International Journal of Engineering Innovation & Research
Volume 9, Issue 1, ISSN: 2277 – 5668

Development of a Prediction Model for Nigerian Stock Exchange using Linear Regression Algorithm

Dr. B.O. Omijeh [1] and Stanley Noah Tams [2*]
[1, 2] Center for Information and Telecommunication Engineering, CITE, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria.
*Corresponding author email id: kpaintoku@gmail.com

Date of publication (dd/mm/yyyy): 27/02/2020

***Abstract* – This research work presents the aftereffects of the advancement of a predictive model for the expectation of stock pattern dependent on a historical data gathered from the Nigerian Stock Exchange. First, is to collect a historical dataset which would be used to develop a model with machine learning algorithm after the historical dataset is trained with the algorithm. It is observed that a thorough exploratory research and preprocessing of the dataset and exploring Python libraries can yield to the development of good a model to predict the trend of stock with minimal error. To test for the precision of the model the historical dataset was split into training set and test set. The model was then evaluated with the root mean square error (RMSE) to see how large or minimal the error is on the prediction, this was seen to be closer to zero which shows that the prediction of stock trend with respect to the dataset is sufficiently okay. With the flexibility of Python a model was developed with the linear regression algorithm and the coefficients and intercept of the model were retrieved.**

***Keywords* – Predict, Dataset, Trend, Algorithm, Model, Linear Regression.**

I. INTRODUCTION

The stock market is one sector that huge amount of data is being generated on a daily basis making it a good source of data for analysis. Analyzing stock data can be said to be a comprehensive assessment and evaluation of a number of data resources of a particular trade instrument, a particular sector or the stock market as a whole to help predict the future pattern or trend of the instrument, sector or market.

It is the result of such an analysis that would empower financial investors or brokers settle on an educated choice on what to buy or sell in the market. These can also be used to gain insight in the economy of a nation, the stock market or a specific instrument.

There are two forms of stock analysis that investors tend to favor, these are:

Technical Analysis

Technical Analysis which examines historical price charts of an instrument and study earlier market pattern in order to predict the future trend.

Fundamental Analysis

Fundamental Analysis which examines data or information from the microeconomic background to assess probable trends of the market and profit from trades [1]. Fundamental analysis focuses on data sources that may be available for the public from sectors and Departments of Government such as regulatory Agencies and data sheets and revenue flows of a particular company.

This research would tend towards machine learning, developing a model that would be used to train the historical data generated by the Nigerian Stock Exchange (NSE) and its patterns in order to predict a probable trend of the market and then gain insight of the trend of some sectors of the economy at the time of review.

II. Related Works

According to (Shah et al 2019) [1], the financial markets are one of the most intriguing inventions of our time. A stock exchange is a regulated and organized financial market where long term securities (bonds, shares and notes) are sold and bought at costs administered by the powers of interest and supply, (Okaro, 2002) [2]. The exchange enforces strict rules, listing and statutory requirements that are binding to all trading and listed companies.

A number of researchers delved into developing model for stock data and analysts keep on making far reaching researches on how the most proficient method to predicting future pattern of stocks utilizing various algorithms. Different researchers had done are still doing detailed research and anticipating the dynamic nature of the adjustments in stock trends and talked about what impacted these changes. While Thawornwong and Enke (2004) in their exploration with the utilization of choice stock index anticipated the stock market utilizing Artificial Neural Network (ANN) [3], in a study by Kumar et al (2013) while contributing to knowledge introduced the combination of time series data with the Artificial Neural Network to predict the stock market [4]. (Mbeledogu et al, 2017) Proposed the use of soft computing techniques to predict the future trend of stock markets using an ANN model [5]. (Nivetha & Dhaya, 2017) Examined three different algorithms, the Support Vector machine [6], Multiple Linear Regression and the Artificial Neural Network and concluded that the ANN was the best and most effective of the three.

III. Design Analysis

3.1 *Problem Definition*

Predicting the stock market which is dynamic in nature is a tricky task with incessant lose for investors after delving into the stock market with little or no prior knowledge of the stocks or instruments to invest in; it is of the essence to investigate the market scientifically, compare stocks and instrument and make prediction for investors to go into the market (stock) with a high level of confidence of not losing their investment; in this case the NSE, developing a model for this purpose which would in turn give investors the needed confidence.

3.2 *Proposed Model*

Multiple linear regression algorithms is used to develop the model based on a dataset collected from the Nigeria Stock Exchange.

