

Asymmetric GARCH Type Models and LSTM for Volatility Characteristics Analysis of Nigeria Stock Exchange Returns

Samuel Olorunfemi Adams^{1,*}, Omorogbe Joseph Asemota², Abdulsalam Ahovi Ibrahim¹

¹Department of Statistics, University of Abuja, Abuja, Nigeria

²National Institute for Legislative Studies, Abuja, Nigeria

Abstract Modeling and predicting volatility of stock returns is very important for financial decision making, and it is of significant application to risk management, option pricing, strategic pair trading and portfolio selection. This study aimed to examine the asymmetric characteristics in the stock return volatility and compare the forecasting accuracy of the Long Short-term Memory (LSTM) and Asymmetric Generalized Autoregressive Conditional Heteroskedasticity (GARCH) type models like the Exponential GARCH (EGARCH) and the Threshold GARCH (TGARCH) Models based on the Nigerian stock market volatility returns. The data utilized in this study is the monthly all share index (ASI) of the Nigerian Stock Exchange (NSE) obtained from the Nigerian Stock Exchange for the period, January, 1985 to December 2023. The model efficiency and performance was measured with the Root Mean Squared Error (RMSE) criteria. The results from EGARCH (1,1) and TGARCH (1,1) models indicated that volatility of stock returns is persistent in Nigeria. The result of the forecasting accuracy of the LSTM and asymmetric GARCHs model based on the Nigerian stock market volatility returns indicated that EGARCH (1,1) outperformed LSTM (160-80-5) model based on the lowest AIC and BIC values. It is recommended that, stockbrokers should invest regularly, be consistent in trading, maintain a diversified portfolio and seek professional advice so as to curb the persistent volatility in Nigeria's stock return.

Keywords Asymmetric GARCH, EGARCH (1,1), LSTM, NSE, TGARCH (1.1)

1. Introduction

Volatility modeling of stock market returns has been gaining great interest by financial markets participants, academia, financial analysts and the general public. Volatility measures the uncertainty and risk which play significant role in modern financial analysis. Measuring and predicting volatility is crucial for financial decision making and has significant applications in areas such as portfolio selection, option pricing, risk management, hedging and strategic pair-trading as well as Value-at-Risk (VaR) estimation. Providing accurate volatility estimates in Nigeria market avails regulators, government, traders and investors the opportunity to formulate better policies and make appropriate financial investment decisions, (Emenike, 2010). Modelling volatility is an important element in pricing equity, risk management and portfolio management. Stock prices reflect all available information and the quicker they are in absorbing accurately new information, the more efficient is the stock market in

allocating resources. Modelling volatility will improve the usefulness of stock prices as a signal about the intrinsic value of securities, thereby, making it easier for firms to raise fund in the market. Also, detection of stock returns volatility-trends would provide insight for designing investment strategies and for portfolio management. The accuracy of volatility estimates and the precision of interval forecast is compromised if structural breaks are ignored (Orabi and Alqurran, 2015; Adewale *et al.*, 2016; Kuhe and Chiawa, 2017). Studies conducted by Perron (1989, 1990); Diebold and Inoue (2001); Gil-Alana *et al.* (2015), among others, revealed that when stationary processes are contaminated with structural breaks, the sum of Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized ARCH (GARCH) terms are always biased to unity. It is therefore, reasonable to incorporate these sudden shifts in variance when modeling and estimating parameters of volatility models. In our opinion, if the causes of shocks in stock prices are identified and accounted for, it will yield more reliable and accurate volatility estimates, this could also help in making good financial reforms that may have positive impacts and direct bearing on financial institutions and the economy.

* Corresponding author:

samuel.adams@uniabuja.edu.ng (Samuel Olorunfemi Adams)

Received: Jul. 17, 2024; Accepted: Aug. 12, 2024; Published: Aug. 29, 2024

Published online at <http://journal.sapub.org/ajms>

Volatility clustering occurs when large stock price changes are followed by large price change, of either sign, and small price changes are followed by periods of small price changes. Leptokurtosis means that the distribution of stock returns is not normal but exhibits fat-tails. In other words, Leptokurtosis signifies that high probabilities for extreme values are more frequent than the normal law predict in a series. Asymmetry, also known as leverage effects, means that a fall in return is followed by an increase in volatility greater than the volatility induced by an increase in returns. This implies that more prices wander far from the average trend in a crash than in a bubble because of higher perceived uncertainty (Mandelbrot, 1963; Fama, 1965; Black, 1976). These characteristics are perceived as indicating a rise in financial risk, which can adversely affect investors' assets and wealth. For instance, volatility clustering makes investors more averse to holding stocks due to uncertainty. Investors in turn demand a higher risk premium in order to insure against the increased uncertainty. A greater risk premium results in a higher cost of capital, which then leads to less private physical investment.

The financial series' trend is implied to be variable and unstable by volatility, which is typically not constant. Because of this, autoregressive conditional heteroskedasticity econometric models, or ARCH-type models, must be used instead of the traditional models with homoscedastic variance. The ARCH model has the benefit of taking conditional volatility into account, but it also has the drawback of typically requiring a very high number of parameters to be estimated, which leads to worse fits for this kind of financial data. Given that the conditional variance depends on both the squares of the observations and the conditional variances of earlier periods, the GARCH model has an advantage over the ARCH model. It does not, however, provide the variation's sign (positive or negative). The EGARCH model solves this problem by making it possible to determine whether price changes are positive or negative. Similarly, statistical models like the Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) model can assist explain time-varying volatility patterns, (Yahaya *et al.* 2022; Mohammed *et al.* 2022; Adams *et al.*, 2023). Although these sophisticated modeling techniques have great potential, there are still obstacles in successfully implementing them in the Nigerian setting. Among other things, non-stationarity, feature selection, and data quality are concerns in exchange rate forecasting.

It is difficult for traditional econometric models to capture the complex patterns of exchange rate changes, especially in a time when financial markets are changing quickly. Deep learning techniques such as Long Short-term Memory (LSTM) methodology, which can handle financial data forecasting, have opened up new prospects to improve the accuracy of currency exchange rate predictions, (Hochreiter and Schmidhuber, 1997). One kind of machine learning model that handles consecutive data inputs and outputs is the recurrent neural network (RNN). By providing feed forward neural networks with feedback, RNN is able to capture the temporal relationship between input/output sequences.

Recurrent layers are added to the neural network (NN) model to create an RNN. A generic RNN model consists of three sets of layers: input, recurrent, and output. The input layers convey the features of the input once the input data have first been converted into a vector. Next, the feedback-providing recurrent layers are implemented. The model then ends with fully connected (FC) layers at the end, also known as RNN model output layers, in the same manner as other NN models. The two main types of RNN models are LSTM-based RNN models and GRU-based RNN models, (Cho, *et al.* 2014). The recurrent layers that are used in each model set them apart from one another. Recurrent neural networks (RNNs) include Long-Short Term Memory (LSTM). It employs gate approaches and recurrent mechanisms to process information in sequences. When modeling time series, including audio and video, LSTM is superior to other feed forward and recurrent neural networks in a number of ways, (Hochreiter and Schmidhuber, 1997). To fully utilize the long short term memory (LSTM) in time series modeling, it is supplemented with other identification techniques. The simulation model, which is utilized for multistep prediction, is incompatible with it. They must make advantage of the previous test results, (Yu, 2018). The NN's output that feeds back to the input enables the modeling of data chains or sequences. The RNN training approach has a Vanishing Gradient problem since it requires back propagation over time. Several gated units are used by LSTM to prevent this. Sequence prediction, natural language processing, audio recognition, and time series modeling are just a few of the fields in which LSTM has been extensively used, (Graves, *et al.*, 2013).

This study aim to examine asymmetric characteristics in the volatility of the Nigerian stock market time series and compare the forecasting accuracy of the LSTM model and GARCHs model based on the Nigerian stock market volatility returns.

2. Literature Review

Modeling volatility of stock market return series using time varying GARCH models proposed by Engle (1982), Bollerslev (1986) and extended by Nelson (1991), Glosten *et al.* (1993), Ding *et al.* (1993), Zakoian (1994), *e.t.c.*, has been gaining attention in recent times by policy makers, academics, financial analysts and researchers among others. This is partly as a result of the fact that GARCH family models have been more successful in capturing stylized facts (statistical regularities) of financial time series such as volatility clustering, volatility shock persistence, volatility mean reversion, leverage effect and risk premium among others; and partly because volatility is an important concept for many economic and financial applications such as risk management, option trading, portfolio optimization and asset pricing. The prices of stocks and other assets depend on the expected volatility of returns. As part of monitoring risk exposure, banks and other financial institutions make use of volatility assessments (Engle and Patton, 2001).

Many scholars have documented evidence relating to

volatility models in the presence of structural breaks across developed and emerging economies. Lamoureux and Lastrapes (1990) found out that not incorporating structural breakpoints in the conditional variance while modelling volatility increases persistence of volatility shocks whereas incorporating the sudden shifts in conditional variance reduces the persistence of shocks in volatility models.

Donaldson (1996) used basic ANN to forecast stock volatility and concluded that ANN has a better performance as ANN can fit complex non-linear relationships that traditional time series models.

Charalambous *et al.* (2000) conducted a comparative analysis of ARCH-type models and artificial neural networks. The results revealed little evidence that ANN prediction outperformed traditional models. Malik *et al.* (2005) investigated the persistence of volatility shocks on the Canadian stock data using heteroskedastic models. Results showed reduction in shocks persistence in volatility when structural break points were incorporated in the conditional variance while estimating volatility. Hammoudeh and Li (2008) also found significant decrease in volatility shock persistence when valid sudden shifts in variance were incorporated while predicting volatility in Gulf Arab countries stock markets.

The study by Hossain and Nasser (2008) compares the GARCH model against the Back Propagation Artificial Neural Network (BPANN) in forecasting four internationals, including two Asian stock market indexes. The fitted GARCH models outperformed the fitted standard BP-ANN models in forecasting the four worldwide market indices, with the exception of one market.

