2021

Stock market volatility forecasting at the Nairobi Securities Exchange: a comparison between asymmetric GARCH models and neural networks.

Rugut, Diana Chemutai
Strathmore Institute of Mathematical Sciences
Strathmore University

Stock Market Volatility Forecasting at the Nairobi Securities Exchange : A comparison between Asymmetric GARCH Models and Neural Networks.

Rugut Diana Chemutai*

122164

Supervisor: Dr. Samuel Chege Maina

A research dissertation submitted in partial fulfillment of the requirements for the Degree of Masters of Science in Mathematical Finance at Strathmore University

*diana.rugut@strathmore.edu

DECLARATION

I declare to the best of my knowledge that this work has not been previously submitted and approved for the award of a degree by this or any other university. To the best of my knowledge, this dissertation contains no material previously published by another person except where due reference is made.

Signature:.. Date: 23rd August 2021

RUGUT DIANA CHEMUTAI - 122164

This work has been presented with my authority as a University supervisor.

Signature:.. Date: 16th August 2021

DR. SAMUEL CHEGE MAINA

LECTURER, STRATHMORE UNIVERSITY

ACKNOWLEDGMENT

I thank the Almighty God for His provision, grace and care this far.

I am thankful to my supervisor, Dr. Samuel Chege Maina who devoted his time to guide and support me throughout the writing of my dissertation.

I am indebted to Strathmore University Institute of Mathematical Sciences(SIMS) fraternity for their continued support.

Lastly, I acknowledge the overwhelming support I received from my parents, siblings, family and my colleagues in the Msc. Mathematical Finance Class of 2019.

DEDICATION

This research is dedicated to our Almighty God for His strength and sufficient grace. To both my parents, Simon Rugut and Beatrice Rugut who have continued to provide their moral, spiritual and financial support. My siblings who have constantly believed in me and prayed for me. You have all been a source of inspiration and I am forever grateful.

Abstract

In this study, we conducted a comparative study on the volatility forecasting models focusing on the asymmetric GARCH models and the Artificial Neural Networks. The study focused on the Nairobi Securities Exchange and used the daily data from the NSE-20 Share Index for a period between January 2012 - February 2021. We took advantage of the year 2020 to investigate whether the models are still effective or they break during a period of turbulence.The parameters of the asymmetric GARCH models were estimated using Maximium Likelihood Estimation(MLE). The asymmetric GARCH models, namely the EGARCH and GJR-GARCH models were compared using AIC and BIC to determine the model with the best fit when it comes to modelling and based on the results of the AIC and BIC, the EGARCH was the better model. The study then compared the forecasting ability of the asymmetric GARCH models to that of the Artificial Neural Networks in terms of the standard deviations as a measure of the actual volatility to determine the better model. The RMSE and MAE measures were used to evaluate the forecasting abilities of the models and based on the results of the RMSE and MAE from the study, the Artificial Neural Networks outperformed the asymmetric GARCH models when it comes to forecasting.

List of Abbreviations

GARCH	Generalized Autoregressive Conditional Heteroskedasticity
ANN	Artificial Neural Networks
NSE	Nairobi Securities Exchnage
ARCH	Autoregressive Conditional Heteroscedasticity
EWMA	Exponentially Weighted Moving Average
GARCH-M	Generalized Autoregressive Conditional Heteroskedasticity-in Mean
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
GJR-GARCH	Glosten, Jagannathan and Runkle-Generalized Autoregressive Conditional Heteroskedasticity
EGARCH-M	Exponential Generalized Autoregressive Conditional Heteroskedasticity-in Mean
ADF	Augmented Dickey Fuller
FNN	Feed Forward Neural Network
RNN	Recurrent Neural Network
LSTM	Long-Short Term Memory
MLE	Maximum Likelihood Estimation
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
RMSE	Root Mean Square Error
MAE	Mean Absolute Error

List of Figures

List of Tables

Table of Contents

1 Introduction

1.1 Background of the Study

1.1.1 Market Volatility

Measuring and predicting stock market volatility has received growing attention from both academics and practitioners over the last years. Market volatility tries to explain the uncertainties of the financial market based on the changes in the values of the securities. It is measured by the standard deviation, which measures the variability around an average. The larger the standard deviation, the higher the volatility, and vice versa. Volatility forecasting is crucial in risk management, asset valuation and investment management for investors and financial market participants. For this reason this area has attracted a lot of research in order to identify good models that are able to predict volatility.

GARCH models are popular in literature when it comes to forecasting volatility because the model is able to capture time-varying volatility. The problem with the standard GARCH model is that it may not predict the distributions effectively and this is because the relationship between the current and previous variance may not be linear. The asymmetric GARCH models have since been developed so as to capture the asymmetries in volatility that the standard GARCH model was not able to capture.

The GARCH models have underlying assumptions such as normality on error distributions. Artificial Neural Networks have been developed to predict and define relationship between variables without relying so much on such assumptions. Artificial Neural Network have the capability of learning and modelling non-linear and complex relationships which is important.

This study will focus on the assymetric GARCH models and the Artificial Neural Networks where we will compare the forecasting ability of the assymetric GARCH models and the Artificial Neural Networks and examine which model is a better fit at forecasting volatility.

1.1.2 Overview of the Nairobi Securities Exchange

The Nairobi Securities Exchange which was formerly known as the Nairobi Stock Exchange was registered as a voluntary association of stockbrokers in 1954 under the Societies Act and was responsible for developing the securities market and regulating trading activities. The three main types of indices in the NSE market are;

1. NSE 20-Share Index which is designed to represent the geometric mean prices of the top 20 companies listed in the NSE.

2. NSE 25-Share Index which is designed to represent the performance of Kenyan companies listed in the NSE.

3. NSE All Share Index which is designed to represent all companies listed in the NSE, therefore being an overall indicator of the markets performance.

The NSE also plays an important role in the growth of Kenya's economy as it encourages savings and investments.

1.2 Problem Statement

Forecasting of market volatility is crucial in the financial markets as it is important in the areas of risk management, pricing, and among others. It is therefore necessary to identify a model that is efficient and approximate at forecasting the market volatility. The research available in the Kenyan market is limited. Research has been done to investigate whether there are stylized facts in the Kenyan market. However, their is limited evidence in the models used when forecasting volatility. For instance whether the asymmetric GARCH model is a better model compared to the symmetric model and also the use of the neural networks techniques in stock market volatility forecasting. This research will try to fill the literature gap on the asymmetric GARCH models and the Neural Networks techniques as well as investigate which of the models can accurately forecast volatility.

1.3 Objectives

The main objective of the study is to forecast stock market volatility at the Nairobi Securities Exchange. Some of the specific objectives of the study will include;

1. To investigate the best fit asymmetric GARCH model.

2. To perform a comparative analysis between the best fit asymmetric GARCH model and the Artificial Neural Networks.

1.4 Significance of the Study

The study will;

- Illustrate which model is more accurate at forecasting the stock market volatility and this will be beneficial to the financial market players.

- Contribute to the limited literature available for forecasting stock volatility in emerging and developing markets.

