PRINT ISSN 1119-8362
Electronic ISSN 2659-1499

Full-text Available Online at
https://www.ajol.info/index.php/jasem
https://www.bioline.org.br/ja

J. Appl. Sci. Environ. Manage.
Vol. 28 (6) 1913-1925 June 2024

Application of K-Nearest Neighbours and Long-Short-Term Memory Models using Hidden Markov Model to Predict Inflation Rate and Transition Patterns in Nigeria

***[1]NKEMNOLE, EB; [2]WULU, JT; [1]OSUBU, I**

**[1]Department of Statistics, University Lagos, [2]Departement of Mathematics and Statistics, University of Maryland Global Campus, Adelphi USA*

**Corresponding Author Email: enkemnole@unilag.edu.ng*
**ORCID: https//orcid.org/0000-0001-7521-2858*
**Tel: +2347033835855*

Co-Author Email: john.wulu@faculty.umgc.edu; idowuosub@gmail.com

ABSTRACT: It fluctuates in most countries in the world regardless of whether the countries are been developed, developing or underdeveloped. The widespread effects of inflation have a significant impact on most countries including Nigeria because it is a global phenomenon influenced by variety of factors such as economic growth and monetary policy. This paper proposes the application of K-Nearest Neighbours (KNN) and Long Short-Term Memory (LSTM) models using Hidden Markov Model (HMM) to predict the inflation rate and its transition patterns in Nigeria with Secondary data collected from National Bureau of Statistics (NBS) official website. Empirical analysis revealed that GDP per capita show a significant influence in inflation rate and contributes to inflation forecasting in Nigeria.

Keywords: I DOI: https://dx.doi.org/10.4314/jasem.v28i6.33

Cite this Article as: NKEMNOLE, E. B; WULU, J. T; OSUBU, I. (2024). Application of K-Nearest Neighbours and Long-Short-Term Memory Models using Hidden Markov Model to Predict Inflation Rate and Transition Patterns in Nigeria. *J. Appl. Sci. Environ. Manage.* 28 (6) 1913-1925

Dates: Received: 20 April 2024; Revised: 15 May 2024; Accepted: 18 June 2024 Published: 27 June 2024

Keywords: Inflation; Modelling; Forecasting; Hidden Markov Models,

Inflation is a phenomenon where the prices of goods and services increase over a certain period, resulting in a decline in the purchasing power of customers. The World Bank's global database of inflation covers up to 209 countries over the period 1970-2022 and includes six measures of inflation in three frequencies (annual, quarterly, and monthly). The Consumer Price Index (CPI) is a key barometer of inflation, measuring how the prices of anything from fruits and vegetables to haircuts and concert tickets are changing across the economy. Inflation has been a significant issue in Nigeria, with the annual inflation rate reaching an 18-year high of 25.8% in August 2023, due to various factors such as the Nigerian Naira (NGN) depreciation, higher fuel and food prices, logistics costs, and money supply growth. The Central Bank of Nigeria has been taking measures to address inflation, including raising interest rates, but the impact of these measures remains to be seen. Several studies have been done using HMM for various real-life applications (Aron and Muellbauer, (2012), Altug and Cakmakli (2016)). Analyzing the inflation rate for a specific period of time using machine learning with HMM is a novel approach. In this paper, we develop a machine learning technique with Hidden Markov Model (HMM) to analyze inflation scenario from a given data set in Nigeria for predicting inflation rates. The rest of this paper is structured as follows: Section 2 discusses related work. In Section 3, the methodology is presented describing how machine learning models can be combined with HMM. Section 4 presents results. The appropriate parameters of machine learning models with HMM were explored and the inflation rate with respect to economic indicator is

**Corresponding Author Email: enkemnole@unilag.edu.ng*
**ORCID: https//orcid.org/0000-0001-7521-2858*
**Tel: +2347033835855*

predicted. Conclusion is presented in Section 5. A consistent and ongoing increase in the average price of goods and services is referred to as inflation. On the other hand, inflation is a persistent and ongoing decline in the value of money. In general, a few hypotheses around the idea of inflation have been established. According to the demand-pull paradigm, inflation occurs when overall demand for goods and services exceeds overall supply, and this excess demand cannot be satisfied by depleting current stockpiles, shifting supplies from the export to the domestic market, boosting imports, or delaying demand. According to the cost-push theory, rising manufacturing costs lead to higher pricing. According to this idea, rising salaries lead prices of products and services to increase. The Classical to Neo-structuralist expositions of orthodox economists embody the theoretical foundations of inflation drivers. Prominent monetarists such as Brunner and Meltzer (1976); Friedman (1956, 1970); and Parking (1975) proposed that the rate of inflation and its acceleration can be explained by changes in the money supply and growth rate. The money supply is therefore the primary cause of inflation. Keynesians believed that excessive spending in relation to the supply of goods and services at a period of full employment was the cause of inflation. In other words, inflation is essentially caused by a money supply that exceeds potential output (Javed *et al*. 2010). Structuralists came to believe that imbalances in an economy, particularly in emerging nations, are what primarily cause inflation. These imbalances include the development of monopolistic and oligopolistic market structures, as well as infrastructure bottlenecks.

Researchers including Altug and Cakmakli, (2016) ; Aron and Muellbauer, (2012), Medel *et al*. (2016), Mandalinci, (2017); Pincheira and Medel (2015) ; Balcilar *et al*. (2015) and Chen *et al*. (2014) were forced to propose means of maintaining stable, economic and monetary policies as solutions to the aftermath of the high inflation in nations like Canada, China, France, Germany, Japan, Korea, UK, USA, Australia, Chile, New Zealand, South Africa, etc. Research on the inflation rate in Nigeria and around the world in general has been done by Popoola *et al*. (2017), Mustapha and Kubalu (2016), Olajide *et al*. (2012), Kelikume and Salami (2014), Okafor and Shaibu (2013), Adebiyi *et al.* (2010), and Etuk *et al.* (2012) . A variety of techniques have been used to model the dynamics, causes, and impact of the rate of inflation on the Nigerian economy. Olubusoye and Oyaromade (2008) investigated the primary factor influencing Nigerian inflation. The authors used data from 1970 to 2003 and the Error Correction Mechanism. The study's conclusions demonstrate that the prices of petroleum products, anticipated inflation, and exchange rates have a major impact on Nigeria's inflation tendency. Imimole and Enoma (2011) conducted study on how Nigeria's inflationary process is affected by currency rate depreciation. The Autoregressive Distributed Lag (ARDL) model was utilized by the author with data spanning from 1986 to 2008. Their findings indicated that the money supply, inflation inertia, real gross domestic product (GDP), and exchange rate depreciation all contributed to Nigeria's inflation rate. John and Patrick (2016) used data from January 2000 to June 2015 to model Nigeria's inflation rate. They suggested using the ARIMA (0, 1, 0) x (0, 1, 1) model to predict inflation in Nigeria.