Kumar and Mahfeswaran (2012) show that the estimates of stock returns volatility are considerably affected by sudden structural breakpoints or sudden regime shifts which occur as a result of domestic and external shocks. Su *et al.* (2012) applied the long short-term memory (LSTM) process to the stock sequence prediction of China A-stock. The study found that LTSM can accurately learn the pattern of the China A-stock market. Idowu *et al.* (2012) utilized an artificial neural network to investigate how to predict the Nigerian stock market. The study examined what happened when artificial neural networks were used to predict the value of specific Nigerian Stock Exchange (NSE) bank market indexes. A multilayer feed forward neural network transmits the intended reaction to input signals to each output unit. The study demonstrated that artificial neural networks can forecast future stock prices. Mantri *et al.* (2014) used ANN to model the Indian market volatility, but the conclusion indicates a trivial difference in the forecast between ANN and other models. Falat *et al.* (2015) used ARCH and GARCH models in a comparative analysis of out-of-sample predictions and discovered that neural network models performed nearly as well as standard statistical models. This implies that they are fair and acceptable in economic models. In another study on the neural stochastic volatility model. Charef, and Ayachi (2016) offered an alternate way to predicting daily exchange rates using an artificial neural network (ANN). The analysis is based on a set of daily data

from Tunisia. To assess this strategy, ANN was compared to a generalized autoregressive conditional heteroskedasticity (GARCH) model in terms of performance. The results show that the suggested nonlinear autoregressive (NAR) model is an accurate and fast prediction approach. Luo *et al.* (2017) proposed a new method for calculating volatility using deep recurrent neural networks. When tested on real-world financial datasets, NSVM outperformed various regularly used models such as GARCH, EGARCH, GJRGARCH, and TARCH. Liu *et al.* (2017) introduced recurrent neural networks (RNN) combining sentiment data from the online forum, and the results showed a significant improvement in prediction performance. Yu *et al.* (2018) compared the performance of LSTM and GARCH models in predicting the fluctuations of China's equity market. Huang *et al.* (2019) constructs time series models and deep learning model, respectively, and compares the prediction results of the two types of models from the perspective of dynamic and static forecasting based on SSE index data. The results show that the forecasting methods of the models affect their forecasting effects, and the GARCH model has the highest average forecasting accuracy in static forecasting, while the LSTM model has the most accurate forecasting effect in dynamic forecasting, with an RMSE value of only 6.32%. Nasrin *et al.* (2020) investigated the forecast accuracy of data-driven models for monthly stream flow in the Urmia lake basin, employing the Autoregressive conditionally heteroskedastic time-series model. The results showed that ARCH-CHAID models outperformed all other models in the two stations under consideration. Shaik and Aditya (2020) investigated the efficacy of artificial neural networks (ANN) and GARCH models in predicting the volatility of Indian stock market indexes. The most recent 20% of the observations were subjected to out-of-sample testing using the GARCH (1,1) and recurrent neural networks. The results clearly revealed that the GARCH model outperformed the ANN. The ANN may be a superior indicator in periods of low volatility, but its effectiveness deteriorates in moments of high volatility. Zhou *et al.* (2021) used a deep learning model to forecast the volatility of SSE50 ETF options. A comparison between the deep learning and GARCH models showed that deep learning performs better. Shen, *et al.* (2021) evaluated out-of-sample forecasting accuracy and risk management efficiency. The results show that RNN outperforms GARCH and EWMA on average predicting performance. However, it is less effective in capturing the Bitcoin market's extreme occurrences. Furthermore, the RNN performs poorly in Value at Risk forecasting, implying that it may not work as well as econometric models in explaining severe volatility. Petrozziello *et al.* (2022) investigate the profitability of a deep Long Short-Term Memory (LSTM) Neural Network for forecasting daily stock market volatility using a panel of 28 assets representative of the Dow Jones Industrial Average index combined with the market factor proxied by the SPY and, separately, a panel of 92 assets belonging to the NASDAQ 100 index. The Dow Jones plus SPY data are from January 2002 to August 2008, while the NASDAQ 100 is from

December 2012 to November 2017. The study compared the volatility forecasts generated by the LSTM approach to those obtained through use of widely recognized benchmarks models in this field, in particular, univariate parametric models such as the Realized Generalized Autoregressive Conditionally Heteroskedastic (R-GARCH) and the Glosten–Jagannathan–Runkle Multiplicative Error Models (GJR-MEM). The results demonstrate the superiority of the LSTM over the widely popular R-GARCH and GJR-MEM univariate parametric methods, when forecasting in condition of high volatility, while still producing comparable predictions for more tranquil periods. Zahid *et al.* (2022) provided a forecast for bitcoin volatility using hybrid GARCH models and machine learning. Both the in-sample and out-of-sample results of their study demonstrate that these hybrid models can accurately predict how much Bitcoin's price will change. Chatterjee *et al.* (2022) investigated stock volatility prediction using time series and deep learning methods, introducing several volatility models based on the Exponential general autoregressive conditional heteroscedasticity (EGARCH), Glosten-Jagannathan-GARCH (GJR-GARCH), and generalized autoregressive conditional heteroscedasticity (GARCH) frameworks. Three types of GARCH models and the LSTM model were tested for their ability to predict stock volatility in three distinct industries. The findings revealed that the LSTM model performed better in predicting volatility in the pharmaceutical sector than in the banking and IT sectors. Goel *et al.* (2023) concluded that the Multilayer Perceptron Neural Network outperformed the GARCH model in forecasting monthly returns.

Cai *et al.* (2023) investigates stock price prediction based on GARCH and LSTM. The primary objective of the study was analyze the performance and respective advantages of the GARCH model and various neural network models integrating market sentiment indicators for stock volatility forecasting. Forecasting models include not only technical indicators but also market sentiment data. According to the analysis, simple RNN and LSTM predicted volatilities have the distribution characteristics closest to the actual situation, while traditional time series regression method GARCH has the poorest prediction performance. Although other studies have shown that the effect of layer normalization is better for RNNs and LSTM, adding batch normalization to all models has a better effect on performance improvement for the data of this study. Adams and Uchema (2024) compared Exponential Generalized Autoregressive Conditional Heteroskedasticity with order $p=1$ and $q=1$, (EGARCH (1,1)) and Recurrent Neural Network (RNN) based on long short term memory (LSTM) model with the combinations of $p=10$ and $q=1$ layers to model the volatility of Nigeria naira exchange rate with Euro, pounds and US dollars. The study intends to determine the preferred model for predicting Nigeria's Naira exchange rate volatility with Euro, Pounds and US Dollars. The results indicated that the Nigeria exchange rate volatility is asymmetric, and leverage effects are evident in the results of the EGARCH (1, 1) model. It was observed also that there is a steady increase in the Nigeria Naira exchange rate with

the euro, pounds sterling and US dollar from 2016 to its highest peak in 2023. Result of the comparative analysis indicated that, EGARCH (1,1) performed better than the LSTM model because it provided a smaller MSE values of 224.7, 231.3 and 138.5 for euros, pounds sterling and US Dollars respectively.

A summary of the reviewed studies above indicates that no study have been conducted on modeling Nigeria's stock exchange volatility using recurrent neural network based on long short-term memory (LSTM) and Asymmetric Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. This is the first study that straightforwardly compared LSTM approach with asymmetric GARCH model for Nigeria's all share index stock exchange volatility.

3. Materials and Methods

3.1. Source of Data

The data used in this research work are the monthly all share index (ASI) of the Nigerian Stock Exchange (NSE) obtained from the Nigerian Stock Exchange (NSE, 2023) for the period January, 1985 to December 2023 making a total of four hundred and sixty-eight (468) observations.

3.2. ARCH/GARCH MODEL

3.2.1. The Autoregressive Conditional Heteroskedasticity (ARCH) Family of Models

Every model in the ARCH or GARCH family needs two distinct specifications: the equations for mean and variance. The ARCH model can be used to model the conditional mean equation, which describes how the dependent variable, y_t changes over time. The model expresses the mean equation in the following manner:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \mu_t \quad \mu_t \sim N(0, \sigma_t^2) \quad (1a)$$

$$\sigma_t^2 = \beta_0 + \beta_1 \mu_{t-1}^2 \quad (1b)$$

Where; μ_t is the error from the mean equation at time t . This equation also applies to other GARCH family models. According to equation (1b), the ARCH model models the "autocorrelation in volatility" by allowing the conditional variance of the error term, σ_t^2 , to depend on the value of the squared error that came before it. Because the conditional variance is only dependent on a single lagged squared error, the model above is called an ARCH (1).

The ARCH (q) model was first proposed by Engle (1982). In this model, the conditional variance σ_t^2 is a linear function of the lagged squared residuals μ_t . The following is the formula for an ARCH model of order q 's variance equation:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^q \beta_i \mu_{t-i}^2 \quad (2)$$

Where $\beta_0 > 0$; $\beta_i \geq 0$; $\forall i = 1, \dots, q$

The ARCH (q) model in equation (2) allows for a time-varying variation in conditional variance about previous errors. The unconditional distribution of t in ARCH models is always leptokurtic. The required lag q often turned out

to be quite large in applications of the ARCH (q) model. The generalized ARCH (p, q) model (GARCH (p, q)) was developed by Bollerslev (1986) to achieve a more constrained parameterization.

3.2.2. Asymmetric Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Models

In financial markets, if “bad news” has a more pronounced effect on volatility than “good news” of the same magnitude, then a symmetric specification such as GARCH or GARCH-M is not appropriate, since only squared residuals ε_{t-1}^2 enter the equation, the signs of the residuals or shocks have no effects (in other words, by squaring the lagged error in GARCH, the sign is lost) on conditional volatility. In other word, the model assumes good and bad news have the same effect. However, a stylized fact of financial volatility is that bad news (negative shocks) tends to have a larger impact on volatility than good news (positive shocks). In the case of equity returns, such asymmetries are typically attributed to leverage effects, whereby negative shocks cause the value of the firm to fall which raises the debt-equity ratio thereby increasing the risk of bankruptcy (debt-equity ratios are a key predictor of the probability of default in credit scoring models), (Asemota and Ekejiuba, 2017). This makes shareholders, who bear the residual risk of the firm, to perceive their future cash low stream as being relatively more risky. In order to account for the leverage effects observed in stock returns, the asymmetric models which include: [TGARCH (1,1) and EGARCH(1,1)] are employed.