2 Literature Review

2.1 Introduction

The efficient market hypothesis is closely associated with Fama[20] when he stated that a market in which the security prices "fully reflect" all available information is an "efficient market" which implies that the successive price changes are independent and identically distributed. The notion that the market is efficient is associated with a random walk where new information reflects randomly on the current stock price which will ultimately have an effect on the future stock price. Fama[20] also categorized the efficient market hypothesis into three tests; weak form test which studies information of historical prices, semi-strong form test which studies the speed of price adjustment to publicly available information, and strong form test which studies monopolistic access to information about prices. From this he notes that there is no important evidence against the weak and semi-strong form, and only limited evidence against the strong form.

According to Thompson and Ward [28], econometric models have been used before when testing the efficient market hypothesis and since some of the models test for dependency of the share prices, it invalidates the efficient market hypothesis.

Fama [12] showed that financial time series are known to display stylized facts which are critical to model estimation and forecasting. Sewell [27] defines stylized fact as a term in economics used to refer to empirical findings that are so consistent across markets that they are accepted as truth. To get reliable future forecasts of volatility, it is necessary to account for stylized facts. Examples of the stylized facts include; volatility clustering, fat-tails, leverage effects, long memory, skewness and leptokurtic.

2.2 Empirical Evidence

2.2.1 GARCH models

Engle [11] noted the importance of modelling time-varying second-order moments and he proposed a model where the variance depended on past realizations; the ARCH model which recognized the difference between conditional and unconditional variance, and allowed the conditional variance to change overtime as a function of past errors [7] . Bollerslev [7] noted that the ARCH model, in taking account of the long memory had a declining lag structure in the conditional variance and because of this he modified the ARCH model to GARCH model by allowing a flexible lag structure.

Forecasting market volatility has attracted a lot of research and for this reason there has been many applications of the GARCH models.

Akgiray [1] presented new evidence on the time-series behavior of stock prices using US stock data and found that forecasts done using the GARCH model were found to be superior.

Tse's [29] study on forecasting volatility in the Tokyo exchange market compared three models. Two of them being the GARCH model and the EWMA model. The study concluded that the EWMA model was more accurate than the GARCH model. Kuen and Hoong [18] had a similar study in the Singapore stock market and also concluded that the EWMA model outperformed the GARCH model.

Awartani et. al [6] predicted the volatility of the S& P-500 stock index via GARCH models and from their study, the asymmetric GARCH model performed better than the symmetric GARCH models. Alberg, Shaliti and Yosef [3] validated this when they estimated stock market volatility using asymmetric GARCH models and concluded that the asymmetric GARCH model outperformed the symmetric GARCH models. Different studies have since been done to compare the GARCH model with the asymmetric GARCH models and the studies found that the asymmetric GARCH models outperformed the GARCH model.

Gokcan [13] states that stock market return displays skewed distribution which the liner GARCH model is not able to manage. Gokcan's [13] study was focused on forecasting volatility of emerging stock market and compared both the linear and non-linear GARCH models. The study concluded that in emerging markets, the linear model outperformed the non-linear model despite the stock market return displaying skewed distribution.

Alagidede, P., and Panagiotidis [2] modelled stock returns in Africa's emerging markets and from their study, they tested and rejected the random walk hypothesis. The study used GARCH, GARCH-M and EGARCH-M to model conditional variance and found that volatility clustering, leptokurtosis and leverage effects were present in the African stock index returns. Their study also noted that there was not enough evidence to reject the weak form efficiency in these markets.

Omar [25] studied GARCH modeling in monthly foreign exchange and share prices for specific companies in Kenya and noted that GARCH(1,1) successfully captures volatility clustering and can be modified to allow other stylized facts. Korir [17] validates this in her research where she concluded that the GARCH(1,1) fully captures the stylized characteristics of financial time series.

Ogum et. al [24] investigated the emerging equity market volatility in the Kenyan and Nigerian markets using the EGARCH model. They found that the asymmetric volatility found in the U.S and other developed countries was also identified in the Nigerian Stock Market however, the asymmetric volatility in the Kenyan market was positive indicating that positive shocks in the market had more impact on the volatility as opposed to negative shocks of the same magnitude.

Maqsood et. al [21] modeled stock market volatility at the Nairobi Securities Exchange using GARCH models and concluded that the asymmetric models performed better than the symmetric models.

2.2.2 Artificial Neural Networks

John McCarthy coined the term artificial intelligence in 1955 which is about creating machines that demonstrate the intelligence that is usually displayed by humans [16]. There are different branches in artificial intelligence namely; machine learning, neural networks, robotics, fuzzy logic, expert systems and natural language processing.

Dixon et. al [9] in there book stated that machine learning can be viewed as a highly efficient data compression technique designed to provide predictors in complex settings where relations between input and output variables are non-linear and input space is often high-dimensional. In addition they stated that machine learning tries to balance filtering data with the goal of making accurate and robust decisions, often discrete and as a categorical function of input data. Dixon et. al [9] noted that this differs from maximum likelihood estimators used in standard statistical models, which assumes that the data was generated by the model and typically have difficulty with over-fitting, especially when applied to high-dimensional datasets.

The study will focus on the neural networks which is a set of algorithms that are used to find the elemental relationships across the bunches of data via the process that imitates the human brain operating process. A neuron is a mathematical function that gathers and classifies information according to a particular structure and implements various techniques to accomplish a task. It has been used in market research, risk analysis, stock market prediction among others [23].

Neural Networks are valuable in forecasting because; they are capable of performing nonlinear modeling without priori knowledge of the relationship between input and output variables, and it can generalize making it a more general and flexible modeling tool for forecasting [31]. In addition, many researchers have combined the neural networks with the GARCH models in their forecasting volatility study. [22] [30]

Donaldson and Kamstra [10]investigated the usefulness of a semi-nonparametric GARCH model to capture nonlinear relationships, their study evaluated its ability to forecast stock return volatility in London, New York, Tokyo and Toronto. Based on their results they concluded that the ANN model performs better than all the other model.

Atsalakis and Valavanis [5] surveyed stock market forecasting techniques with a focus on soft computing techniques. They surveyed 100 different published articles that applied the neural networks and neuro-fuzzy models in forecasting. From the articles surveyed, the study noted that the the models are suitable for stock market forecasting and that they outperformed the conventional models. Maciel et. al [19] analyzed neural networks for financial time series forecasting and compared it to the traditional GARCH models. Their study focused on the North American, European and Brazilian stock market and they concluded that the neural networks performed better than the GARCH models.

Bucci's [8] study focused on realized volatility forecasting using neural networks and concluded that the neural networks outperformed the traditional econometric methods. Arneric et. al [4] also compared the performance of the GARCH model to the Neural Networks and found that the Neural Networks performed better than the GARCH model. Harrilall [14] study focused on the South African stock market and found that neural networks had superior forecasting abilities as compared to the traditional forecasting models.

3 Methodology

3.1 Data Collection

The study will use secondary data from the Nairobi Securities Exchange. The daily data from the NSE 20 share index will be used in the study from January 2012-March 2019 to calibrate the models while the data from April 2019-February 2021 will be used for out of sample testing.

3.2 Data Analysis

The basic assumption in time series is that the underlying series is stationary in nature. The test for stationarity of the series under consideration will be done using the Augmented Dickey Fuller(ADF) test statistics. Pre-testing will be done on the data to test for normality using the Jacque-Bera test as the efficient market hypothesis implies that returns are normally distributed with zero skewness and a kurtosis of three.