Inam (2017) used the VAR model to model inflation in Nigeria. The analysis included data from 1990 to 2012 on changes in real output and import prices, as well as information on the money supply, inflation rate, fiscal deficit, interest rate, and real exchange rate. Important conclusions showed that the past inflation lag value has a substantial impact on both the present and future inflation rates in Nigeria. Other studies that modelled the dynamics, determinants, and effect of the inflation rate inflation rate in other continents have been employed using HMM. Almosova and Andreson (2023) used nonlinear inflation forecasting with recurrent neural networks. The authors found that a combination of a Markov Chain and a neural network was able to outperform other inflation prediction models, including traditional statistical models and machine learning models that did not use Markov Chains. Papadamou, (2019) focused on inflation forecasting in emerging market economies using a Markov-switching dynamic factor model. The study sheds light on the specific challenges and factors influencing inflation in regions. Hossain, *et al*. (2012) used monthly inflation rates to examine inflation throughout various time periods. They divided the continuous inflation data into classes and transformed it into discrete data. In order to estimate inflation rates, three hidden states are selected. The HMM is then trained for various time periods, then tested using inflation data from 2010 to 2011 in order to identify any patterns in past years' inflation behavior.

Inge (2013) built the HMM using inflation data from Sweden between 1831 and 2012. He used the annual percentage changes in the Consumer Price Index (CPI) as a gauge of inflation, the EM-algorithm to estimate parameters, and AIC and BIC to identify the best HMM. Jochman (2010) looked for structural discontinuities in the U.S. inflation data by applying the Infinite Hidden Markov Model (IHMM). Song (2011) modeled the U.S. inflation data using an IHMM as well. He introduced the sticky double hierarchical Dirichlet process hidden Markov model (SDHDP-HMM) in his research and compared it to other alternative regime switching and structural change models that are currently in use. This research was conducted in conjunction to Jochman's (2010) research. According to the study, SDHDP-HMM produces more accurate forecasts than structural break and regime switching models. Machine learning with Markov models has the potential tool for predicting inflation rate for a given data set in Nigeria. In this study the K-Nearest Neighbours (KNN) and Long Short-Term Memory (LSTM) models with HMM have been developed to predict the inflation rate and its transition patterns. Hence, this paper proposes the application of K-Nearest Neighbours (KNN) and Long Short-Term Memory (LSTM) models using Hidden Markov Model (HMM) to predict the inflation rate and its transition patterns in Nigeria

MATERIALS AND METHODS

Data Collection: The data used in this was obtained from the Nigeria's National Bureau of Statistics (NBS) website, Macrotrends and the database website of Central Bank of Nigeria statistics on inflation rate and the economic indicators. The aspects investigated were the Nigeria Annual gross income known as GDP and the annual average standard of living which is known as Gross Domestic Product per capita from year 1960 to 2022.

Data were analyzed using machine learning models with Hidden Markov Model. In addition, tables, charts, graphs, correlation analysis and trend analysis were used to explore the moving pattern of the inflation rate.

Methodology: Machine learning models that incorporate Markov chains are known as Markov Chain Monte Carlo (MCMC) methods. These techniques are used in a wide range of applications, including statistical modelling, Bayesian inference, and simulations. This paper presents an overview of how machine learning models can be combined with HMM. Some of the supervised machine learning models are Long Short-Term Memory (LSTM), Neural Networks, Decision trees, Support Vector Machines (SVMs), K-nearest neighbors (KNN) and random forest.

Long Short-Term Memory (LSTM): A theoretical mathematical representation of a Long Short-Term Memory (LSTM) cell can be complex due to the many components and operations involved. Here is a representation of an LSTM cell. The LSTM architecture was introduced by Hochreiter (1997). Their work significantly contributed to the development of LSTMs, which have become a fundamental component of many sequential data processing tasks in machine learning and artificial intelligence. The mathematical expression for an LSTM unit is as follows:

$$f_t = \boldsymbol{\sigma}(W_f[h_{t-1}, x_t] + b_f \quad (1)$$

$$i_t = \boldsymbol{\sigma}(W_i[h_{t-1}, x_t] + b_i \quad (2)$$

$$o_t = \boldsymbol{\sigma}(W_o[h_{t-1}, x_t] + b_o \quad (3)$$

$$c_t = f_t * c_{t-1} + i_t * \tanh(W_c[h_{t-1}, x_t] + b_c \quad (4)$$

$$h_t = o_t * \tanh(c_t) \quad (5)$$

Where: f_t is the forget gate ; i_t is the input gate;
o_t is the output gate ; c_t is the cell state
h_t is the hidden state;
x_t is the input vector at time t;
W_f, W_i, W_o *and* W_c are weight matrices.
b_f, b_i, b_o *and* b_c are bias vectors;
h_{t-1} is the hidden state at time $t - 1$

Assumptions:
h_t represents the hidden state of the LSTM cell at time t.
$c_t$represents the cell state or memory of the LSTM cell at time t.
x_t represents the input at time t.
W and b represent weight matrices and bias vectors specific to different LSTM components.

Hidden State Update: The hidden state (h_t) is computed based on the cell state (c_t) and the output gate

$$o_t.h_t = o_t \odot \tanh(c_t) \quad (6)$$

Where, σ represents the sigmoid activation function, and tanh represents the hyperbolic tangent activation function. ⊙ represents element-wise multiplication.
Memory cells, gates (input, output, and forget gates) that regulate information flow into and out of the cells, and a hidden state that is transmitted from one-time step to the next are all components of the LSTM architecture.

Utilizing LSTM training and a loss function developed by Dongxu *et al*. (1997), this study computes the difference between the expected and actual outputs. A vector of memory cells, denoted as $c_t^l \boldsymbol{\epsilon} \boldsymbol{R^n}$, contains the "long term" memory. Explicit memory cells are a feature shared by all LSTM designs that enable long-term information storage. The memory cell can be overridden, retrieved, or retained for the subsequent time step by the LSTM. Given by the equation as:

$$\text{LSTM: } h_t^{l-1}, h_{t-1}^{l}, c_{t-1}^{l} \longrightarrow h_t^l, c_t^l$$
$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} sigm \\ sigm \\ sigm \\ tanh \end{pmatrix} T_{2n,4n} \begin{pmatrix} h_t^{l-1} \\ h_{t-1}^{l} \end{pmatrix}$$
$$c_t^l = f * c_{t-1}^l + i * g$$
$$h_t^l = o * \tanh(c_t^l)$$

Where sigm and tanh are applied element-wise.