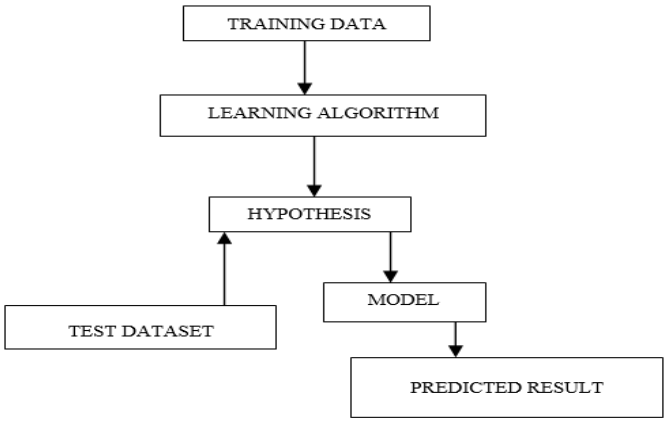

Fig. 1. Model Representation.

3.3 *Data Collection*

The method of data collection used to get the dataset for this research is secondary, since the data is a historical data collected and stored by the NSE.

The NSE operates a database that contains historical data of stock indices and other related data for research and other purposes. The lay down procedure had been followed to obtain these historical data for the purpose of this research work, hence; the NSE is the source of the dataset used for this research. The data collected span a total of ten (10) years of historical data of up to six (6) stocks from three indices, NSE Banking, NSE Oil and Gas, and NSE Consumer Goods.

3.4 *Choosing Algorithm*

A scatter diagram is used for the exploratory analysis which in viewing the dataset and gain insight on the connection between the variables and see the range of data flow to establish its linearity. In addition to the scatter diagram the pattern or movement of the historical stock data set was also drawn as shown in **Figure 2** and **Figure 3** below.

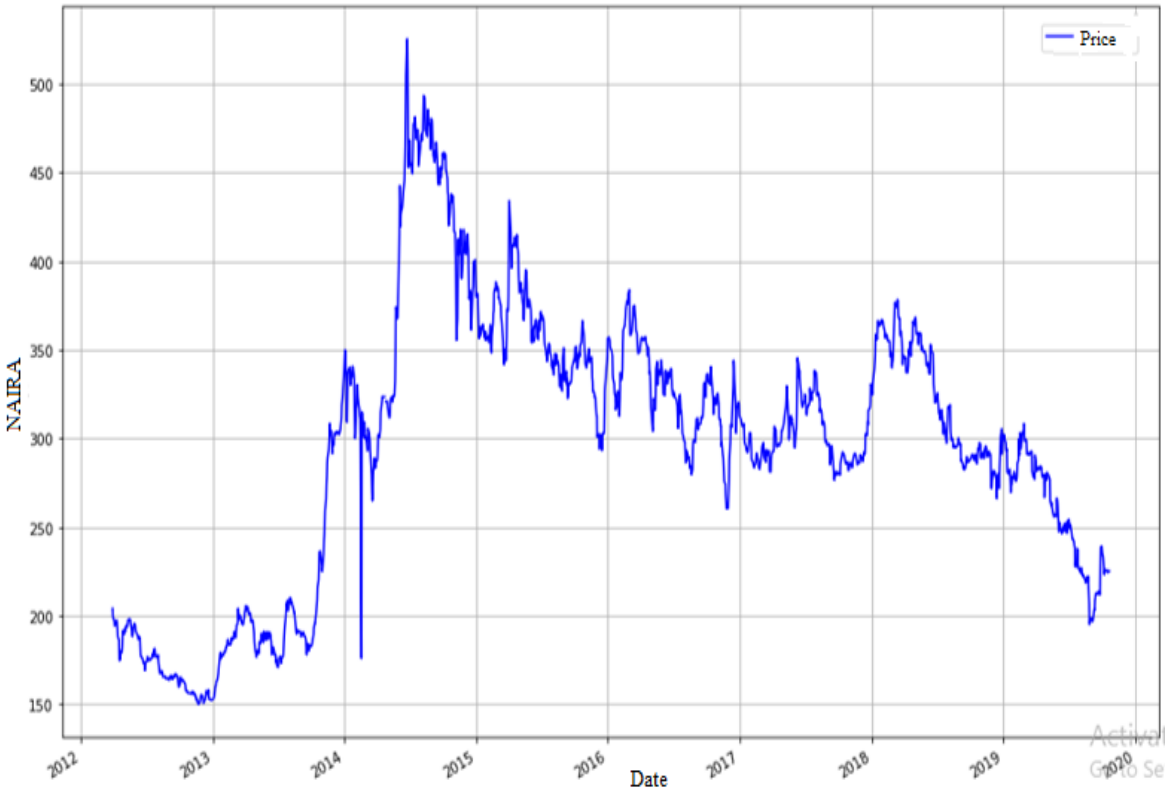

Fig. 2. Data Trend of collected Data.