3.2.3. The Threshold GARCH (TGARCH (1,1)) Model

Asymmetry and leverage (the fact that volatility is negatively correlated with changes in stock returns) are not taken into account by either the ARCH or GARCH models. Although GARCH (p, q) models provide adequate fits for the majority of equity-return dynamics, these models frequently fail to accurately model the volatility of stock returns due to their assumption of a symmetric response between returns and volatility. As a result, the leverage effect of stock returns cannot be taken into account by GARCH models. When it comes to stocks, it is frequently observed that when the market experiences a decline of the same magnitude, volatility is higher than when it experiences a rise of the same magnitude. The threshold GARCH (TGARCH) model was developed by Zakoian (1994) to take into account the existing leverage effect.

The TGARCH model was defined by Zakoian (1994) by permitting the conditional standard deviation to depend on a single lag in innovation. Parameter restrictions that guarantee the conditional variance to be positive are not shown in the specification. However, the TGARCH model's parameters must be restricted and the error distribution chosen to account for stationarity to guarantee stationarity. The GJR- GARCH model is another name for the threshold GARCH model. A straightforward addition to the GARCH model, the GJR model adds a term to account for potential asymmetries.

Using TGARCH (p, q), the following is the generalized specification for the conditional variance:

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^q \beta_i \mu_{t-i}^2 + \sum_{i=1}^q \gamma_i \mu_{t-i}^2 d_{t-i} + \sum_{j=1}^p \alpha_j \sigma_{t-j}^2 \quad (3)$$

Where $d_{t-i} = 1$ if $\mu_{t-i} < 0$ and 0 if $\mu_{t-i} \geq 0$.

The condition for non-negativity will be $\beta_0 > 0$, $\beta_i > 0$, $\alpha_j \geq 0$, and $\beta_i + \gamma_i \geq 0$

In this model, good news implies that $\mu_{t-i}^2 > 0$ and has an impact of β_i and bad news implies that $\mu_{t-i}^2 < 0$ with an impact of $\beta_i + \gamma_i$. Bad news increases volatility when $\gamma_i > 0$, which implies the existence of leverage effect in the i th order, and when $\gamma_i \neq 0$ the news impact is asymmetric. These two shocks of equal size have different effects on the conditional variance.

The first order representation of TGARCH (p, q) is TARCHG (1, 1) given as:

$$\sigma_t^2 = \beta_0 + \beta_1 \mu_{t-1}^2 + \gamma_1 \mu_{t-1}^2 d_{t-1} + \alpha_1 \sigma_{t-1}^2 \quad (4)$$

In this model, good news has a positive impact of β_1 and negative news has a negative impact of $\beta_1 + \gamma_1$.

3.2.4. The Exponential GARCH (EGARCH (1,1)) Model

Nelson's (1991) exponential GARCH (EGARCH) model allows for asymmetric effects between asset returns that are positive and negative. Three major shortcomings of the GARCH model were proposed to be addressed by the EGARCH, which takes into account the asymmetric properties of volatility and returns. They are: Restrictions on parameters that guarantee positive conditional variance, insensitivity to volatility's asymmetric response to shock, and difficulty measuring persistence in a strong stationary series.

The EGARCH model's log of the conditional variance indicates that the leverage effect is exponential rather than quadratic. A major advantage of the EGARCH model over the symmetric GARCH model is that it does not restrict the sign of the model parameters because volatility is specified in terms of its logarithmic transformation. As a result, there are no restrictions on the parameters to ensure that the variance is positive. The following is a general description of the EGARCH (p, q) model's conditional variance.

$$\log(\sigma_t^2) = \beta_0 + \sum_{i=1}^q \beta_i \left| \frac{\mu_{t-i}}{\sigma_{t-i}} \right| + \sum_{i=1}^q \gamma_i \frac{\mu_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^p \alpha_j \log(\sigma_{t-j}^2) \quad (5)$$

In this model, good news implies that μ_{t-i} is positive with total effects $(1 + \gamma_i)|\mu_{t-i}|$ and bad news implies μ_{t-i} is negative with the total effect $(1 - \gamma_i)|\mu_{t-i}|$. When $\gamma_i < 0$, the expectation is that bad news would have a higher impact on volatility (leverage effect is present) and the news impact is asymmetric if $\gamma_i \neq 0$. The EGARCH model achieves covariance stationarity when $\sum_{j=1}^p \alpha_j < 1$. The EGARCH (1, 1) is specified as:

$$\log(\sigma_t^2) = \beta_0 + \beta_1 \left| \frac{\mu_{t-1}}{\sigma_{t-1}} \right| + \gamma_1 \frac{\mu_{t-1}}{\sigma_{t-1}} + \alpha_1 \log(\sigma_{t-1}^2) \quad (6)$$

The total effect of good news for EGARCH (1, 1) is $(1 + \gamma_1)|\mu_{t-1}|$ and the total effect of bad news for

EGARCH is $(1 - \gamma_1)|\mu_{t-1}|$. If the null hypothesis that $\gamma_1 = 0$ is rejected then, a leverage effect is present, that is, bad news has a stronger effect than good news on the volatility of the stock index return, and the forecasts can be tested by the hypothesis $\gamma_1 < 0$.

3.3. Long Short-Term Memory (LSTM) Deep Learning Model

In this study, LSTMs with better design and gates to control information flow will be used to function as memory units. A typical RNN contains a single hidden state that is updated over time, making it difficult for the network to learn long-term dependencies. The LSTM model addresses this issue by incorporating a memory cell, which is a container capable of storing information for an extended length of time. LSTM architectures can learn long-term dependencies in sequential data, making them ideal for tasks like language translation, speech recognition, and time series forecasting, (Sundermeyer, *et al.*, 2012). In LSTM designs (see Figure 1), the memory cell is controlled by three gates: the input gate, the forget gate, and the output gate. These

gates determine what information is added to, removed from, and output from the memory cell. (i) The input gate determines what information is added to the memory cell. (ii) The forget gate determines which information is erased from the memory cell. (iii) The output gate determines what information is output by the memory cell. The LSTM maintains a concealed state that functions as the network's short-term memory. The input layer is represented at the bottom, the output layer is represented at the top and the unfolded recurrent layers are represented horizontally. A unit layer is called a cell that takes external inputs, inputs from the previous time cells in a recurrent framework, produces outputs, and passes information and outputs to the cells ahead in time. The cell state is defined as the information that flows over time in this network (as recurrent connections) with the information content having a value of $c(t)$ at time t . The cell state would be affected by inputs and outputs of the different cells, as we go over the network (or more concretely in time over the temporal sequences). Similarly, the network passes the output $y(t)$ from the previous time to the next time as a recurrent connection, (Kala, 2021).

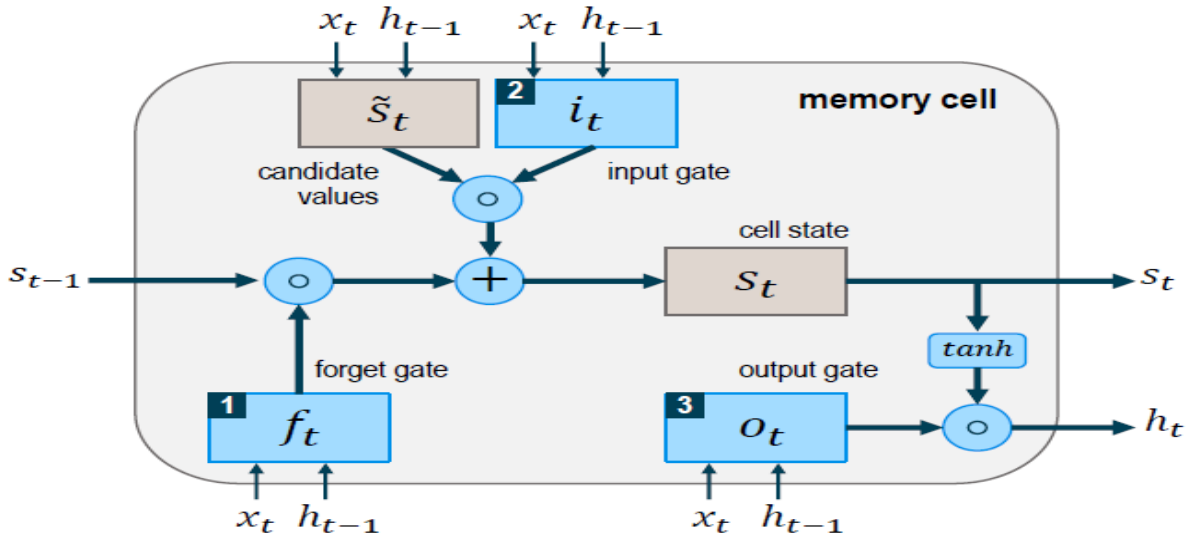


Figure 1. Architecture with Long Short-Term Memory (LSTM) Network (Source: Memory Cell in LSTM network (Fischer & Krauss, 2017))