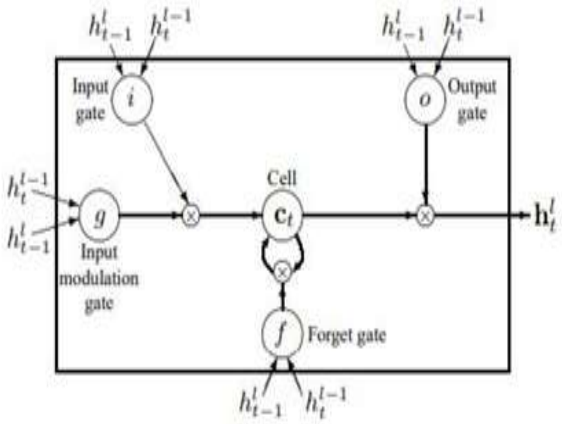

Fig 1: A graphical representation of LSTM memory cells by Wojciech Zaremba (2015)

This work uses loss function with the components (input gate, output gate, forget gate and cell gate) of LSTM to formulate machine translation and predict image caption generation Hochreiter and Schmidhuber (1997). Leveraging on this concept, we would apply LSTM using loss function to reduce over- fitting and predict inflation rate considering GDP and Per Capita as the economic indicators.

The LSTM model with a loss function mathematically using the variables:

$$\hat{Y} = H(X; W, b)$$
$$L = L(Y, \hat{Y})$$
$$min_{W,b} L$$

H represents the LSTM model with its weights and biases

Where; X represents your input data, which includes historical GDP and Per Capita values; Y represents your target variable, which is the historical inflation rates: W represents the weights of the LSTM model, including those of the LSTM cells: b represents the biases of the LSTM model.

Loss Function: L represents the loss function, such as Mean Squared Error (MSE). The loss function is used to measure the difference between the predicted and actual inflation rates.

Predictions ($\hat{Y}$): represents the predicted inflation rates.

K- Nearest Neighbors Classifiers: As a kind of supervised learning, K-NN classification aims to classify an item represented by many measurements into one of a limited number of classes. A specific amount of training examples are available for each class, and these are utilized to train the classifier in order to categorize novel patterns (Ethem Alpaydin 1997). Because it is straightforward and simple to apply, the method is one of the most popular categorization algorithms. Furthermore, it is frequently employed in numerous domain issues as the baseline classifier (Jain *et al*. 2000). The K-NN method immediately searches through all of the training instances for a specific testing example during the classification stage by computing the distances between the testing example and all of the training data in order to find its nearest neighbors and produce the classification output (Mitchell 1997). The KNN was used to measure the distance data points by using the Euclidean distance of two data points X_1 and X_2 which can be expressed mathematically as:

$$d(X_1, X_2) = \sqrt{\sum_1^n (X_{1i} - X_{2i})^2}. \quad (7)$$

This study use a classification approach to predict inflation rate, expressed mathematically as:

$$C(x_{new}) = arg \max_c \sum_{i=1}^{k} I(Y_i = c) \quad (8)$$

Where, c represents the class. Y_i is the class label of the i^{th} nearest neighbor. I is an indicator function that returns 1, if $Y_i = c$ and 0 otherwise

To estimate the parameter of the model:

$$Y(X_{new}) = \frac{1}{k} \sum_{i=1}^{k} y_i \quad (9)$$

Where X: represents your dataset, which contains data points with features and labels (class for classification or target values for the model. x_{new} : is the observation for which you want to make a prediction (classify or regress).

k: represents the number of nearest neighbours to consider for making predictions. It is a hyperparameter that you set before applying the algorithm. $C(x_{new})$: is the predicted class for the observation. It is determined by the most common class among the k-nearest neighbors of x_{new}; $Y(x_{new})$: is the predicted value for the observation. It is typically the mean values of the k-nearest neighbours of x_{new}

In determining the optimal k for predicting inflation rates based on GDP and per capita data involves a combination of domain knowledge, experimentation, and cross-validation. You should aim to strike a balance between model complexity and predictive accuracy. Experiment with different k values and select the one that provides the best balance for your specific problem and dataset. Organize your dataset, ensuring that you have historical data for inflation rates, GDP, per capita, and class labels representing the inflation rate categories (e.g., "Low," "Moderate," "High").

$$P(I_t|I_{t-1}, I_{t-2}, \dots, I_{t-k}) = \frac{1}{k}\sum_{i=1}^{k} P(I_t|I_{t-1}, I_{t-2}, \dots, I_{t-k}) \quad (10)$$

To show the mathematical expression of KNN using classification to predict inflation rate, this study uses the following steps:

a. Define the terms:

I_t is the inflation rate at time t

$I_{t-1}, I_{t-2}, \dots, I_{t-k}$ are the inflation rates at the previous k time steps

$P(I_t|I_{t-1}, I_{t-2}, \dots, I_{t-k})$ is the probability of the inflation rate being *It* at time *t*, given that the inflation rates were $I_{t-1}, I_{t-2}, \dots, I_{t-k}$ at the previous *k* time steps.

k is number of most similar data points to use in the prediction

b. Write down the Markov Chain property:

$$P(X_{t+1} \mid X_t, X_{t-1}, X_{t-2}, \dots, X_0) = P(X_{t+1} \mid X_t) \quad (11)$$

Where:

X_t *is* the state of the Markov chain at time t.

$P(X_{t+1} \mid X_t)$ *is* the transition probability from state X_t to state X_{t+1}

The KNN prediction equation is as follows:

$$Y_{pred} = \frac{1}{k}\sum_{i=1}^{k} y_i \quad (12)$$

Where:

Y_{pred} is the predicted value for the new data point ;

k is the number of most similar data points to use in the prediction;

y_i is the value of the ith most similar data point in the training set

d. Combine the Markov Chain property and the KNN prediction equation:

Combining the Markov Chain property and the KNN prediction equation, we get the following expression:

$$P(I_t|I_{t-1}, I_{t-2}, \dots, I_{t-k}) = \frac{1}{k}\sum_{i=1}^{k} P(I_t|I_{t-1}, I_{t-2}, \dots, I_{t-k})$$

This expression states that the probability of the inflation rate being *It* at time *t*, given that the inflation rates were $I_{t-1}, I_{t-2}, \dots, I_{t-k}$ at the previous *k* time steps, is equal to the average of the probabilities of the inflation rate being I_t at time *t*, given that the inflation rates were $I_{t-1}, I_{t-2}, \dots, I_{t-k}$ at the previous *k* time steps, for each of the *k* most similar data points in the training set.