The study will focus on two models for the analysis namely the asymmetrical GARCH models, and the Neural Networks technique for the purpose of evaluating the forecasting ability of the different models and later on compare the accuracy of the methods used. For the asymmetric GARCH models, the study will focus on the EGARCH and GJR-GARCH model.

3.2.1 Augmented Dickey Fuller Test

The Dickey Fuller test is a unit root test that tests the γ coefficient of the first lag of Y. The equation is as shown below,

$$y_t = \theta + \alpha t + \gamma y_{t-1} + \phi_1 \Delta y_{t-1} + \epsilon_t . \quad (1)$$

The Augmented Dickey Fuller test expands the Dickey-Fuller test by including higher regressive process in the model.

$$y_t = \theta + \alpha t + \gamma y_{t-1} + \phi_1 \Delta y_{t-1} + \phi_2 \Delta y_{t-2} + ... + \phi_{p-1} \Delta y_{t-p+1} + \epsilon_t , \quad (2)$$

where θ is a constant, α is the coefficient on a time trend and p is the lag order of the autoregressive process. This test is supposed to test the null hypothesis that a unit root is present in a time series data sample that is $\gamma = 0$ and the alternative hypothesis is that $\gamma < 0$. A unit root is a characteristic of time series that makes it non-stationary. The test statistic is computed and if the test statistic is less than the critical value at a confidence interval then the null hypothesis is rejected.

3.2.2 Jacque-Bera Test

The Jacque-Bera test is used to test for normality. Its a goodness of fit test to test whether the skewness and kurtosis of the data matches that of normal distribution.The test statistic of Jacque-Bera test is

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right), \tag{3}$$

where n is the number of observations, S is the sample skewness and K is the sample Kurtosis. The test is supposed to test the null hypothesis that the skewness of the data is zero and kurtosis of the data is also zero which is similar to a kurtosis of 3. Any deviation of this increases the Jacque-Bera statistic and thus rejecting the null hypotheis and concluding that the error terms are not normally distributed.

3.3 The Models

3.3.1 EGARCH Model

The Exponential Generalized Autoregressive Conditional Heteroscedasticity was developed by Nelson 1991 to capture the asymmetries in volatility. The natural logarithm of the condition variance is allowed to vary over time as a function of the lagged error terms rather than lagged squared errors. EGARCH(p,q) is specified as

$$\ln\sigma_t^2 = \omega + \sum_{j=1}^{p}\alpha_j\left|\frac{u_{t-j}}{\sqrt{\sigma_{t-j}^2}}\right| + \sum_{j=1}^{p}\gamma_j\frac{u_{t-j}}{\sqrt{\sigma_{t-j}^2}} + \sum_{i=1}^{q}\beta_i\ln\sigma_{t-i}^2, \tag{4}$$

The $\ln\sigma_t^2$ makes the leverage effect exponential rather than quadratic and this ensures that the estimates are non-negative. ω is a constant, α represents the ARCH effects, γ the asymmetric effects and β the GARCH effects.
The parameter γ is important when testing for asymmetries and the hypothesis is that if $\gamma_j > 0$ then positive shocks generates higher volatility than negative shocks and vice versa whereas if $\gamma_j = 0$ then the model is symmetric.

3.3.2 GJR-GARCH Model

The GJR-GARCH model was introduced by Glosten, Jagannathan and Runkle 1993 with the aim of capturing asymmetries in terms of negative and positive shocks. A multiplicative dummy variable is added into the variance equation to check whether there is statistically significant difference when shocks are negative. Unlike the EGARCH, the effect is captured in a linear manner.

$$\sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \gamma_i u_{t-i}^2 I_{t-i}\,, \tag{5}$$

where if $u_{t-i} > 0$

$$I_{t-i} = 1\,, \tag{6}$$

otherwise

$$I_{t-i} = 0\,, \tag{7}$$

γ is the asymmetry or leverage term. If it is significant and positive, then negative shocks will have larger effects on σ^2 than positive shocks. The condition for no negativity will be $\omega > 0, \alpha > 0, \beta \geq 0$, and $\alpha + \gamma \geq 0$. That is, the model is still admissible, even if $\gamma < 0$, provided that $\alpha + \gamma \geq 0$.

3.3.3 Neural Networks Technique

Neural Networks uses a non-parametric method of forecasting meaning that the underlying non-linear function is not prescribed and therefore the model is not limited to a restrictive list of non-linear functions. The most popular ANN has been the single-layer feed forward. In the FNN, information which is suitably weighted, is passed from a point of entry to a further layer of hidden neurons which is also assigned a weight, and finally reaches the output layer which represents the forecast. Let $p_t, t = 1,2,3,.....$ be the daily stock price. The daily returns are then computed by $r_t = \log(p_t) - \log(p_{t-1})$. The output y of a single layer FNN is then given by

$$y_t = S\left[a_0 + \sum_{i=1}^{a} a_i . g_{t,i}\right]\,, \tag{8}$$

where

$$g_{t,i} = G(b_{i0} + \sum_{j=1}^{n} b_{ij} r_{t-j} + \sum_{k=0}^{p} C_{ik} \Delta_{t-k}^{\frac{1}{2}}), (i = 1,2,......q)\,, \tag{9}$$

Function G and S use the tan-sigmoid function and the purelin function respectively. The tan-sigmoid functiona normalizes the values of each neuron to be in the interval (-1,+1) while the linear output layer lets the network produce values outside the range. Under the error back-propagation method, the neural network runs through all the input data over an initial “training” period and produces a list of outputs. The weights are then re-valuated using a recursive “gradient” descent method, so that the mean-squared error between the observed output and the predicted one is minimized.

The trained neural network is then applied over a different data set which is the "validation" period. This is done in order to evaluate the ability of the trained network to avoid over-fitting. In order to capture dynamics, the FNN includes lagged variables as explanatory variables. This problem is being addressed by the Recurrent Neural Networks which is a network with feedbacks from the hidden neurons, g, to the input layer with delay.

The RNN memorize this information since its output depends on both current and prior inputs.We will focus on the RNN with a single hidden layer and feedback connection from the output of the hidden layer to its input. Therefore in a RNN model equation (7) can be re-written as

$$g_{t,i} = G(b_{i0} + \sum_{j=1}^{n} b_{ij} r_{t-j} + \sum_{k=0}^{p} C_{ik} \Delta_{t-k}^{\frac{1}{2}} + \sum_{l=1}^{m} \delta_{il} g_{l,t-1}, (i = 1, 2, ...q)\,, \tag{10}$$

With back substitution it can be shown that the output y_t depends on the entire history of the inputs r and $\Delta h^{\frac{1}{2}}$.

The study will use the RNN model with LSTM(Long-Short Term Memory) technique. The LSTM model was introduced by Hochreiter and Schmidhuber [15] and functions through the use of operators that limit the follow of information in long sequences. The LSTMs have a separate cell memory C_t, in addition to a hidden state h_t. The cell memory is updated to include a forget gate, $\hat{\alpha}_t$, input gate $\hat{z}_t$ and cell gate $\hat{c}_t$

$$c_t = \hat{\alpha}_t c_{t-1} + \hat{z}_t \hat{c}_t\,, \tag{11}$$

According to [9],when the forget gate $\hat{\alpha}_t = 0$, then the cell memory depends solely on the cell memory gate update $\hat{c}_t$. By the term $\hat{\alpha}_t c_{t-1}$, the cell memory has long-term memory which is only forgotten beyond lag s if $\hat{\alpha_{t-s}} = 0$.

