic Forecasting*, *1*, 35–59.

Subramaniam, A. (2009, December). *Forecasting Techniques*. Retrieved from http://www.slideshare.net/guest865c0e0c/forecasting-techniques

Suykens, J. A. K., Vandewalle, J. P. L., & Moor, B. L. de. (2012). *Artificial Neural Networks for Modelling and Control of Non-Linear Systems*. Springer Science & Business Media.

Swain, M. (2012). An Approach for IRIS Plant Classification Using Neural Network. *International Journal on Soft Computing*, *3*(1), 79–89. http://doi.org/10.5121/ijsc.2012.3107

Tang, L., Dai, W., Yu, L., & Wang, S. (2015). A Novel CEEMD-Based EELM Ensemble Learning Paradigm for Crude Oil Price Forecasting. *International Journal of Information Technology & Decision Making*, *14*(1), 141–169. http://doi.org/10.1142/S0219622015400015

Thacker, B. H., Anderson, M. C., Senseny, P. E., & Rodriguez, E. A. (2005). The role of nondeterminism in model verification and validation. *International Journal of Materials and Product Technology*, *25*(1–3), 144–163.

Tsai, M.-T., & Chen, C.-H. (2010). A forecasting system of electric price using the refined Back propagation Neural Network. In *2010 International Conference on Power System Technology (POWERCON)* (pp. 1–6). http://doi.org/10.1109/POWERCON.2010.5666458

van der Maas, J. (2014). Forecasting inflation using time-varying Bayesian model averaging: Forecasting inflation using time-varying BMA. *Statistica Neerlandica*, *68*(3), 149–182. http://doi.org/10.1111/stan.12027

van Wyk, B. (2012). Research design and methods Part I. Retrieved from http://www.uwc.ac.za/Students/Postgraduate/Documents/Research_and_Design_I.pdf

Vizureanu, P. (2010). *Expert systems*. Rijek, Crotia: InTech. Retrieved from http://www.intechweb.org/books/show/title/expert-systems

Weather Online. (2016). Precipitation - Climate Robot Kenya. Retrieved April 1, 2016, from http://www.weatheronline.co.uk/weather/maps/city?WMO=63740&FMM=1&LMM=12&FYY=2002&LYY=2015&LANG=en&R=0&LEVEL=162&MOD=MOA&ART=PRE&OFFSET=08&NOREGION=0

Youn, H., & Gu, Z. (2010). Predict US restaurant fi rm failures: The artifi cial neural network model versus logistic regression model. *School of Merchandising and Hospitality Management, University of North Texas, Vol. 10*, *3*, 171–187.

Appendix A: Raw Data

Table A.1: Exchange Rates

This captures the data collected and normalized for exchange rates per year as from 2002-2015.

year	Normalized
2002	0.091854461
2003	0.535957916
2004	0.423518528
2005	0.260202783
2006	0.150305701
2007	0
2008	0.049354658
2009	0.317168103
2010	0.378902922
2011	0.687478047

2012	0.547937491
2013	0.599435388
2014	0.657186188
2015	1

Table A.2: Food Prices (Maize Prices)

This captures the data collected and normalized for food prices per year as from 2002-2015

year	Normalized
2002	0.341744135
2003	-0.602478973
2004	0.45396193
2005	0.452412572
2006	0.462594068
2007	0.421425409
2008	0.585436034
2009	0.609340416
2010	0.554448871
2011	0.839309429
2012	0.921868083
Year	Normalized
2013	0.93625498
2014	1
2015	0

Table A.3: Oil Prices

This captures the data collected and normalized for oil prices per year as from 2002-2015

year	Normalized
2002	-0.693815766
2003	-0.625170083
2004	-0.498491927
2005	-0.312443306
2006	-0.188418451

2007	-0.094373639
2008	0.282009706
2009	-0.233751361
2010	0.015659017
2011	0.302056876
2012	0.287565766
2013	0.306184234
2014	0.193498276
2015	-0.397122188

Table A.4: Output Gap

This captures the data collected and normalized for the output gap per year as from 2002-2015

year	Normalized
2002	-1.64285714
2003	-1.85714286
2004	-0.57142857
2005	-0.64285714
2006	-0.28571429
Year	Normalized
2007	0
2008	-0.42857143
2009	-0.64285714
2010	-1.14285714
2011	0
2012	-1
2013	-1
2014	-1
2015	-1

Table A.5: Money Velocity (Effects of using M-pesa)

This captures the data collected and normalized for annual money velocity per year as from 2002-2015

year	Normalized
2002	0
2003	-0.09917355
2004	-0.05785124
2005	-0.03305785
2006	-0.09090909
2007	-0.18181818
2008	-0.16528926
2009	-0.18181818
2010	-0.27272727
2011	-0.23140496
2012	-0.23966942
2013	-1
2014	-1
2015	-1

Table A.6: Rainfall Data

This captures the data collected and normalized for the annual average rainfall per year as from 2002-2015

year	Normalized
2002	0
2003	0
2004	0.386763
2005	0.323677
2006	0.649891
2007	0.42976
2008	0.448298
2009	0.393177
2010	0.391344
2011	1
2012	0.394516
2013	0.392613
2014	0.351801
2015	0.644252

Appendix B: Originality Turnitin Report

Turnitin Originality Report

Artificial Neural Network Model for Inflation Forecasting in Kenya by Carolyn Wanja Mwangi

From MSc IT Thesis Proposal (MSc. IT Thesis Proposal 2015)

- Processed on 31-Mar-2016 12:55 AM EAT
- ID: 652121032
- Word Count: 18058

Similarity Index
19%
Similarity by Source

Internet Sources:
13%
Publications:
11%
Student Papers:
9%

sources:

1 1% match (publications)
Frederic Stahl. "An overview of the use of neural networks for data mining tasks", Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, 05/2012

2 1% match (Internet from 24-Feb-2014)
http://www.cse.unr.edu/~harryt/CS773C/Project/895-1697-1-PB.pdf

3 1% match (Internet from 27-May-2011)

Appendix C: Data Collection Reference Letter

Strathmore
UNIVERSITY

FACULTY OF INFORMATION TECHNOLOGY

Our Ref.: FIT/MSIT/RL/16/43

14th March, 2016

To whom it may Concern:

<u>Re: Mwangi Carolyn Wanja - 085082</u>

This is to confirm that the above named is a student at Strathmore University pursuing *Master of Science in Information Systems (MSc IT)* since May 2014.

Carolyn is a research scholar who is currently in her 2nd (final year) of study and is doing a research pertaining her masters degree which is entitled: **Forecasting Model Based on Artificial Neural Network for Inflation Prediction in Kenya**

This research being a mandatory requirement towards successful completion of her studies, it would be great if you accord Carolyn the necessary support that she may need from your organization to enable her complete this task.

Any assistance accorded her shall be highly appreciated.

In case you would wish to clarify any issues with us, please feel free to do so.

Yours faithfully,

Danny Nyatuka (Mr.)
Administrator, Faculty of Information Technology
dnyatuka@strathmore.edu

Ole Sangale Rd, Madaraka Estate. PO Box 59857-00200, Nairobi, Kenya. Tel +254 (0)703 034000
+254 (0)703 034200 Email info@strathmore.edu www.strathmore.edu