Hidden Markov Model (HMM): When the states that produced the observations are not readily observable, a hidden Markov chain (HMM) is a statistical model that is used to represent the sequence of observations. Natural language processing, machine learning, and speech recognition are just a few of the many uses for HMMs.

The mathematical expression for an HMM is as follows:

$$\lambda = (S,\ M,\ A,\ B,\ P(x_t|s_t),\ P(\mathrm{s_{t+1}}|\mathrm{s_t}) \quad (13)$$

Where**:**

S is the set of hidden states ;

M the set of observable states;

A is the set of emissions ;

B *is* the transition matrix;

$P(s_0)$is the initial state distribution;

$(x_t|s_t)$ is the emission probability distribution;

$P(S_{t+1}|S_t)$ is the state transition probability distribution

The complete description of an HMM is given by its parameters = (A, B, π), where A is the transition probability matrix, B is the emission probability matrix, and π is the initial state probability matrix.

Let I be the number of our hidden states. The Hidden states are denoted by:

$$\boldsymbol{I_0} = \{\boldsymbol{I_1}, \boldsymbol{I_2}, \boldsymbol{I_3}, \boldsymbol{I_4}\} \quad \textbf{(14)}$$

Where I = 4, I_1= Low GDP and Low GDP Per Capita (LL), I_2= Low GDP and High GDP per capita (LH), I_3= High GDP and Low GDP Per Capita (HL), I_1= High GDP and High GDP per capita (HH)

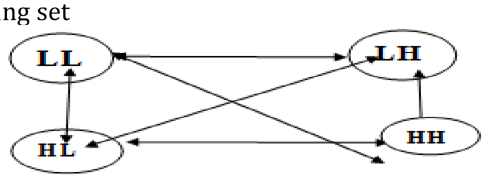

Fig 2: Transition Diagram for the Hidden Markov model for both Low and High inflation rate.

O is the number of observations. The set of possible observations is denoted by:

$$G_0 = \{g_1, g_2, g_3, \}$$

Where g_1= Low, g_2= Moderate, and g_3= High

We define X as a stationary state sequence of length I, and its equivalent observation sequence as G_0

$$X = \{x_1,\ x_2, x_3, x_4\} \quad G_0 = \{g_1, g_2, g_3\}$$

Transition and Emission Matrix: Let A be the $N \times N$ dimension transition probability matrix. It records the likelihood that state j will come after the state in the a_{ij} cell:

$$a_{33} = P(X_4 = s_4|r_4 = s_4) \qquad (15)$$

$$\begin{array}{c} \\ LL \\ LH \\ HL \\ HH \end{array} \overbrace{\begin{array}{cccc} LL & LH & HL & HH \\ \end{array}} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

In this case, HH is High|High, LL is Low|Low, while LH and HL are Low|High and High|Low. The observation output, or emission, probability matrix of dimension N×N is represented by B. The probability of producing observation k at state i, which is independent of time instant t, is defined as $bi_{(k)}$.

$$b_3(k) = P\left(r_k = X_t|X_{t+1} = s_j\right) \qquad (16)$$

$$\begin{array}{c|ccc} & L & M & H \\ \hline LL & b_{11} & b_{12} & b_{13} \\ LH & b_{21} & b_{22} & b_{23} \\ HL & b_{31} & b_{32} & b_{33} \\ HH & b_{41} & b_{42} & b_{43} \end{array}$$

π is my initial state probability vector.

$$\pi = (\pi_0, \pi_1, \pi_2, \pi_3)$$

Methods in Hidden Markov model (hmm: Based on the lessons Jack Ferguson taught in the 1960s, Rabiner (1989) proposed that three primary difficulties should describe hidden Markov models.

The evaluation problem, the decoding problem, and the learning problem are the three fundamental issues. First introduced in a major tutorial by Rabiner (1989), the concept that hidden Markov models should be characterized by three fundamental difficulties was based on tutorials by Jack Ferguson in the 1960s.

The Evaluation Problem: The chance that a sequence of observations, $Y_0^T = r_0^T$,, is produced by a Hidden Markov Model, let's say λ, is given by $P_\lambda(Y_0^T = r_0^T)$. When dealing with numerous competing models, the evaluation problem's solution yields a superior model that better fits the observation sequence.

Likelihood Problem: Find the likelihood $P(O|\lambda)$ given an HMM.
$\lambda = (A, B)$ and an observation sequence O.
Decoding Problem: Find the optimal hidden state sequence Q from an observation sequence O and an HMM $\lambda = (A, B)$.

Learning Problem: Learn the HMM parameters A and B from the set of states in the HMM and an observation sequence O.

Out of the three problems in Hidden Markov Model (HMM), this research work would focus on the likelihood problem using the Forward-Backward Algorithm for training and for predicting.

LSTM and KNN with Hidden Markov Model: Combining LSTM networks and KNN with HMMs can result in a significant improvement in performance for a variety of tasks, such as speech recognition, natural language processing, and machine translation. Hainan Xu *et al.* (2014) used LSTM networks and KNN to improve the performance of HMMs for machine translation. They found that LSTM networks and KNN were able to learn the transition probabilities and emission probabilities in the HMM more accurately than traditional methods.

This research uses two supervised machines learning language (LSTM and KNN) to predict inflation rate and compare which of the model is better.
Mathematically, the use of LSTM networks and KNN to improve the performance of HMMs can be expressed as follows:

$$P(X_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0) = \sum_{s\epsilon S} P(X_t\ |S_t)\ P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0) \qquad (17)$$

Where: X_t is the observation at time t;
S_t is the hidden state at time t; S is the set of all possible hidden states

The first term on the right-hand side of the equation is the probability of the observation at time t given the hidden state at time t. This term can be learned using an LSTM network. The second term on the right-hand side of the equation is the probability of the hidden state at time t given the observations from time 0 to time t−1. This term can be learned using an LSTM network or a KNN classifier.