The hidden state h_t is defined by the following equation

$$h_t = \hat{r}\, tanh(c)_t\,, \tag{12}$$

If $\hat{r}_t = 0$ then the hidden state is reset otherwise if $\hat{r}_t = 1$ then the cell memory is determined by the hidden state.

3.4 Parameter Estimation

The parameters of the GARCH model will be estimated using the maximum likelihood method using the student *t* likelihoods. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed. The study will then use the AIC and BIC to select the model with the best fit. The study used both AIC and BIC because AIC is better in situations when a false negative finding would be considered more misleading than a false positive, and BIC is better in situations where a false positive is as misleading as, or more misleading than, a false negative.

3.5 Simulation

The parameters estimated will be used to simulate the model using data from January 2017-Decemeber 2020. The forecasting ability of the models will be judged on the basis of root mean square error(RMSE) and Mean Absolute Error(MAE).

The forecasts obtained from the assymetric GARCH models and the Neural Networks will be compared in terms of the standard deviation as a measure of the actual volatility and the results will show which model provides better volatility forecasts.

4 Data Analysis, Results and Discussions

4.1 Data Description

In this study, daily NSE20 share index data was used. The data obtained was from the Nairobi Securities Exchange website and it is for the period ranging between 3rd January 2012 to 22nd February 2021, a total of 2,273 observations.

4.2 Data Analysis

In the study we used the return term defined as follows

$$r_t = log\left(\frac{P_t}{P_{t-1}}\right),$$

where P_t is the price and r_t is the return.

The descriptive analysis of the data is described in Table 1 below;

Variables	Count	Mean	Std.Dev	Min	Max	Skewness	Kurtosis
P_t	2272	3661.3526	1007.5358	1723.9600	5499.6400	-0.1208	-0.9298
r_t	2272	-0.0236	0.7226	-8.6022	8.6344	-0.3755	21.8080

Table 1: Summary statistics of NSE 20 Share Index between January 2012-February 2021.

Table 1 above shows that the NSE 20 Share Index price achieved a high value of 5499.64 and a low value of 1723.96. The skewness of both the share index price and the return are non-zero and negative which indicates that it is not normally distributed and skewed to the left. The kurtosis of the share index price is less than 3 which indicates that it has a platykurtic distribution meaning that it has thinner tails than a normal distribution whereas that of the share index return is more than three which indicates that it has a leptokurtic distribution meaning that it has thicker tails than a normal distribution.

Studying the graphical representation of the share index price shown in Figure 1 for the period covered, we notice a persistent increase in price since 2012 to early 2015 followed by a persistent decrease until 2017, then an increase in price from 2017 to 2018 and then a persistent decrease since 2018 to early 2021.The highest price of the share index for the period covered was observed in 2015.

Figure 1: NSE20 Daily Prices

A series of daily returns was calculated from the daily prices of the share index. It can be observed from Figure 2 below, that the highest volatility was observed in 2014. In order to identify the unique characteristics of the share index, we plot a histogram of the daily returns as shown in Figure 3. The histogram shows that the returns of the share index for the period covered is slightly skewed to the left.

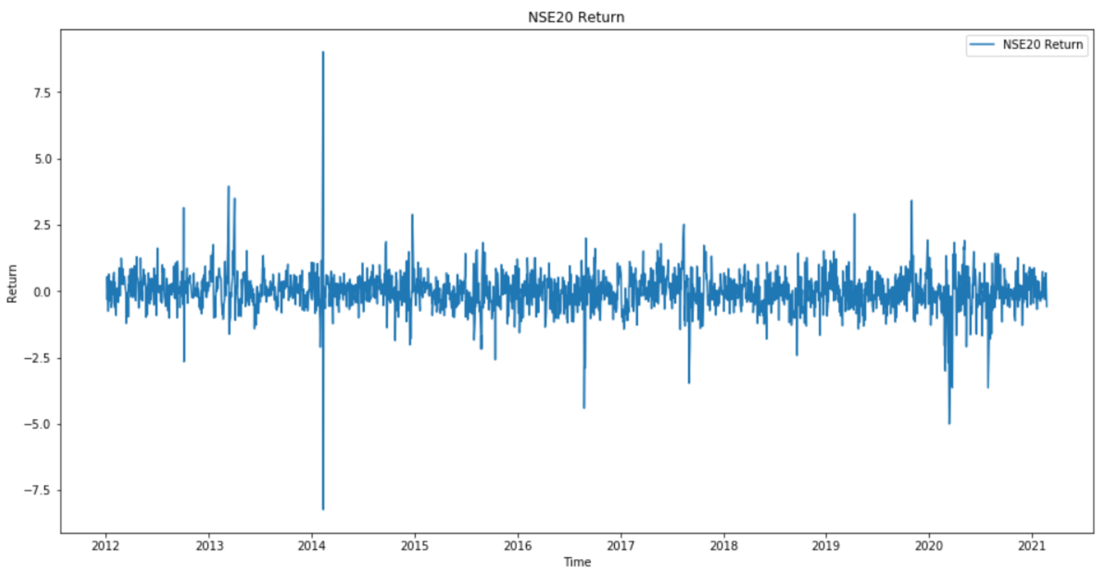

Figure 2: NSE20 Daily Returns

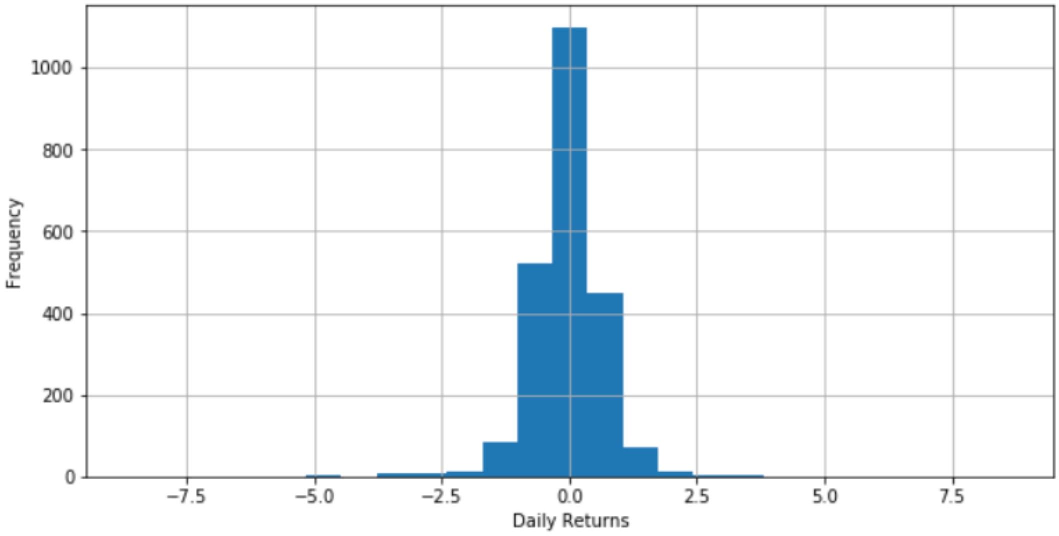

Figure 3: Histogram of Daily Returns

4.2.1 Results of Normality test

The Jarque-Bera test of goodness of fit to the normal distribution was used in this study. It tests whether the data has a skewness or kurtosis matching that of a normal distribution. The results for the Jarque-Bera test for the share index daily returns is shown in Table 2 below. Since the test statistic is far from zero, we reject the null hypotheis therefore concluding that the daily returns are not normally distributed.