According to the former settlement, Hochreiter and Schmidhuber (1997) introduced a new architecture: the Long Short-Term Memory (LSTM). This structure works in a similar way of the Elman process, as a matter of fact, it uses one input layer, one hidden layer and one output layer. But the fully self-connected hidden layer presents a complex internal processing unit, named “memory cell”, with a series of corresponding regulators, usually called “gates” that supervise and manage the entering and exiting flow of information. The memory cell, which is depicted in Figure 1, is responsible for keeping note of all the dependencies between inputs. It is usually composed by three gates: a forget gate (f_t), an input gate (i_t), and an output gate (o_t). Each of them, at every time step t , interacts with one element of the input sequence, x_t , and the output of the memory cell from the previous time step, h_{t-1} . Specifically, the forget gate defines which information to delete from the cell state,

the input gate which data to add to the memory and the output gate which results to use as output from the cell state. Each gate operates similarly to the nodes of the neural network, in fact, they act on the received signals based on the inputs and their weights, and they block or pass the information using an activation function (Fischer & Krauss, 2017). Figure 2 gives a representation of all the steps in which the work of the memory cell is organised and their explanations from a mathematical point of view. The first step (A) is to decide which information should be deleted from the network through the forget gate. So, the activation values f_t of such gate is calculated and scaled in a range between 0 and 1 by a sigmoid function, meaning respectively “completely forget” or “completely remember”. If the value is a number between them, the information is deleted or kept depending whether the produced value is nearer 0 or 1. The meaning behind the forget gate is ascribable to the fact that,

sometimes, the memory cell recognizes as useless, for the assigned task, some collected information and wants to proceed by deleting them. The next step (B) is deciding which data should be added to the network's cell state, st . This operation happens through two phases: the use of a tanh function which designs new candidate variables that could be added to the unit cell and the application of a sigmoid function, which defines the activation values, it , of the input gate. In the following step (C), there is the need to update the

old state of the memory cell by combining the results of the two previous steps and by transferring them into the cell. The final stage (D) consists in creating results as the output of the memory cell. This concluding step is done, first, through the use of a sigmoid function, that define which data contained in the memory cell should be presented as the output, and then, through the application of tanh function, which delimits the values of the output in a range between -1 and 1 (Olah, 2015).

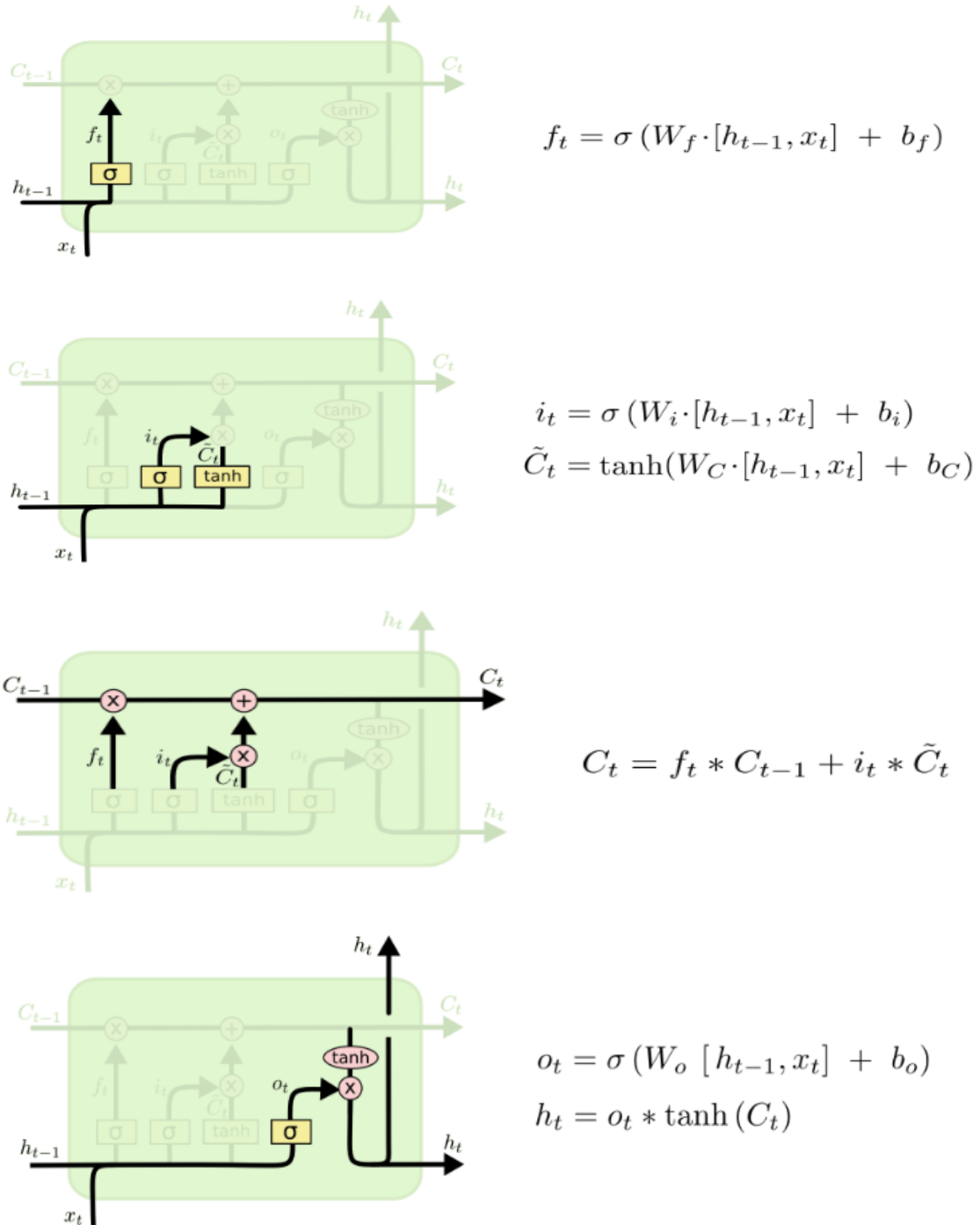


Figure 2. Mode of operation of a memory cell in LSTM network (Olah, 2015)

3.4. Model Selection Criteria

The best model for each of the two Nigeria stock exchange volatility will be selected based on the following criteria; (i) the Akaike information Criterion (AIC), Akaike, H., (1973), (ii) Schwarz information criterion (SIC), Schwarz, G. (1978) and root mean square (RMSE). The volatility of Nigeria stock exchange will based on estimated coefficients of the best conditional variance models, and the model with the least value for these three criteria across the error distributions is adjudged the best fitted. This selection produces the best-fitted conditional variance models for stock returns.

- i. The Akaike information criterion (AIC) is:

$$2k + n \ln \left(\frac{RSS}{n} \right) \quad (7)$$

- ii. Schwarz information criterion (SIC) is:

$$n \ln \left(\frac{RSS}{n} \right) + k \ln n \quad (8)$$

- iii. The RMSE will produce the preferred model from the two competing models., it is written as;

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (9)$$

4. Result

4.1. Preliminary Result

4.1.1. Descriptive Analysis of Nigeria Stock Exchange Return

Table 1 shows the descriptive statistics of the NSE return series. The average monthly return is 5070.728. The monthly standard deviation is 5707.306, reflecting a high level of volatility in the market. The wide gap between the maximum (20682.29) and minimum (4.82) returns gives support to the high variability of change in the NSE. Under the null hypothesis of normal distribution, Jarque-Bera is zero. The Jarque-Bera value of 56.84852 deviated from normal distribution. Similarly, skewness and kurtosis represent the nature of departure from normality. In a normally distributed series, skewness is 0 and kurtosis is 3. Positive or negative skewness indicate asymmetry in the series and less than or greater than 3 kurtosis coefficient suggest flatness and peakedness respectively, in the returns data. The skewness coefficient of 0.678003 is positively skewed.

Positive skewness implies that the distribution has a long right tail and a deviation from normality. The empirical distribution of the kurtosis is clearly not normal but peaked.

On the whole, the NSE return series do not conform to

normal distribution but display positive skewness and leptokurtic distribution. These results are, however, based on the null hypothesis of normality and provide no information for the parametric distribution of the series.

Table 1. Descriptive Statistics of Nigeria Stock Exchange Return from 1985-2023

| Descriptive Statistics | NSE Returns |
|------------------------|-------------|
| Mean | 5070.728 |
| Median | 1761.680 |
| Maximum | 20682.29 |
| Minimum | 4.820000 |
| Std. Dev. | 5707.306 |
| Skewness | 0.678003 |
| Kurtosis | 1.962428 |
| Jarque-Bera | 56.84852 |
| Probability | 0.000000 |
| Sum | 2373101. |
| Sum Sq. Dev. | 1.52E+10 |
| Observations | 468 |

4.1.2. Unit Root Test Results

Augmented Dicky Fuller and Phillips-Perron unit root test is employed in examining stationarity characteristics of the Nigeria's stock returns from January 1985 – December 2023. The results of Augmented Dicky Fuller and Perron unit root test for data at level and first differenced are presented in Tables 2A-D. The results of Augmented Dicky Fuller and Perron unit root test indicates that the market returns are indeed non-stationary. This is shown by the Augmented Dicky Fuller and Phillips-Perron test statistics being higher than their corresponding asymptotic critical values at 1% and 5% levels. However, the Phillips-Perron unit root test result of the stock returns show evidence of covariance stationarity as the test statistics are all smaller than their corresponding asymptotic critical values at all the designated test sizes both for constant only and for constant and linear trend. These results confirmed the result observed by visual inspection of time plots reported in Figures 4.

4.1.3. Correlogram Test for Heteroskedasticity

Table 3 presents the Ljung-Box Q-statistics for the heteroskedasticity. It is a test of the null hypothesis that the autocorrelation of the lags are equal to zero. The result indicated that the p-values of the five lags were less than the level of significance, an indication that the model does not exhibits a lack of fit.

Table 2A. Unit Root test Augmented Dicky Fuller Test (at Level)

| | t-Statistic | Prob.* |
|--|-------------|--------|
| Augmented Dickey-Fuller test statistic | -0.223961 | 0.9326 |
| Test critical values: | | |
| 1% level | -3.444562 | |
| 5% level | -2.867700 | |
| 10% level | -2.570115 | |

*MacKinnon (1996) one-sided p-values.