To prove that combining LSTM networks and KNN with HMMs can improve the performance of HMMs, we can use the following steps:

Define the terms: X_t $Observation\ at\ time\ t$

X_t Hidden state at time t S: Set of all possible hidden states

$P(X_t\ |S_t)$: Probability of the observation at time *t* given the hidden state at time *t*

$P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0)$: Probability of the hidden state at time *t* given the observations from time 0 to time *t*−1

Write down the Bayes' theorem:

Bayes' theorem states that the probability of the hidden state at time *t* given the observations from time 0 to time *t*−1 can be calculated as follows:

$$P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0) = \frac{P(X_t\ |S_t) P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,\dots,} X_0 - 1)}{P(X_t)} \quad (18)$$

Where:

$P(X_t)$is the probability of the observation at time t

Write down the HMM equation:

The HMM equation states that the observation probability at time *t* given the observations beginning at time 0 to time *t*−1 can be calculated as follows:

$$P(X_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0) = \sum_{s_t \epsilon S} P(X_t\ |S_t)\ P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0)$$

Combine the Bayes' theorem and the HMM equation:

Combining the Bayes' theorem and the HMM equation, we get the following equation:

$$P(X_t\ |X_{t-1}, X_{t-2}, X_{t-3,} \dots, X_0) = \frac{P(X_t\ |S_t)\ \sum_{s_t \epsilon S} P(S_t\ |X_{t-1}, X_{t-2}, X_{t-3,\dots,} X_0 - 1)}{P(X_t)} \quad (19)$$

This equation shows that the observation probability at time *t* given the observations beginning at time 0 to time *t*−1 can be calculated by summing over the probabilities of the observation at time *t* given each possible hidden state and the probabilities of the hidden state at time *t* given the observations from time 0 to time *t*−1.

RESULTS AND DISCUSSIONS

This section presents the analysis of data collected and interpretation of results. Data were analyzed using machine learning models with Hidden Markov Model. In addition, tables, charts, graphs, correlation analysis and trend analysis were used to explore the moving pattern of the inflation rate.

Descriptive Statistics: Table 1 shows the descriptive measure for Inflation rate, GDP, GDP per capita and annual % change between the year.

Table 1: Summary Statistics of the inflation rate, GDP, GDP Per Capita and Annual % Change between the years

STATISTIC	Inflation Rate (%)	GDP (Billions of US $)	GDP Per Capita (US $)
Count	63	63	63
Min.	-3.7263	4.1962	93.3970
1st Quartile	7.5665	27.7656	363.5480
Median	12.0951	54.6041	567.5179
Mean	15.8522	147.1467	1008.1317
Standard Deviation	15.0302	172.6514	892.0328
3rd Quartile	17.3867	258.3579	1880.1500
Max.	72.8355	574.1838	3200.9531

From Table 1, the highest inflation rate in a particular year is 72.8355. The highest GDP is 574.1838 (Billions of US $) and the highest GDP Per Capita is 3200.9531. The average inflation rate is 15.8522 with a standard deviation of 15.0302, the average GDP is 147.1467 with a standard deviation of 172.6514 and the average GDP Per Capita is 1008.1317 with a standard deviation of 892.0328 while the minimum inflation rate in a particular year is -3.7263, this shows that the economy in the year 1967 was terrible and would have had negative effect on the masses. The minimum GDP is 4.1962 (Billions of US $) implies that the gross income was extremely small and the minimum GDP Per Capita is 93.3970. Figure 3 reveals that the lowest inflation rate in Nigeria was in far back of year 1967. There has also been a significant increase over the last seven years shown on the graph from 2016 to 2022 which moves to double digits.

Figure 4 shows that the highest GDP and GDP Per Capita happens in the year 2014. This indicate that there is a significant relationship between the GDP of Nigeria and her Per Capita which indirectly affect the standard living of the citizen and it most likely minimizes the effect of inflation rate on the economy.

The Scatter plot above shows that both the inflation rate and annual % Change is seasonal, why the GDP is regular but the GDP Per Capita is cyclical form of the components of time series. From year to year, it is mostly predictable for inflation rate and the GDP but since Nigeria population play significant roles in determining the GDP Per Capita, this indicator is almost unpredictable as shown in the plot above.

Fig 3: The Trend Analysis of inflation rate between the years 1960 to 2022 (cbn.gov.ng).

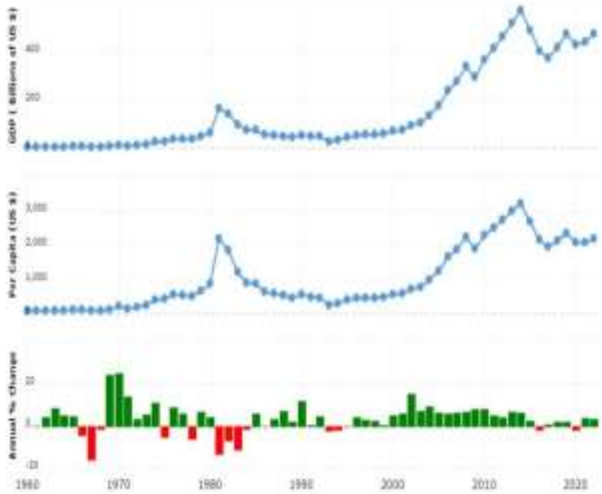

Fig 4: Time series plot of GDP and GDP Per Capita between the years 1960 to 2022 (cbn.gov.ng).

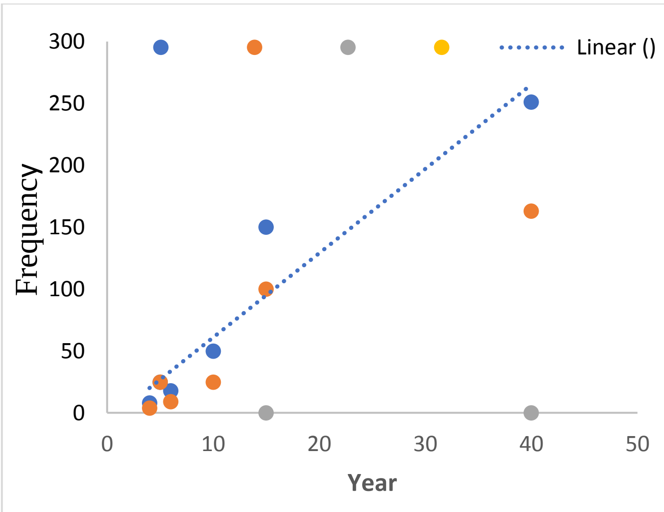

Fig 5: Scatter plot of the inflation rate, GDP, GDP Per Capita and Annual % Change between the years 1960 to 2022.

LSTM and KNN Training and Evaluation Analysis of Inflation Rate: We compare the performance of the Hidden Markov Model (HMM) technique with that of the K-Nearest Neighbors (K-NN) network and the Long Short-Term Model (LSTM) network in our analysis of the data. To enable us to train and test our models with 80% and 20% of the same data set, the datasets were partitioned.

LSTM Training and Evaluation Analysis of Inflation Rate: Data is passed from the input layer to the 45-node LSTM layer at the output. The dense layer 1 (one) receives the output of the LSTM layer and receives 300 nodes as input. Mean square error (MSE) is the loss function and ADAM is the optimizer in the dense layer 1. At last, the output layer—which is likewise a fully linked layer—is connected to the dense layer 1.