Variable	Test-Statistic	P-Value
Jarque-Bera	44866.9458	0.0

Table 2: Jarque-Bera Test for Normality

4.2.2 Results of Stationarity test

After rejecting normality of the daily returns, this indicates that there may be presence of trend in the data and hence the test of stationarity was conducted using the Augmented Dickey Fuller test. The results of the stationarity test is summarized in Table 3 below.

		T-Statistics	P-Value
Augmented-Dickey Fuller	Test-Statistic	-19.811	0.0
Test Critical Values	1%	-3.4332	
	5%	-2.8628	
	10%	-2.5674	

Table 3: Augmented-Dickey Fuller Test for Stationarity

The null hypothesis when testing for stationarity is that a unit root is present in the time series data which is a characteristic that makes it non-stationary. Since the test statistic is less than the critical values at 1%, 5% and 10% confidence interval, then we rejected the null hypothesis and concluded that the dataset is stationary.

4.3 Empirical Results

4.3.1 Asymmetric GARCH Models

The study focused on two asymmetric GARCH models namely, the EGARCH models and the GJR-GARCH model. The EGARCH model was used on the daily returns of the NSE20 share index to capture the asymmetries in volatility and ensure that the estimates are non-negative. The GJR-GARCH on the other hand was used to test for leverage effect and the relationship between the stock market return and volatility. The summary of the results is as shown below in Table 4

Parameter\ Model	EGARCH	GJR-GARCH
ω	-0.2655	0.1657
α	0.4138	0.2968
γ	-0.0133	-0.0237
β	0.7210	0.3158
AIC	3243.21	3250.38
BIC	3276.24	3283.41
Log-likelihood	-1615.60	-1619.19

Table 4: Parameter Estimation Results

The leverage effect is denoted by γ. It is negative and statistically significant for the EGARCH model, therefore the null hypothesis is rejected which means that the negative shocks will generate higher volatility than positive shocks. For the GJR-GARCH, the γ is also negative and statistically significant. However based on the conditions for no negativity, then we conclude that the leverage effect for the GJR-GARCH model is positive and statistically significant hence the null hypothesis is accepted meaning that negative shocks will have larger effects on σ^2 than positive shocks, that is the higher the volatility at the NSE the lower the returns.

In order to compare the model with the best fit, the study used the AIC and BIC measures. Figure 4 and 5 show a graphical representation of the modelled results.

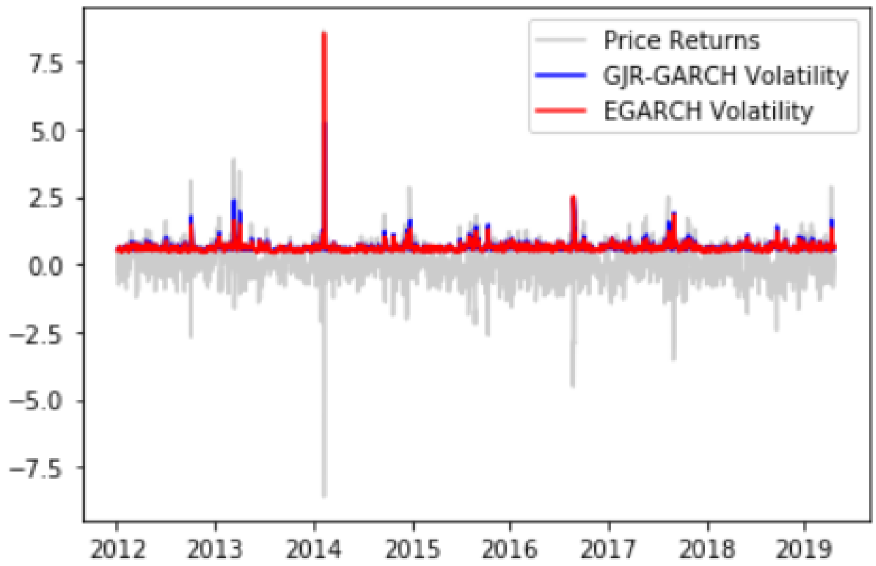

Figure 4: NSE 20 Daily Returns with EGARCH and GJR-GARCH Volatility

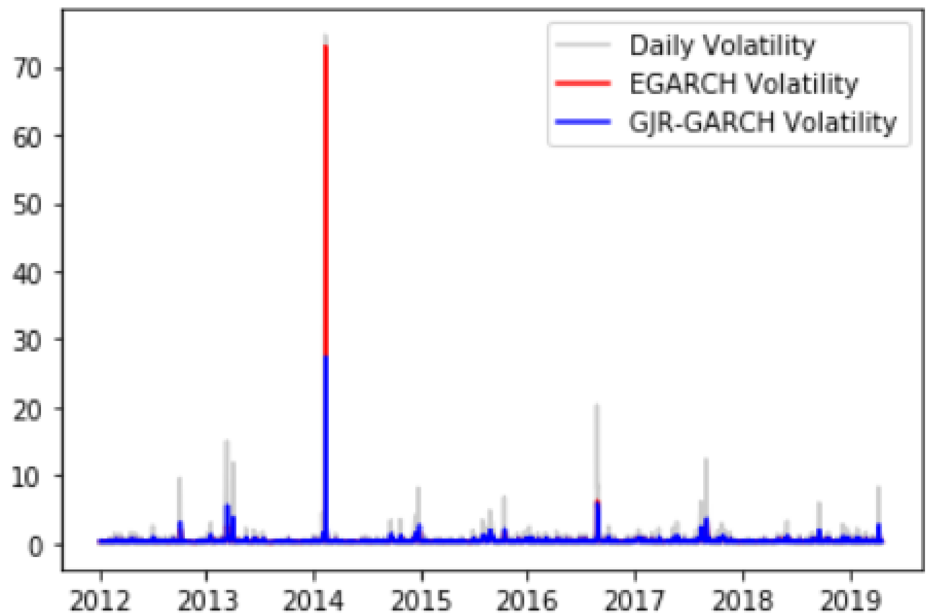

Figure 5: Daily Volatility Comparison with EGARCH and GJR-GARCH Volatility

Both EGARCH and GJR-GARCH are good at fitting the actual data. However, the EGARCH is a better model at fitting the actual data as it has a lower AIC and BIC value.

4.3.2 Artificial Neural Network Analysis

For this study, the Recurrent Neural Network was used because of its recurrent characteristics which works better at predicting time series than any other machine learning algorithm. The data was divided into 80% train and 20% test. From a total of 2,273 observations, 1,818 observations was used as a training set and 455 observations was used as a testing set. For our training data set of 1,818 observations, the model predicted the 1,819 observations and included it in the training dataset while simultaneously excluding the first data point in order to predict the 1,820 observation. This process continued until all the observations in the testing data set was met. The actual values in the testing dataset was then compared to the predicted values to validate the accuracy of the model.