Table 2B. Phillips-Perron test statistic (At Level)

| | Adj. t-Stat | Prob.* |
|--|-------------|----------|
| Phillips-Perron test statistic | -0.760865 | 0.8285 |
| Test critical values: | | |
| 1% level | -3.444128 | |
| 5% level | -2.867509 | |
| 10% level | -2.570012 | |
| *MacKinnon (1996) one-sided p-values. | | |
| Residual variance (no correction) | | 728495.5 |
| HAC corrected variance (Bartlett kernel) | | 336584.3 |

Table 2C. Augmented Dicky Fuller Test (at First Difference)

| | t-Statistic | Prob.* |
|--|-------------|--------|
| Augmented Dickey-Fuller test statistic | -5.871358 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.444691 | |
| 5% level | -2.867757 | |
| 10% level | -2.570145 | |

*MacKinnon (1996) one-sided p-values.

Table 2D. Phillips-Perron test statistic (At First Difference)

| | Adj. t-Stat | Prob.* |
|--|-------------|----------|
| Phillips-Perron test statistic | -35.84907 | 0.0000 |
| Test critical values: | | |
| 1% level | -3.444158 | |
| 5% level | -2.867522 | |
| 10% level | -2.570019 | |
| *MacKinnon (1996) one-sided p-values. | | |
| Residual variance (no correction) | | 534539.7 |
| HAC corrected variance (Bartlett kernel) | | 672978.2 |

Table 3. Correlogram and Ljung-Box Q-statistics Test for Heteroskedasticity

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob. |
|-----------------|---------------------|---|--------|--------|--------|-------|
| **** . | **** . | 1 | -0.509 | -0.509 | 121.59 | 0.000 |
| . *** | . * | 2 | 0.368 | 0.147 | 185.25 | 0.000 |
| ** . | * . | 3 | -0.306 | -0.099 | 229.52 | 0.000 |
| . ** | . * | 4 | 0.345 | 0.178 | 285.72 | 0.000 |
| ** . | * . | 5 | -0.311 | -0.067 | 331.56 | 0.000 |

Table 4. ARCH-LM TEST FOR HETEROSKEDASTICITY

| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|-----------|
| GARCH | 4.30E-08 | 4.83E-05 | 0.000889 | 0.9993 |
| Variance Equation | | | | |
| C | 0.007792 | 0.001100 | 7.085457 | 0.0000 |
| RESID(-1)^2 | 0.150400 | 0.019032 | 7.902623 | 0.0000 |
| GARCH(-1) | 0.917583 | 0.010511 | 87.29815 | 0.0000 |
| R-squared | -0.000039 | Mean dependent var | | -0.010299 |
| Adjusted R-squared | -0.000039 | S.D. dependent var | | 1065.738 |
| S.E. of regression | 1065.759 | Akaike info criterion | | 11.11966 |
| Sum squared resid | 5.30E+08 | Schwarz criterion | | 11.15512 |
| Log likelihood | -2598.001 | Hannan-Quinn criter. | | 11.13361 |
| Durbin-Watson stat | 2.097144 | | | |

4.1.4. Graphical Presentation of Nigeria Stock Exchange Returns

Figure 3 presents the pattern of level data on return series of the NSE for the period under review, which is January 1985 – December 2023. The level data show that the NSE return series is not stationary indicating the need for differencing. But the return series show sign of returning to its mean suggesting that the series are weakly stationary (See Figures 4). From Figure 5, we see that the NSE stock returns distribution is peaked confirming the evidence of non-normal

distribution in Table 1. Peaked distribution is a sign of recurrent wide changes, which is an indication of uncertainty in the return discovery process.

4.2. ARCH-LM Test for Heteroskedasticity

The result presented in Table 4 indicates result of the residual test of heteroskedasticity for ARCH effects. The test rejects the null hypothesis of no ARCH effects in the residuals of returns. This means that the errors are time varying and can only be modeled using heteroskedastic ARCH family models.

NSE

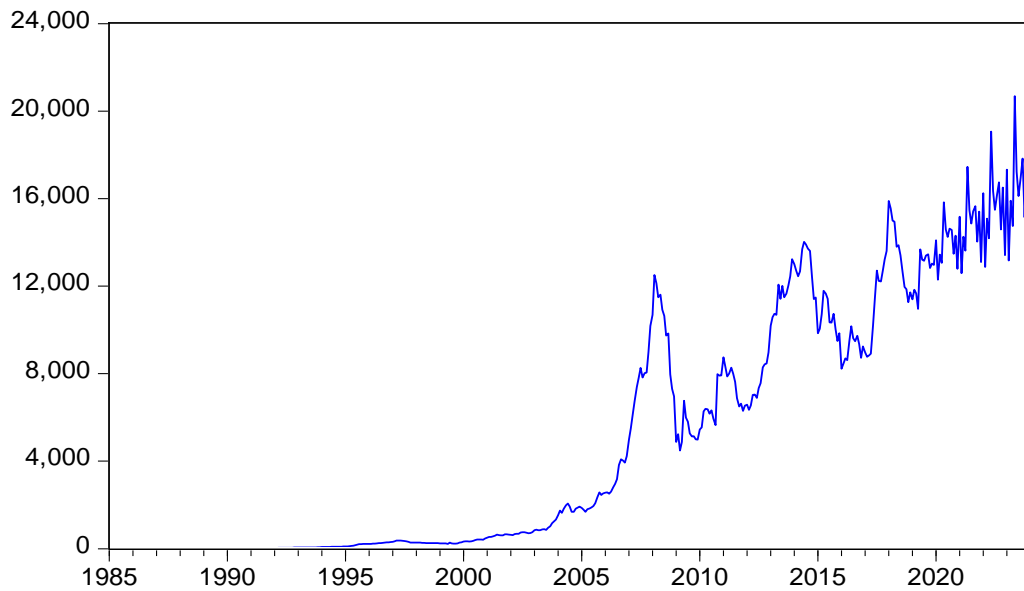


Figure 3. Line Graph of Monthly Index for NSE (Jan.1985 – Dec. 2023)

FIRST

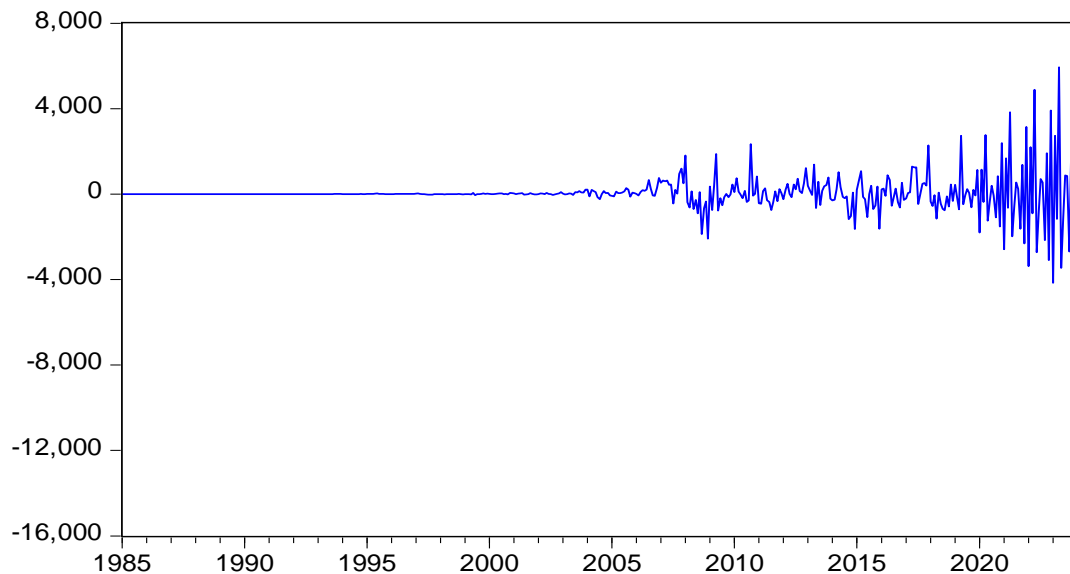


Figure 4. Line Graph of Monthly Index of NSE First Differencing (Jan.1985 – Dec. 2023)

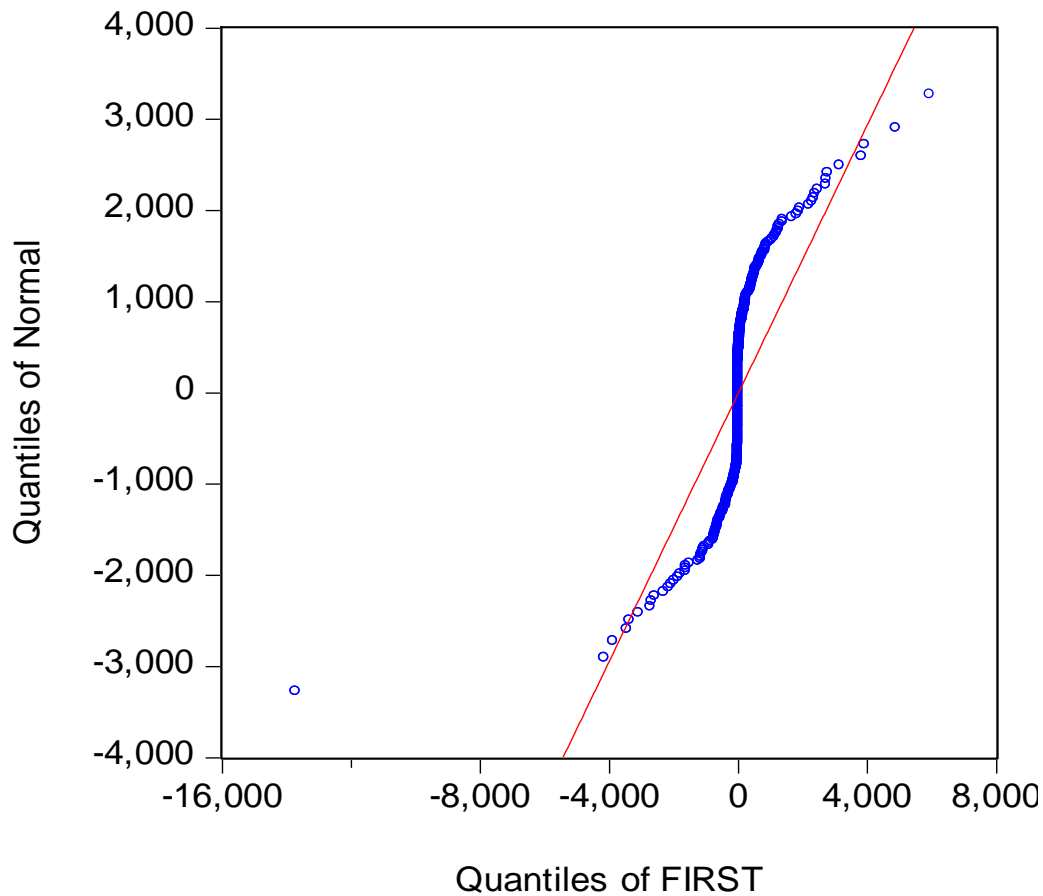


Figure 5. Quantile-Quantile Plot for the First Differenced NSE Monthly Returns (Jan. 1985-Dec. 2023)