Table 2: Some LSTM parameters in training and evaluation

Hidden Layer	Node Or Neurons	Epochs	Loss	Val_Loss	Test Accuracy
1	45	50	0.8500	0.3054	0.6714

From table 2, we monitor the loss and validation loss throughout the training process using this simple code: monitor = EarlyStopping(monitor = 'val_loss', min_delta = 1e4, patience = 6, verbose = 2, mode = 'auto') to save the best performing model. The training loss of 0.8500 indicates that the model is performing reasonably well with respect to the trained data.

The validation loss of 0.3054 is lesser than the training loss and thus, indicates that the model is not over-fitting to the training data and may generalize well to new, unseen data. The test accuracy of 0.6714 suggests that the model is not performing very well on the test data. This means that the model may not be generalizing well to new and unseen data.

The plot of predicted Inflation Rate against GDP and Per Capita helps to understand how well the model predicts inflation based on these economic variables. The trained data at 20% shows that economic indicator (GDP and GDP Per Capita) has significant effect in having moderate inflation rate. If these indicators are not taken into consideration, there will always be a future high rate of inflation.

GDP Per Capita drives this effect more than the GDP because it has a direct impact on the population and standard of living of the masses. There is a significant relationship between the GDP and GDP Per Capita. It is however necessary to understand the economic implications of the indicators because the trend line is not linear based on the model.

K-NN Evaluation Analysis of Inflation Rate: To implement the K-NN mode, we first divide the dataset into dependent and independentvariable:

Independent Variable (X): These independent variables are the 'GDP', and 'GDP Per Capita'.

Dependent Variable (y): The dependent variables is the Inflation Rate' column.

Encoding Categorical Variables: To implement our model, there is a need to ensure that all the variables are in numerical form. For this purpose, we encode our economic indicator and inflation rate column with '1s' and '0s', where 0 represent GDP and 1 represent GDP Per Capita for economic indicator column and 0 represent High, 1 represent Moderate, and 2 represent Low for inflation rate column.

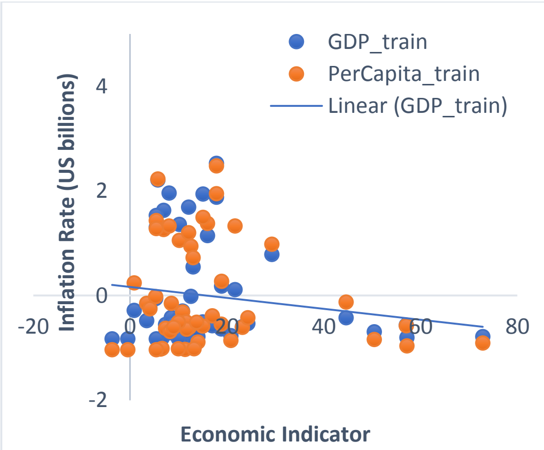

Fig 6: Plot of predicted Inflation Rate against GDP and GDP Per Capita

Feature Scaling: To get a very accurate prediction from our K-NN model, we need to standardize the values and ensure that they are of the same range. To do this, we use:

$$Z = \frac{X - mean(X)}{Std(X)} \quad (20)$$

Train/Test Split: We divided the data into two parts (a ratio of 1:5). One part was used to train the model, while the other part was used to test the performance of the model. These sets of data were respectively called Training set and Test set.

Thus, the Training set consists of 63 samples, while the Test set consists of 6 samples. The division was done randomly using the pandas' python library.

Table 3: Predicted Inflation Rate, GDP and GDP per capita

InflationRate_Predicted	GDP_train	PerCapita_train
20.8128	-0.768	-0.86035
7.3644	-0.55591	-0.62703
11.3964	-0.82791	-1.02977
57.0317	-0.56744	-0.57888
5.4027	1.305398	1.434736
72.8355	-0.78452	-0.90951
14.0318	-0.78292	-0.89277
50.4667	-0.69333	-0.84263
8.5299	-0.42234	-0.14896
10.8261	-0.2968	-0.31749
12.2242	-0.57524	-0.60805
3.4576	-0.47858	-0.15148
14.998	-0.50703	-0.58658
0.8568	-0.28478	0.237742
11.5811	-0.53847	-0.50558
5.7172	2.207591	2.224015
5.2656	-0.05694	-0.02381
6.2791	-0.82641	-1.02619
6.6184	-0.81775	-1.0103
10.156	1.356982	1.051528
12.5378	-0.01567	0.941596
18.8472	-0.64162	-0.54558
8.2252	-0.71045	-0.68827
13.246	-0.82223	-1.01525
11.7097	-0.69317	-0.64782
5.388	-0.83052	-1.03819
6.9333	1.627967	1.259944
-0.4761	-0.8262	-1.03782
9.9964	-0.82412	-1.01979
8.0474	1.955828	1.326321
15.9991	1.143289	1.375457
5.4443	1.527275	1.281409
18.8736	0.178177	0.268109
13.007	0.548091	0.723277
21.7092	0.112216	1.330247
44.5888	-0.42388	-0.12498
23.2123	-0.59758	-0.61102
17.8635	1.876788	1.942206
9.9723	-0.44766	-0.50809
13.7401	-0.64292	-0.51235
57.1653	-0.80274	-0.96554
9.0094	-0.53598	-0.60635
17.8205	2.526141	2.478023
12.0951	1.689593	1.201913
16.9528	-0.57839	-0.39176
24.3	-0.53393	-0.4249
-3.7263	-0.82618	-1.03543
29.2683	0.782619	0.977356
15.0878	1.938926	1.495816
4.1035	-0.23974	-0.25714

Implementation of our K-NN Model: The K-NN model was built into a python library called sklearn. From this library, we imported the K-NN function called K-Neighbors Classifier ().

After fitting the model to the dataset, we obtained an accuracy of 83.1% on the Training set, implying that over 80% of its prediction was correct.