After splitting the data to train and test ,the data was then scaled using the MinMaxScaler in order to train the data. The model was then configured with the number of neurons and timestamps.For optimization, the Adam optimizer was used and it performs gradient descent via backpropargation through time. This was to allow for quicker convergence at the cost of time. The MAE was used as a loss function to minimize the difference between the actual values and the predicted values.

Figure 6 and 8 below shows a graphical representation of the training data, the test predictions and the actual out of sample data.

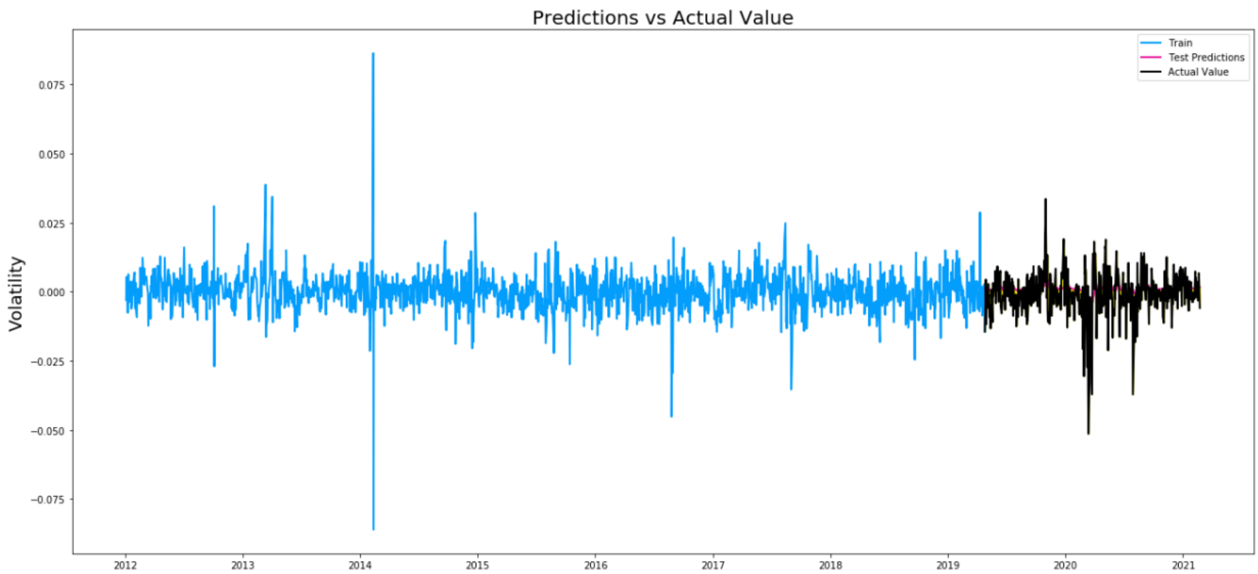

Figure 6: NSE20 Daily Returns-ANN Predictions

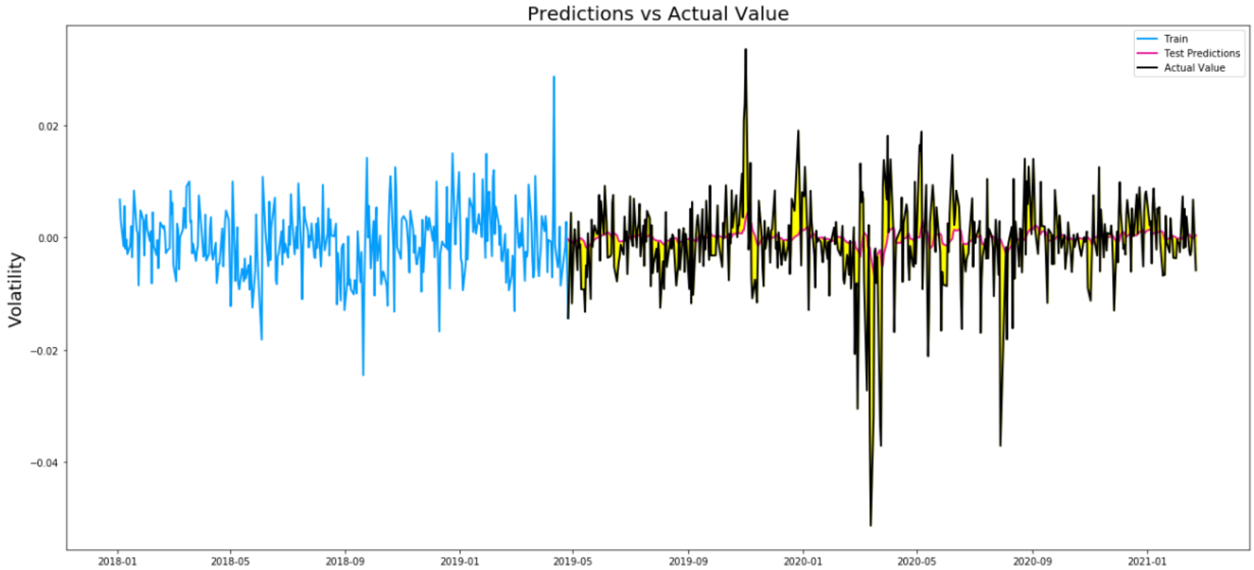

Figure 7: NSE20 Daily Returns-ANN Predictions Close-Up

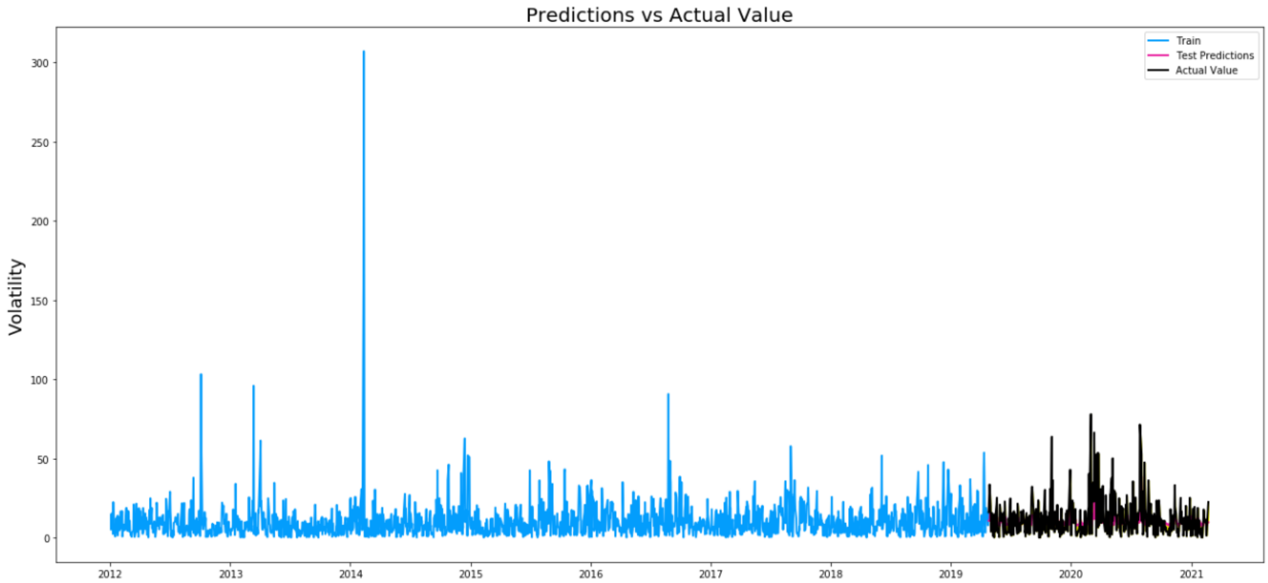

Figure 8: NSE20 Daily Volatility-ANN Predictions

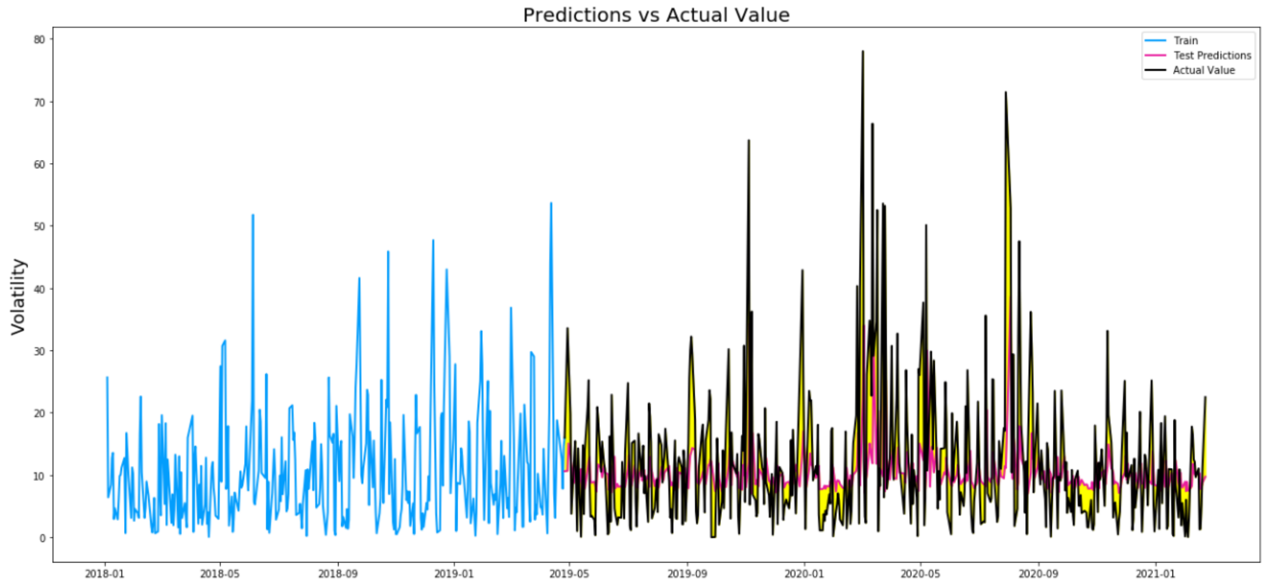

Figure 9: NSE20 Daily Volatility-ANN Predictions Close-Up

4.3.3 Forecast Evaluation

In addition to modelling volatility, the other objective of the study was to forecast volatility and compare which model is best at forecasting volatility between the EGARCH, GJR-GARCH and the Artificial Neural Networks. In order to identify the best model, the study used the MAE and RMSE measures.

The RMSE is the square root of the variance of the residual and it indicates how close the observed data points are to the model's predicted values. The MAE measures the average magnitude of the errors in a set of predictions, without considering their direction.

Model	MAE	RMSE
EGARCH	0.4714	0.0334
GJR-GARCH	0.5124	0.0422
Aritificial Neural Networks	0.0054	0.0003

Table 5: Forecast Evaluation

The lower the value the MAE and RMSE the better the model is at forecasting. Based on the results in Table 5 above, the Artificial Neural Networks is the best model of the three when it come to forecasting as it has the lowest MAE and RMSE.

5 Summary, Conclusion and Suggestion for Further Research

5.1 Summary

The study analyzed the volatility at the NSE using the NSE 20 Share Index from 3rd January 2012 to 22nd February 2021. The daily returns of the prices was calculated using $log\left(\frac{P_t}{P_{t-1}}\right)$. Various market efficiency tests such as the Augmented-Dickey Fuller test and Jarque-Bera test was then used to test for stationarity and normal distribution of the daily returns for the share index for the period under observation. The asymmetric GARCH models used in the study were the EGARCH and the GJR-GARCH model and the results were compared to that of the Artificial Neural Networks.

From the empirical results, the NSE 20 Share Index for the period under observation has fat tails and is skewed to the left therefore it is not normally distributed, the returns were also stationary and therefore we can conclude that the returns are mean-reverting in their first differencing form. Following the results from the parameter estimation of the EGARCH and GJR-GARCH models, the ARCH coefficients are positive and statistically significant which implied the volatility clustering and persistence of the market. For the EGARCH, the results showed that the market reacted differently to positive and negative shocks of the same magnitude whereas for the GJR-GARCH, the leverage effect was detected which implies that the higher the volatilty, the lower the returns. The two asymmetric GARCH models were compared using AIC and BIC to select the model with the best fit and the results showed that the EGARCH model was a better fit compared to the GJR-GARCH model as it had a lower AIC and BIC value.