4.3. Analysis of Data Using Asymmetric GARCH Type-Models

4.3.1. Results of EGARCH (1,1)

The parameter estimates of the asymmetric EGARCH (1, 1) model presented in Table 5 indicates that the coefficient (0.6432) of the ARCH effect is significantly positive at 1% showing that news about previous volatility (past square residual terms) has an explanatory power on current volatility. The GARCH effect coefficient (0.5789) is also significantly positive at 1% level, which shows that past volatility of stock market return is significantly influencing current volatility. The sum of ARCH and GARCH coefficients (1.2221) is a measure of the persistence of variance, and since the value is greater than 1 implies that there is significant high persistence in volatility. As suggested by Engle and Bollerslev (1986), if $\alpha_1 + \beta = 1$, a current shock persists indefinitely in conditioning the future variance. Since the sum of $\alpha_1 + \beta$ represents the change in response function of shocks to volatility persistence, a value greater than unity implies that response function of volatility increases with time and a

value less than unity implies that shocks decay with time (Chou, 1988). Therefore, the above result indicates that memory of shocks is remembered in the NSE.

4.3.2. Results of TGARCH (1,1)

The TGARCH (1,1) is an asymmetric model, which is used to investigate the existence of leverage effect in returns of the NSE. The estimates of the model are reported in Table 5. The asymmetric term in the TGARCH (1, 1) model is positive and significant suggesting the existence of leverage effect in returns during the study period. This implies that previous period's positive and negative shocks have a different effect on the conditional variance.

For the effect of the previous period's bad news to be greater than the effect of good news of the same magnitude, Gamma should be significant and have a negative sign. On the other hand, for the effect of the previous periods' positive news to be greater than the effect of bad news of the same magnitude, Gamma should be significant and have a positive sign. Table 6 are the estimates of TGARCH (1,1) model. The asymmetric effect of the TGARCH (1,1), captured by the parameter estimate γ , is positive and significant.

Table 5. EGARCH(1,1) MODEL FOR NSE (1985-2023)

| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|---------------------------|-------------|-----------------------|-------------|----------|
| Mean Equation | | | | |
| Constant | -0.315672 | 1.769082 | -0.4.3864 | 0.0005 |
| 1 st order Lag | -1.628064 | 5.742004 | -0.283536 | 0.0012 |
| 2 nd order Lag | 1.978368 | 2.440239 | 0.810732 | 0.0000 |
| 3 rd Order Lag | -0.279379 | 2.439418 | 0.451200 | 0.0032 |
| 4 th Order Lag | 0.997669 | 0.115867 | 8.610459 | 0.0000 |
| Variance Equation | | | | |
| Omega (constant) | -0.4532 | | | 0.0000 |
| Alpha (ARCH) | 0.6432 | | | 0.0000 |
| Gamma (Leverage) | 0.5789 | | | 0.0000 |
| Beta (GARCH) | 0.4325 | | | 0.0000 |
| R-squared | 0.89532 | Mean dependent var. | | 5070.728 |
| Adjusted R-squared | 0.87543 | S.D. dependent var. | | 5707.306 |
| S.E. of regression | 7629.943 | Akaike info criterion | | 16.31139 |
| Sum squared resid | 2.72E+10 | Schwarz criterion | | 16.34684 |
| Log likelihood | -3812.864 | Hannan-Quinn criter. | | 16.32534 |
| Durbin-Watson stat | 0.012553 | | | |

Table 6. TGARCH MODEL FOR NSE (1985-2023)

| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|---------------------------|-------------|-----------------------|-------------|----------|
| Mean Equation | | | | |
| Constant | -0.119089 | 1.854168 | -0.634242 | 0.0005 |
| 1 st order Lag | 0.762398 | 4.925836 | 0.514866 | 0.0000 |
| 2 nd order Lag | -0.088128 | 3.251636 | -0.278961 | 0.0000 |
| 3 rd Order Lag | 0.133848 | 2.964602 | 0.452149 | 0.0003 |
| 4 th Order Lag | 2.150102 | 2.972451 | 0.723653 | 0.0000 |
| Variance Equation | | | | |
| Omega (constant) | -0.5436 | | | 0.0000 |
| Alpha (ARCH) | 0.3221 | | | 0.0000 |
| Gamma (Leverage) | 0.5532 | | | 0.0000 |
| Beta (GARCH) | 0.0321 | | | 0.0000 |
| R-squared | 0.791056 | Mean dependent var. | | 5070.728 |
| Adjusted R-squared | 0.787229 | S.D. dependent var. | | 5707.306 |
| S.E. of regression | 7629.943 | Akaike info criterion | | 16.31563 |
| Sum squared resid | 2.72E+10 | Schwarz criterion | | 16.35996 |
| Log likelihood | -3812.858 | Hannan-Quinn criter. | | 16.33307 |
| Durbin-Watson stat | 0.012553 | | | |

4.4. Selection of the Best Fitting Asymmetric GARCH Model

To select the best fitting asymmetric GARCH models with suitable distributional assumption, information criteria such as Akaike information criterion (AIC) due to Akaike (1978) and Schwarz information criterion (SIC) due to Schwarz (1978) are employed in conjunction with log likelihoods (LogL). The best fitting model is one with largest log likelihood and minimum information criteria. Table 7 displays the asymmetric GARCH results of EGARCH (1,1) and TGARCH (1,1)

models. The information criteria together with the log likelihood optimally selects asymmetric EGARCH (1,1) as the best candidates to model the stock return volatility in Nigerian stock market.

Table 7. Asymmetric GARCH Model Order Selection

| S/N | Model | LogL | AIC | BIC |
|-----|-------------|-----------|----------|----------|
| 1. | EGARCH(1,1) | -3812.864 | 16.31139 | 16.34684 |
| 2. | TGARCH(1,1) | -3812.858 | 16.31563 | 16.35996 |

Table 8. Comparison of different architectures for the LSTM model with AIC

| Structure of LSTM Model | Train Data | Validation Data | Test Data |
|-------------------------|-----------------|-----------------|-----------------|
| 5 | -10774.8 | -4567.8 | -10551.4 |
| 10 | -10779.1 | -4566.6 | -10564.3 |
| 20 | -10787.5 | -4568.3 | -10517.3 |
| 40 | -10799.3 | -4573.9 | -10484.6 |
| 80 | -10789.1 | -4575.5 | -10468.8 |
| 160 | -10584.3 | -4568.4 | -10484.7 |
| 160-5 | -10667.4 | -4565.2 | -10323.4 |
| 160-10 | -10702.6 | -4576.2 | -10379.8 |
| 160-20 | -10730.4 | -4580.1 | -10581.5 |
| 160-40 | -10656.8 | -4582.3 | -10674.6 |
| 160-80 | -10737.6 | -4578.5 | -10565.7 |
| 160-80-5 | -10790.2 | -4575.3 | -10584.0 |
| 160-80-10 | -10732.9 | -4588.8 | -10436.2 |

Table 9. Comparison of Different Architectures for the LSTM Model with BIC

| Structure of LSTM Model | Train Data | Validation Data | Test Data |
|-------------------------|-----------------|-----------------|-----------------|
| 5 | -10764.6 | -4534.8 | -10668.9 |
| 10 | -10769.4 | -4542.6 | -10685.5 |
| 20 | -10777.3 | -4557.3 | -10646.4 |
| 40 | -10779.2 | -4542.9 | -10636.1 |
| 80 | -10769.2 | -4552.5 | -10677.6 |
| 160 | -10564.9 | -4543.4 | -10695.8 |
| 160-5 | -10637.3 | -4546.2 | -10594.0 |
| 160-10 | -10712.2 | -4538.2 | -10540.3 |
| 160-20 | -10730.1 | -4541.1 | -10643.6 |
| 160-40 | -10651.4 | -4553.3 | -10597.5 |
| 160-80 | -10724.7 | -4562.5 | -10607.3 |
| 160-80-5 | -10781.7 | -4575.3 | -10696.8 |
| 160-80-10 | -10702.5 | -4545.8 | -10686.1 |

This study compares econometrics models such as GARCH, TARCH, and EGARCH with machine learning techniques such as LSTM. It is necessary to develop a good LSTM model, which significantly depends on the number of layers to be used and the number of nodes present in each layer. Since the LSTM model is compared with GARCH models, to have a fair comparison, the lag order of the ΔVIX_t series for the LSTM model is kept the same as that of the GARCH models, i.e., 4. By comparing the AIC and BIC values of training, validation and test data, the number of layers and nodes in each layer of the LSTM was manually determined. The AIC and BIC values for several layer and node combinations and chose the one that performed the best. ReLU activation function is used between the layers, and the optimizer used to minimize the training error is Adam because Adam has been proven to be the best among all the 1st order optimizers available for neural networks (Kandel *et al.* 2020). Table 8 and 9 present the AIC and BIC performance of the LSTM model using different combinations of layers and nodes, and the best combination for the ΔVIX_t series is

found to be three LSTM layers with 160, 80, and 5 nodes in each layer, respectively.