Calculate classification report (including precision, recall, and F1-Score)

report = classification report (actual labels, predicted labels, target_names=["Low", "Moderate", "High"])

Table 4: Classification Report on Model Performance

Classification Report:				
Inflation Rate	**precision**	**recall**	**f1-score**	**Support**
Low	1	1	1	2
Moderate	1	1	1	2
High	1	1	1	2
Accuracy	0.83	6		
macro avg	1	1	1	6
weighted avg	1	1	1	6

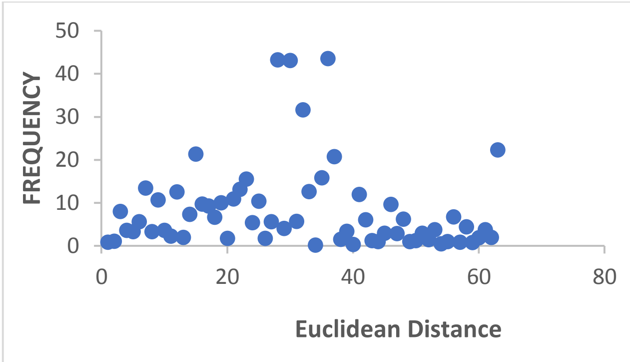

Fig 7: KNN Selection model

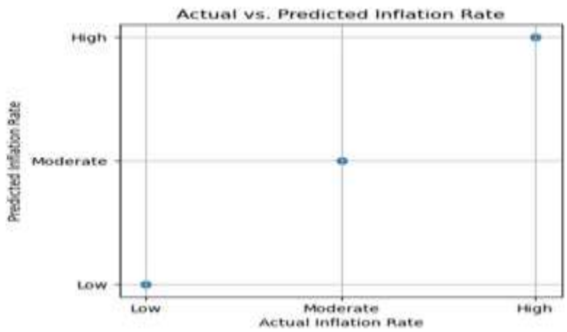

Fig 8: KNN model for comparing the predicted values to the actual inflation rate

The code splits the data into a training set (80% of the data) and a test set (20% of the data) to evaluate the model's performance. The KNN classifier with k=3 was applied on the training data. The model makes predictions on the test data, and the accuracy of the model on the test data is 0.83, indicating that it correctly predicted the Inflation Rate for all test data points. We then apply the KNN model to predict the Inflation Rate for each data point in the test set by comparing the predicted values to the actual Inflation Rate to evaluate the model's performance using a confusion matrix which is a table that is often used to describe the performance of a classification model. It provides a summary of the model's predictions versus the actual class labels. Here the predicted levels are “Low” “Moderate” and “High” Inflation Rates. The Confusion Matrix shows that the classification report provides additional insights into model performance, including metrics like precision, recall (sensitivity), and F1-Score for each class ("Low," "Moderate," and "High") which tells us how many of the predicted values for each class were correct and how many of the actual values for each class were correctly predicted hence, indicating a stable measure of a model's performance.

LSTM and K-NN with Hidden Markov Model for Inflation: Using K-NN with Hidden Markov Models to predict future years, this section uses Nigeria's Inflation rate to get the transition probabilities, confusion matrix, stationary state, and steady state with the number of hidden states (GDP and Per Capita) and observables variable (Low, Moderate and High). The summary of the data is presented in Table 1. **Figures should be numbered as 1, 2, 3, 4

Validity Test for the Model: The Hidden Markov Model (HMM) procedure analyzes the outcome. The transition matrix must be ascertained. After the data was imported into Python, the transition matrix for GDP and GDP Per Capita showed the following outcomes.

Table 5: Transition matrix for GDP and GDP Per Capita

	LL	*LH*	*HL*	*HH*
LL	0.3302	0.5377	0.1320	0.0001
LH	0.3701	0.4294	0.0045	0.1960
HL	0.0056	0.5591	0.3409	0.0944
HH	0.0001	0.1542	0.1212	0.7245

Table 5 reveals that the probability of remaining in state 1 (Low GDP, Low Per Capita) is 0.3302, the probability of remaining in state 2 (Low GDP, High Per Capita) is 0.4294. The probability of remaining in state 3 (High GDP, Low Per Capita) is 0.3409 and the probability of remaining in state 4 (High GDP, High Per Capita) is 0.7245.

The initial state probability is given as:

Table 6: Initial state probability

LL	*LH*	*HL*	*HH*
0.3345	0.1335	0.0255	0.5065

According to the underlying model, for Nigeria, the stable probability transitioning to state 1 is estimated at *33.45%, while the probability transitioning to state 2 and state 3 is estimated at 13.35% and 2.55% respectively, and the probability transitioning to state 4 is 50.65%.*

Most Likely States Using K-NN: The following are the most likely states determined by estimating the observed data: State 1 has a low GDP and low per capita, while state 3 has a high GDP and low per capita. 3 1 3 1 3 1 1 1 3 1 3 1 3 1 3 1 3 1 3 1 3 1

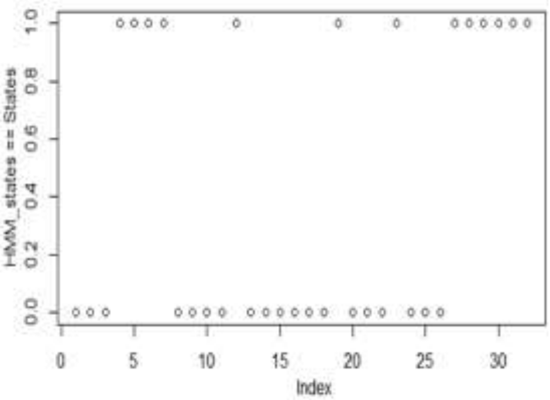

Fig 9: The plot demonstrates that the observable components can only be attributed to the hidden states.

Confusion Matrix: This confusion matrix shows that the model correctly classified 5 instances of Low Inflation Rate, 10 instances for Moderate and 6 instances for High Inflation. It also shows that the model incorrectly classified 10 and 2 instances of Low rate as Moderate and High respectively. While 3 and 4 instances of Moderate were classified as Low and High rate respectively. Also, 4 and 3 instances of High rate were classified as Low and Moderate. This indicates that 41.3% of the observable elements of inflation rate (Low, Moderate, High) can be attributed to the hidden state (LL, LH, HL, HH). The confusion matrix probabilities shows that the model correctly classified 26 out of 63 samples (41%). The model is most accurate at predicting state 2 and 3, with a success rate of 80%. The model is less accurate at predicting state 3, with a success rate of 30%.

$$
\begin{array}{cccc}
 & L & M & H \\
LL & 5 & 10 & 2 \\
LH & 3 & 10 & 4 \\
HL & 6 & 5 & 5 \\
HH & 4 & 3 & 6
\end{array}
$$

$$
\begin{array}{cccc}
 & L & M & H \\
LL & 0.2941 & 0.5882 & 0.1177 \\
LH & 0.1765 & 0.5882 & 0.2353 \\
HL & 0.3750 & 0.3125 & 0.3125 \\
HH & 0.3077 & 0.2308 & 0.4615
\end{array}
$$

Steady State: A Markov chain's steady-state behavior is the probability that it will persist in each state over an extended period of time. It is assumed that all states have the same beginning probabilities. The probability of a stable state as a result are:

Table 7: Steady State probability matrix

L	*M*	*H*
0.3199	0.5916	0.0885

This indicates that after a long period of time the system (inflation rate) will be Low of aorund 32%, at Moderate rate with about 59% and at High rate of about 9%. Considering the Gross income (GDP) and annual per capita will consistently predict the inflation rate.