For the Artificial Neural Networks, the model was trained using *tensorflow* in python. The study used the Recurrent Neural Network with the Long Short Term Memory Technique. The model was trained and predicted values were generated. The predicted values were then compared with the out of sample data and using RMSE, MAE and MSE, the forecasting ability of the Neural Networks was compared to that of the EGARCH and the GJR-GARCH model and based on the results of the RMSE and MAE the Neural Networks outperformed both the EGARCH and the GJR-GARCH models.

The main limitation of the study was the explainability of the neural networks because it contains many hidden layers of nodes that can be trained to identify and exploit extremely complex relationships in data. Since this process is not easily interpretable then it becomes difficult to explain how the predictions were made. In order for decision-makers to trust a model they need to understand and support how the model makes decisions and this can be challenging if there is no explanation on how these decisions were made.

5.2 Conclusion and Suggestion for Further Research

This main objective of the study was to forecast stock market volatility at the NSE and a number of conclusions can be drawn from the results of the data. First, the stylized facts of the financial time series data such as leptokurtosis, volatility clustering, volatility persistence and leverage effect was detected. It is necessary to account for stylized facts in order to get reliable future forecasts of volatility. Secondly, for the asymmetric GARCH models the EGARCH model was a better fit at parameter estimation compared to the GJR-GARCH. Lastly, when it comes to forecasting volatility, based on the results we conclude that the Artificial Neural Networks outperforms both the EGARCH and GJR-GARCH models.

Rojat, Thomas, et al. [26] stated that although a lot of work has been done on explainability in the computer vision and natural language processing fields, there is still a lot of work to be done to explain methods applied on time series. Further research can be done on Explainable Artificial Intelligence methods to try and solve the limitations of the explainability of the Neural Networks by using the Layer-wise Relevance Propagation technique. This will increase the trust in machine learning models and improve the overall troubleshooting. Since this study focused on the technical analysis when modelling and forecasting volatility as it focused on the movement of the stock market return given historical data, further research can also be done to incorporate fundamental analysis which can help the investor to determine the intrinsic value of the stock given the technical analysis.