Table 10. Performance Comparison using RMSE Criterion

| Model | RMSE on one Month Data | RMSE on Two Months Data |
|-----------------|------------------------|-------------------------|
| EGARCH (1,1) | 2.677 | 6.077 |
| LSTM (160-80-5) | 5.112 | 9.544 |

4.5. Comparison of LSTM and GARCH Models Using out of Sample Forecasting Results

This study compares the performance of the models by forecasting the India VIX volatility from each model for the next two months in the future and then analyzing the results based on three performance metrics, i.e., mean directional accuracy (MDA), mean absolute percentage error (MAPE), and root mean squared error (RMSE). The reason for keeping three performance metrics is to compare the performance of models in different aspects; for example, MDA would tell the performance of models in predicting the direction of volatility, whereas RMSE would explain how far the predictions are from actual values. Table 10 presents the out-of-sample version of models (LSTM and GARCH models) on the MAPE metric in predicting the India VIX values for the next two month.

5. Discussion of Finding

Overall results from this study provide evidence to show that volatility of stock returns is persistent in Nigeria stock returns data. These results is in tune with the findings from (Enow, 2023) who revealed that stock market volatility will persist at least for some time from the ARCH and GARCH output results. Active traders and market makers need to adapt their strategies in response to the expected volatility persistence. Higher levels of persistence may call for adjustments such as widening stop-loss orders to accommodate larger price swings or using more extended timeframes to capture sustained trends. Portfolio managers may also opt for strategies that thrive in volatile market conditions such as breakout trading or mean reversion strategies. The results also support the evidences of volatility persistence in Nigeria stock exchange as provided by (Godfrey & Agwu, 2018; Ajayi *et al.*, 2019) and existence of leverage effects in Nigeria stock returns provided by Okpara and Nwezeaku (2009), but disagree with their conclusion that stock returns volatility is not quite persistent in Nigeria.

The result of this study also shows that asymmetric EGARCH (1,1) outperformed LSTM model based on the lowest AIC and BIC forecasting accuracy criteria for the Nigerian stock market volatility returns. This result agrees with findings from (Adams and Uchema, 2023) in a study that compared Exponential Generalized Autoregressive Conditional Heteroskedasticity with order $p=1$ and $q=1$, (EGARCH (1,1)) and Recurrent Neural Network (RNN) based on long short term memory (LSTM) model with the

combinations of $p = 10$ and $q = 1$ layers to model the volatility of Nigerian exchange rates. The results indicated that the Nigeria exchange rate volatility is asymmetric, and leverage effects are evident in the results of the EGARCH (1, 1) model. It was observed also that EGARCH (1,1) performed better than the LSTM model because it provided a smaller MSE values of 224.7, 231.3 and 138.5 for euros, pounds sterling and US Dollars respectively. The study by and The findings of this study is however contradicted by (Chatterjee, *et al.* (2022) that investigated stock volatility prediction using time series and deep learning methods. The study introduced several volatility models based on the Exponential general autoregressive conditional heteroscedasticity (EGARCH), Glosten-Jagannathan-GARCH (GJR-GARCH), and generalized autoregressive conditional heteroscedasticity (GARCH) frameworks. The three types of GARCH models and the LSTM model were tested for their ability to predict stock volatility in three distinct industries. The findings revealed that the LSTM model performed better in predicting volatility in the pharmaceutical sector than in the banking and IT sectors.

6. Conclusions

This study investigated the volatility of stock market returns in Nigeria using EGARCH (1,1), TGARCH (1,1) and the LSTM models. It also examined asymmetric characteristics in the volatility of the Nigerian stock market returns series from January 1985, to December 2023 and develop a robust forecast model that can properly and timely notify Nigerian stockbrokers of fluctuations in the market. The asymmetric GARCH results of EGARCH (1,1) and TGARCH (1,1) models indicated that asymmetric EGARCH (1,1) is the best candidate model for stock return volatility in Nigerian stock market. The performance of the LSTM model using different combinations of layers and nodes, and the best combination for the ΔVIX_t series discovered that three LSTM layers with 160, 80, and 5 nodes in each layer, respectively were the best among the other combinations. The results from EGARCH (1,1) and TGARCH (1,1) models indicated that volatility of stock returns is persistent in Nigeria. The result of the forecasting accuracy of the LSTM and asymmetric GARCHs model based on the Nigerian stock market volatility returns indicated that EGARCH (1,1) outperformed LSTM (160-80-5) model based on the lowest AIC and BIC. EGARCH (1,1) was found to be the appropriate model for forecasting Nigerian stock market volatility returns. Based on the findings of this study, it is recommended that;

- (i) Stockbrokers should invest regularly, be consistent in trading, maintain a diversified portfolio and seek professional advice. The strategies will curb the persistent volatility in Nigeria stock return.
- (ii) In the comparison between econometric asymmetric GARCH and deep learning's LSTM, the EGARCH (1, 1) performs better than the LSTM model.
- (iii) for modeling and predicting the volatility of Nigeria

stock return, EGARCH (1,1) provides the best statistical method.

REFERENCES

- [1] Adams, S. O., & Uchema, J. I. (2024). Nigeria Exchange Rate Volatility: A Comparative Study of Recurrent Neural Network LSTM and Exponential Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Artificial Intelligence and Big Data*, 4(2), 61–73. <https://doi.org/0.31586/jaibd.2024.983>.
- [2] Adams, S.O., Zubair, M.A. & Ezike, M.F. (2023). Mathematical Modeling of the Price Volatility of Maize and Sorghum between 1960 and 2022, *Journal of Mathematical Letters*, 1(1), 1–19. <https://doi.org/10.31586/jml.2023.801>.
- [3] Adewale, A. O., Olufemi, A. P. and Oseko, M. S. (2016). Modelling volatility persistence and asymmetry with structural break: Evidence from the Nigerian stock market, *Journal of Economics and Behavioral Studies*, 8(6), 153-160.
- [4] Ajayi, A., Adams, S.O., Akano, R.O. (2019). Modelling Nigeria Naira Exchange Rate against some selected Country's Currencies Volatility: Application of GARCH Model, *Asian Journal of Probability and Statistics*, 5(1), 1 – 13. 582-0230. <https://doi:10.9734/AJPAS/2019/v5i130128>.
- [5] Akaike, H., (1973). Information Theory and Extension of the Maximum Likelihood Principle, In: B.N. Petrov and F. Csaki (eds), *International Symposium on Information Theory*. 2nd Edition. Budapest: Akademiai Kiado, 267-87.
- [6] Aimin Z, Rui G. (2021). Research on Volatility Prediction Based on Deep Learning Algorithm Example of SSE 50ETF Option, *Huabei Finance*, 09 (07).
- [7] Asemota, O.J. and Ekejiuba, U.C. (2017). An Application of Asymmetric GARCH Models on Volatility of Banks Equity in Nigeria's Stock Market, *CBN Journal of Applied Statistics (JAS)*, 8(1), 73-99. <https://dc.cbn.gov.ng/jas/vol8/iss1/4>.
- [8] Bartholomew, D.C., Orumie, U.C., Obite, C.P., Duru, B.I. and Akanno, F.C. (2021). Modeling the Nigerian Bonny Light Crude Oil Price: The Power of Fuzzy Time Series, *Open Journal of Modelling and Simulation*, 9(4), 370–3900. <https://doi.org/10.4236/ojmsi.2021.94024>.
- [9] Black, F. (1976). Studies of stock price volatility changes, *Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section. American Statistical Association*, Washington, DC 177181.
- [10] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, 307-327.
- [11] Cai, W., Liu, K. and Lu, K. (2023). Forecasting Daily Stock Volatility: A Comparison between GARCH and Recurrent Neuro-networks, *BCP Business & Management*, 38, 427-438.
- [12] Charalambous, C., Charitou, A. and Kaourou, F. (2000). Comparative Analysis of Artificial Neural Network Models: Application in Bankruptcy Prediction, *Annals of Operations Research*, 99, 403-425. <https://doi.org/10.1023/A:1019292321322>.
- [13] Charef, F. and Ayachi, F. (2016). A Comparison between