Performance Evaluation: The findings presented in the next subsection display measures for the two models' predictions using different metrics. Table 4 displays the measures' values for each of the models.

Table 4: Measures for KNN and LSTM models

Models	MAPE	RMSE	Corr.	SEM	MSE
K-NN	10.94	3.21	-0.73	5.19	0.77
LSTM	12.56	4.23	-0.72	6.34	0.83

Comparing the evaluation metrics for both models. The model with lower MAPE, MSE, SEM or RMSE is generally considered better in terms of predictive accuracy. However, the choice might depend on your specific objectives. According to the study's conclusions, which can be observed by looking at the data in the table 4.3, the K-NN performed better than the LSTM. The K-NN prediction by all the used evaluation performance is lower compared to the LSTM prediction. This shows that K-NN is better in predictive accuracy. This on the other hand can depend on the fact that we are working with structured dataset with historical GDP and inflation values, with a known past pattern, so we can predict future inflation, K-NN might be a more straightforward and interpretable approach. The correlation between the two models is less than -0.70, which shows they are negatively correlated. The choice between K-NN and LSTM for predicting inflation based on GDP depends on the amount of historical data, the relationships between GDP and inflation, and the complexity of those relationships. If you have a time series of GDP data and believe there are intricate temporal dependencies, LSTM may be a more suitable choice. Considering the findings, when contrasting the HMM with LSTM and K-NN for predicting inflation rate, the K-NN performs better in predicting inflation rate.

Conclusion: The study established the KNN and the LSTM with Hidden Markov Model for predicting inflation rate in Nigeria**.** The KNN and LSTM were then used to forecast the inflation rate with respect to economic indicator (GDP and Per Capita) with steady-state probabilities of 32%, 59% and 9% for Low, Moderate and High rate respectively. Standard error of the mean (SEM), mean absolute percentage error (MAPE) and root mean squared error (RMSE) were used to assess the performance of the two models. With an RMSE of 3.21 and a MAPE of 10.94, the results of the analysis indicated that the KNN performed better than the LSTM. The future forecast of the model shows that Per Capita has a significant influence in the high rate of inflation in Nigeria and therefore recommend to the Federal Government of Nigeria to increase interest rates to reduce the money supply, subsidized essential goods such as food and fuel.

REFERENCES

Adebiyi, MA; Adenuga, AO; Abeng, MO; Omanukwe, PN; Ononugbo, MC (2010). Inflation forecasting models for Nigeria, Central Bank of Nigeria Occasional Paper No. 36, Abuja, Research and Statistics Department.

Altug, S; C. Cakmakli, C (2016). Forecasting inflation using survey expectations and target inflation, evidence for Brazil and Turkey, *Int. J. Forecast.* 32: 138-153.

Aron, J; Muellbauer, J (2012). Improving forecasting in an emerging economy, South Africa: Changing trends, long-run restrictions, and disaggregation, *Int. J. Forecast.* 28: 456-476.

Almosova, A; Niek, A (2023). "Nonlinear inflation forecasting with recurrent neural networks," *J. Forecasting, John Wiley & Sons, Ltd*., vol. 42(2), pages 240-259.

Balcilar, M; Gupta, R; Kotze, K (2015). Forecasting macroeconomic data for an emerging market with a nonlinear DSGE model, *Eco. Model.* 44: 215-228.

Chen, YC; Turnovsky, JS; Zivot E (2014). Forecasting inflation using commodity price aggregates, *J. Econom.* 183: 117-134.

Etuk, EH; Uchendu, B; Victoredema, UA (2012). Forecasting Nigeria inflation rates by a seasonal ARIMA model, *Canad. J. Pure and Appl. Sci*. 6 (3), 2179-2185.

Hochreiter, S; Schmidhuber, J (1997). Long Short-Term Memory. *Neural Computation* 9(8):1735-1780.

Hossain, B; Ahmed, M; Rabbi, MF (2012). A novel approach for inflation analysis using hidden markov model. *Int. J. Compt. Sci. Iss,* 9(2): 619–623.

Imimole, B; Enoma, A (2011). Exchange rate depreciation and inflation in Nigeria (1986-2008), *Bus. Eco. J.* Vol. 2011: BEJ-28.

Inam, U. S. (2017). Forecasting inflation in Nigeria: A Vector Autoregression Approach. *Int. J. Eco. Comm. Manag.* 5(1): 92-104.

Inge, A (2013). Hidden Markov Models: Theory and simulation, Mathematical Statistics Stockholm University.

Jochmann, M. (2010). Modeling U. S. inflation dynamics: A bayesian nonparametric approach. *Econom. Reviews*, 34: 3-10. DOI:10.1080/07474938.2013.806199

John, EE; Patrick, UU (2016). Short-term forecasting of Nigeria inflation rates using seasonal ARIMA Model. *Sci. J. Appl. Maths. Stat.* 4(3): 101-107.

Kelikume, I; Salami, A (2014). Time series modelling and forecasting inflation: Evidence from Nigeria. *The Int. J. Bus. F. Res.* 8(2): 91-98.

Mandalinci, Z. (2017). Forecasting inflation in emerging markets: An evaluation of alternative models, *Int. J. Forecast.* 33(4): 1082-1104.

Medel, CA; Pedersen, M; P. M. Pincheira, PM (2016). The elusive predictive ability of global Inflation. *Int. Finance* 19(2): 120-146.

Mustapha, AM; Kubalu, AI (2016). Application of box-inflation dynamics. *Ilimi J. Arts Soc. Sci.* 2(1).

Okafor, C; Shaibu, I (2013). Application of ARIMA models to Nigerian inflation dynamics, *Res. J. Fin. Acct.*, 4(3): 138-150.

Olajide, JT; Ayansola, OA; Odusina, MT; Oyenuga, IF (2012). Forecasting the inflation rate in Nigeria: Box Jenkins Approach, *IOSR J. Maths*.

Olubusoye, OE; Oyaromade, R (2008). Modeling the inflation process in Nigeria, AERC Research Paper 182. African Economic Research Consortium, Nairobi, Kenya

Pincheira, PM; Medel, CA (2015). Forecasting inflation with a simple and accurate benchmark: The case of the U.S. and a set of inflation Targeting Countries, Czech. *J. Eco. Fin.* 65(1).

Popoola, OP; Ayanrinde, AW; Rafiu, AA; Odusina, MT (2017). Time series analysis to model and forecast inflation rate in Nigeria. *Annals. Comput. Sci. Series*, 15(1).

Song, Y. (2011). Development and application of hidden markov models in the bayesian framework, Ph.D. thesis, University of Toronto,