References

[1] Vedat Akgiray. Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of business*, pages 55–80, 1989.

[2] Paul Alagidede and Theodore Panagiotidis. Modelling stock returns in africa's emerging equity markets. *International review of financial analysis*, 18(1-2):1–11, 2009.

[3] Dima Alberg, Haim Shalit, and Rami Yosef. Estimating stock market volatility using asymmetric garch models. *Applied Financial Economics*, 18(15):1201–1208, 2008.

[4] Josip Arnerić, Tea Poklepović, and Zdravka Aljinović. Garch based artificial neural networks in forecasting conditional variance of stock returns. *Croatian Operational Research Review*, pages 329–343, 2014.

[5] George S Atsalakis and Kimon P Valavanis. Surveying stock market forecasting techniques–part ii: Soft computing methods. *Expert systems with applications*, 36(3):5932–5941, 2009.

[6] Basel MA Awartani and Valentina Corradi. Predicting the volatility of the s&p-500 stock index via garch models: the role of asymmetries. *International Journal of Forecasting*, 21(1):167–183, 2005.

[7] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[8] Andrea Bucci. Realized volatility forecasting with neural networks. *Journal of Financial Econometrics*, 18(3):502–531, 2020.

[9] Matthew F Dixon, Igor Halperin, and Paul Bilokon. *Machine Learning in Finance*. Springer, 2020.

[10] R Glen Donaldson and Mark Kamstra. An artificial neural network-garch model for international stock return volatility. *Journal of Empirical Finance*, 4(1):17–46, 1997.

[11] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.

[12] Eugene F Fama. The behavior of stock-market prices. *The journal of Business*, 38(1):34–105, 1965.

[13] Suleyman Gokcan. Forecasting volatility of emerging stock markets: linear versus non-linear garch models. *Journal of forecasting*, 19(6):499–504, 2000.

[14] Ushir Harrilall. *Forecasting volatility in the South African stock market: A comparison of methods*. PhD thesis, 2014.

[15] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[16] Ng Jeffrey and Shah Subhas. *Hands-On Artificial Intelligence for Banking: A practical guide to building intelligent financial applications using machine learning techniques*. Packt Publishing, 2019.

[17] Betty Korir. Evaluation, modeling and forecasting volatility of daily and weekly returns on nairobi securities exchange using garch models. *IOSR Journal of Economics and Finance*, 9(6):10–21, 2018.

[18] Tse Yiu Kuen and Tung Siew Hoong. Forecasting volatility in the singapore stock market. *Asia Pacific Journal of Management*, 9(1):1–13, 1992.

[19] Leandro S Maciel and Rosangela Ballini. Neural networks applied to stock market forecasting: An empirical analysis. *Journal of the Brazilian Neural Network Society*, 8(1):3–22, 2010.

[20] Burton G Malkiel and Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

[21] Arfa Maqsood, Suboohi Safdar, Rafia Shafi, and Ntato Jeremiah Lelit. Modeling stock market volatility using garch models: A case study of nairobi securities exchange (nse). *Open Journal of Statistics*, 7(2):369–381, 2017.

[22] Soheil Almasi Monfared and David Enke. Volatility forecasting using a hybrid gjr-garch neural network model. *Procedia Computer Science*, 36:246–253, 2014.

[23] Tyagi Neelam. 6 major branches of artificial intelligence(ai).

[24] George Ogum, Francisca Beer, and Genevieve Nouyrigat. Emerging equity market volatility: An empirical investigation of markets in kenya and nigeria. *Journal of African Business*, 6(1-2):139–154, 2005.

[25] S Omar. Garch modeling in monthly foreign exchange and share prices for specific companies in kenya. *American Journal of Mathematics and Applications*, 1(1):5–12, 2013.

[26] Thomas Rojat, Raphaël Puget, David Filliat, Javier Del Ser, Rodolphe Gelin, and Natalia Díaz-Rodríguez. Explainable artificial intelligence (xai) on timeseries data: A survey. *arXiv preprint arXiv:2104.00950*, 2021.

[27] Martin Sewell. Characterization of financial time series. *Rn*, 11(01):01, 2011.

[28] Alan Raymond Thompson and MJD Ward. The johannesburg stock exchange as an efficient market. *Journal For Studies in Economic and Econometrics*, 19:33–63, 1995.

[29] Yiu Kuen Tse. Stock returns volatility in the tokyo stock exchange. *Japan and the World Economy*, 3(3):285–298, 1991.

[30] Dimitrios I Vortelinos. Forecasting realized volatility: Har against principal components combining, neural networks and garch. *Research in international business and finance*, 39:824–839, 2017.

[31] Lean Yu, Wei Huang, Kin Keung Lai, YOSHITERU NAKAMORI, and Shouyang Wang. Neural networks in finance and economics forecasting. *International Journal of Information Technology Decision Making (IJITDM)*, 06:113–140, 03 2007.

Curiginal

Document Information

Analyzed document	Stock-Market-Volatility-Forecasting-at-the-Nairobi-Securities-Exchange---A-comparison-between-Asymmetric-GARCH-Models-and-Neural-Networks-122164.pdf (D111584267)
Submitted	8/23/2021 8:40:00 AM
Submitted by	
Submitter email	Diana.Rugut@strathmore.edu
Similarity	11%
Analysis address	library.strath@analysis.urkund.com

Sources included in the report

W	URL: http://213.55.95.56/bitstream/handle/123456789/19465/Belete%20Dejen%20%202018.pdf?sequence=1&isAllowed=y Fetched: 8/10/2021 8:07:40 PM	1
W	URL: http://wiredspace.wits.ac.za/jspui/bitstream/10539/21053/1/ThesisReportZamaniMasinga%20325319.pdf Fetched: 1/4/2021 5:22:02 PM	2
W	URL: https://mpra.ub.uni-muenchen.de/95443/1/MPRA_paper_95443.pdf Fetched: 8/23/2021 8:43:00 AM	4
W	URL: http://erepository.uonbi.ac.ke/bitstream/handle/11295/58990/Modeling%20and%20forecasting%20stock%20market%20volatility%20at%20Nairobi%20Securities%20Exchange.pdf?sequence=3 Fetched: 7/14/2021 8:35:32 PM	6
W	URL: https://www.sciencedirect.com/science/article/pii/S1877050914013362 Fetched: 8/23/2021 8:43:00 AM	1
W	URL: https://core.ac.uk/download/pdf/160827876.pdf Fetched: 8/23/2021 8:43:00 AM	1
W	URL: https://core.ac.uk/download/pdf/132797589.pdf Fetched: 8/23/2021 8:43:00 AM	1
SA	**2193230_ECON5004P_1.docx** Document 2193230_ECON5004P_1.docx (D21472632)	2
W	URL: https://www.researchgate.net/publication/276728529_GARCH_based_artificial_neural_networks_in_forecasting_conditional_variance_of_stock_returns Fetched: 8/23/2021 8:43:00 AM	1
SA	**2241256.docx** Document 2241256.docx (D30182353)	2
SA	**MURA_EDITED_EFFI_1Thalassinos_muratoglu_ugurlu_PRAGUE.doc** Document MURA_EDITED_EFFI_1Thalassinos_muratoglu_ugurlu_PRAGUE.doc (D10639766)	1

2177417_ECON5004P.docx