- Neural Networks and GARCH Models in Exchange Rate Forecasting, *International Journal of Academic Research in Accounting, Finance and Management Sciences*, 6(1), 94-99. <https://dx.doi.org/10.6007/IJARAFMS/v6-i1/1996>.
- [14] Chatterjee, A., Bhowmick, H. and Sen, J. (2022). Stock Volatility Prediction using Time Series and Deep Learning Approach, *Computational Finance*, <https://doi.org/10.48550/arXiv:2210.02126>.
- [15] Chen, H., Zhang, J., Tao, Y. and Tan, F. (2019). Asymmetric GARCH Type Models for Asymmetric Volatility Characteristics Analysis and Wind power Forecasting, *Protection and Control of Modern Power Systems*, 4(29), 1 – 11. <https://doi.org/10.1186/s41601-019-0146-0>.
- [16] Cho, K., Bart Van M., Caglar G., Dzmitry B., Fethi B., Holger S., and Yoshua B. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation, *Paper presented at the EMNLP 2014 Conference on Empirical Methods in Natural Language Processing*, Doha, Qatar, October 25–29, 1724–1734.
- [17] Chou, R.Y. (1988). Volatility Persistence and Stock Valuations: Evidence Using GARCH, *Journal of Applied Econometrics*, 3(4), 279 – 294. <https://www.jstor.org/stable/2096644>.
- [18] Diebold, F.X. and Inoue, A. (2001). Long Memory and Regime Switching, *Journal of Econometric*, 105 (1), 131-159. [https://doi.org/10.1016/s0304-4076\(01\)00073-2](https://doi.org/10.1016/s0304-4076(01)00073-2).
- [19] Ding, Z., Granger, C.W.J., Engle, R. F. (1993). A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, 1(1), 83 – 106. [https://doi.org/10.1016/0927-5398\(93\)90006-D](https://doi.org/10.1016/0927-5398(93)90006-D).
- [20] Donaldson R G, and Kamstra M. (1996). Forecast combining with neural networks, *Journal of Forecasting*, 15(1), 49-61.
- [21] Emenike, K.O. (2010). Modelling Stock Returns Volatility in Nigeria Using GARCH Models, *Munich Personal RePEc Archive (MPRA)*. <https://mpra.ub.uni-muenchen.de/22723/>.
- [22] Enow, S.T. (2023). Forecasting Volatility Persistence: Evidence from International Stock Markets Dinasti, *International Journal of Economics, Finance and Accounting*, 4(3), 383 – 389. <https://doi.org/10.38035/dijefa.v4i3>.
- [23] Engle R. F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the Variance of the United Kingdom inflation, *Econometrics* 50, 987-1008.
- [24] Falat, L., Stanikova, Z., Durisova, M., Holkova, B. and Potkanova, T. (2015). Application of Neural Network Models in Modelling Economic Time Series with Non-constant Volatility, *Procedia Economics and Finance*, 34,600-607. [https://doi.org/10.1016/S2212-5671\(15\)01674-3](https://doi.org/10.1016/S2212-5671(15)01674-3).
- [25] Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, 38, 34(105).
- [26] Graves, A., Mohamed, A., Hinton, A. (2013). Speech Recognition with Deep Recurrent Neural Networks, *arXiv*: 1303.5778.
- [27] Fischer, T., & Krauss, C. (2017). Deep learning with long short-term memory networks for financial market predictions. *FAU Discussion Papers in Economics*.
- [28] Godfrey, O.U. & Agwu, E.C. (2018). Volatility Clustering and Persistence of Shocks in the Nigeria Stock Market: Evidence from Univariate GARCH Models. *Research Journal's Journal of Finance*, 6(3), 1-11.
- [29] Goel, A., Goel, A.K. and A. Kumar, A. (2023). The Role of Artificial Neural Network and Machine Learning in Utilizing Spatial Information, *Spatial Information Research*, 31, 275–285, 2023. <https://doi.org/10.1007/s41324-022-00494-x>
- [30] Hsieh, D.A. (1991). Chaos and Nonlinear Dynamics: Application to Financial Markets. *The Journal of Finance*, 46, 1839 – 1877. <https://doi.org/10.1111/j.1540-6261.1991.tb04646.x>.
- [31] Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.* 9, 1735–80.
- [32] Hossain, A. and Nasser, M. (2008). Comparison of GARCH and Neural Network Methods in Financial Time Series Prediction. *Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008)*.25-27 December, 2008, Khulna, Bangladesh.
- [33] Huang, X., You, P., Gao, X. and Cheng, D. (2023). Stock Price Prediction Based on ARIMA-GARCH and LSTM, *J. ICBIS*, 14, pp. 438–448. https://doi.org/10.2991/978-94-6463-198-2_45.
- [34] Idowu, P., Osakwe, C., Kayode, A. and Adagunodo, E. (2012). Prediction of Stock Market in Nigeria Using Artificial Neural Network, *International Journal of Intelligent Systems and Applications*, 4, 68-74. <https://doi.org/10.5815/ijisa.2012.11.08>.
- [35] Kala, R. (2021). An Introduction to Machine Learning and Deep Learning, *Emerging Methodologies and Applications in Modelling*, 569-625. <https://doi.org/10.1016/B978-0-443-18908-1.00022-4>.
- [36] Kandel, I., Mauro Castelli, M. and Popović, A. (2020). Comparative Study of First Order Optimizers for Image Classification Using Convolutional Neural Networks on Histopathology Images, *Journal of Imaging*, 6(9), 92. <https://doi.org/10.3390/jimaging6090092>.
- [37] Kuhe, D. A., and Chiawa, M. A. (2017). Modelling volatility of asset returns in Nigerian stock market: Applications of random level shifts models. *Asian Research Journal of Mathematics*, 7(4), 1-14.
- [38] Liu Y, Qin Z, Li P. (2017). Stock volatility prediction using recurrent neural networks with sentiment analysis. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, Springer, 192-201.
- [39] Luo, R., Zhang, W., Xu, X. and Wang, J. (2017). A Neural Stochastic Volatility Model, in *Proceeding of AAAI Conference on Artificial Intelligence*, 32(1), 10-16. <https://doi.org/10.1609/aaai.v32i1.12124>.
- [40] Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:394-414.
- [41] Mantri J. K., Gahan, P., and Nayak B. B. (2014). Artificial neural networks—an application to stock market volatility. *Soft-Computing in Capital Market, Research and Methods of Computational Finance for Measuring Risk of Financial Instruments*, 179.
- [42] Mohammed, T., Yahaya, H.U., Adams, S.O. (2022). Modeling the Volatility for Some Selected Beverages Stock Returns in Nigeria (2012-2021): A GARCH Model Approach. *Matrix, Science Mathematic (MSMK)*, 6(2), 41-50.

<https://doi.org/10.26480/msmk.02.2022.41.50>.

[/stable/2958889](#).

- [43] Nasrin, F.A., Quoc, B.P., Sajad, F.N., Mohammad, R., Chow, M., Ali, N.A. and E. Ahmed, E. (2020). Enhancing the Prediction Accuracy of Data-Driven Models for Monthly Streamflow in Urmia Lake Basin Based upon the Autoregressive Conditionally Heteroskedastic Time-Series Model, *Applied. Science*, 10(20), 571, 2020. <https://doi.org/10.3390/app10020571>.
- [44] Nelson D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347–370.
- [45] NSE (2023). All Share Index, Nigerian Stock Exchange. Retrieved from, www.nse.ng.org.
- [46] Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 1361-1401.
- [47] Olah, C. (2015). Understanding LSTM Networks. Retrieved from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- [48] Okpara, G.C. and Nwezeaku, N.C. (2009), "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns: Evidence from Nigeria", *European Journal of Economics, Finance and Administrative Sciences*, 17, 1-10.
- [49] Orabi, M.M.A., and Abed-Alkareem, T. (2015). Effect of Volatility Changes on Emerging Stock Markets: The Case of Jordan, *Journal of Management Research*, 7(4), 132 – 143. <https://doi.org/10.5296/jmr.v7i4.7463>.
- [50] Parasuraman, N., & Janaki Ramudu, P. (2011). Historical and Implied Volatility: An Investigation into NSE Nifty Futures and Options. *Australian Journal of Business and Management Research*, 1(7), 112-120.
- [51] Perron, P. (1990). Testing for a unit root in a time series with a changing Mean, *Journal of Business and Economic Statistics*, 8, 153-162.
- [52] Shaik, M. and Aditya (2020). The Comparison of GARCH and ANN Model for Forecasting Volatility: Evidence based on Indian Stock Markets: Predicting Volatility using GARCH and ANN, *The Journal of Prediction Markets*, 14(2), 103-121. <https://doi.org/10.5750/jpm.v14i2.1843>.
- [53] Shen, Z., Wan, Q. and Leatham, D.J. (2021). Bitcoin Return Volatility Forecasting: A Comparative Study between GARCH and RNN. *Journal of Risk and Financial Management* 14(337), 1-28. <https://doi.org/10.3390/jrfm14070337>.
- [54] Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2) 461-464. <https://www.jstor.org/stable/2958889>.
- [55] Schwert, G.W. (1989). Why Does Stock Market Volatility Change Over Time? *The Journal of Finance*, 44(5), 1115 – 1153. <https://doi.org/10.1111/j.1540-5261.1989.tb02647.x>.
- [56] Su, Z., Xie H., Han L. (2021). Multi-factor RFG-LSTM Algorithm for Stock Sequence Crediting, *Computational Economics*, 57(4), 1041-1058.
- [57] Sullivan, (2019). Stock Price Volatility Prediction with Long Short-Term Memory Neural Networks, *IEEE*, 1-4.
- [58] Sundermeyer, M., Schluter, R. and Ney, H. (2012). LSTM Neural Networks for Language Modeling. *Proceedings of Interspeech*, 194-197. <https://doi.org/10.21437/interspeech.2012-65>.
- [59] Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2) 461-464. <https://www.jstor.org/stable/2958889>.
- [60] Taylor, S. J. (1987). Forecasting the Volatility of Currency Exchange Rates, *International Journal of Forecasting*, 3(1), 159-170.
- [61] Yahaya, H. U., Oyinloye, J. S., & Adams, S. O. (2022). Modeling and Forecasting Cryptocurrency Returns and Volatility: An Application of GARCH Models, *Universal Journal of Finance and Economics*, 2(1), 71–90. <https://doi.org/10.31586/ujfe.2022.497>.
- [62] Gil-Alana, L.A., Yaya, O.S. and Shittu, O.I. (2015). GDP Per Capita in Africa before the Global Financial Crisis: Persistence, Mean Reversion and Long Memory Features, *CBN Journal of Applied Statistics (JAS)*, 6(1b), 219-239.
- [63] Yu, W. (2018). Non-Linear System Modeling Using LSTM Neural Networks, *IFAC-PapersOnline*, 51(13), 485-489. <https://doi.org/10.1016/JIFACOL.2018.07.326>.
- [64] Yu, S., and Li, Z. (2018). Forecasting Stock Price Index Volatility with LSTM Deep Neural Network. In: Tavana, M., Patnaik, S. (eds) *Recent Developments in Data Science and Business Analytics*. Springer, 265–272. https://doi.org/10.1007/978-3-319-72745-5_29.
- [65] Zahid, M., Iqbal, F. and D. Koutmos, D. (2022). Forecasting Bitcoin Volatility Using Hybrid GARCH Models with Machine Learning, *Risks*, 10, 237-244. <https://doi.org/10.3390/risks10120237>.
- [66] Zakoian, J. M. (1994). Threshold Heteroscedastic Models, *Journal of Economic Dynamics and Control*, 18, 931-955.