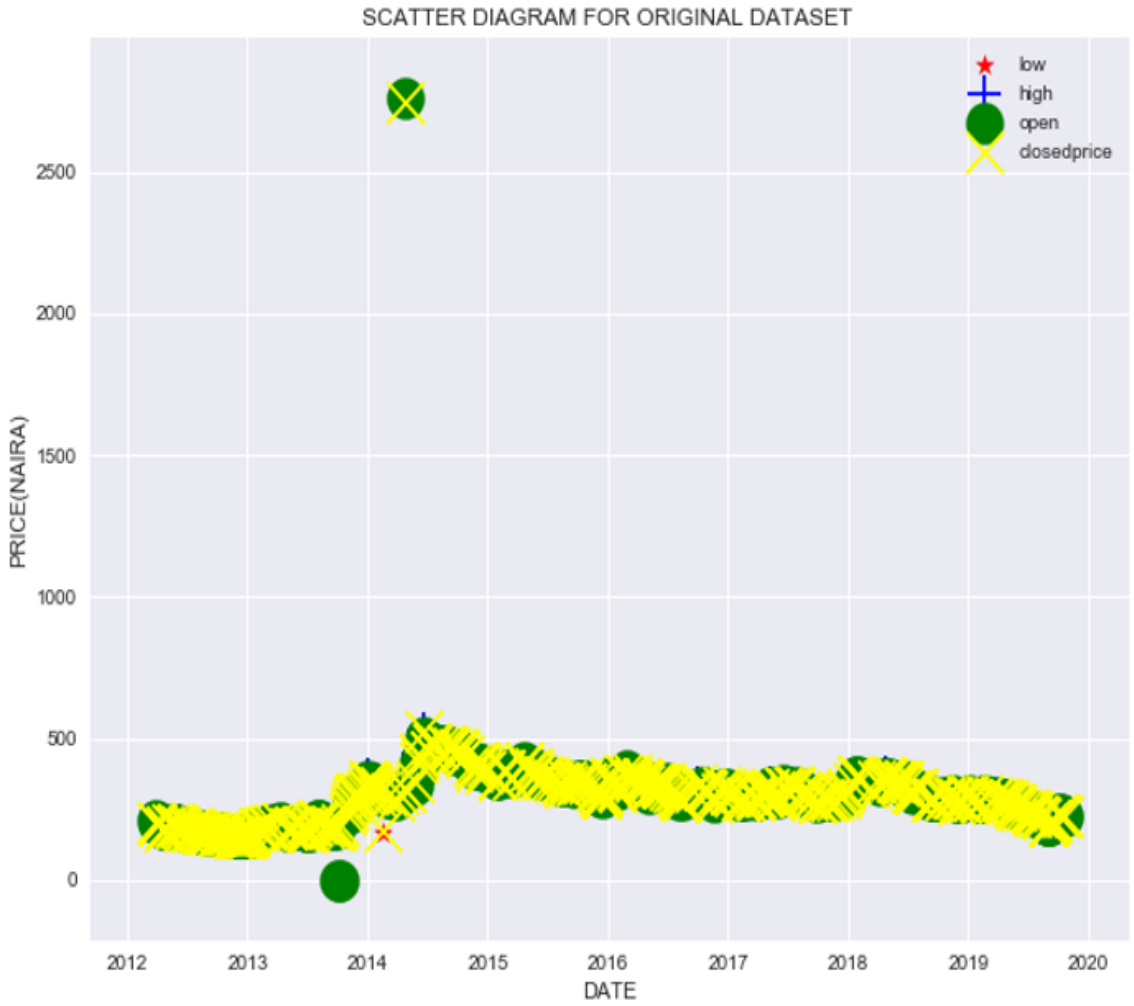

Fig. 3. Scatter Diagram for collected Data.

3.4 *Developing Hypothesis*

During exploratory study the dataset was seen to have the following three (3) features which are independent variables, namely **open, high and low** and there is the dependent variable being the feature to be predicted, **price,** hence there would be three coefficients, three unknown variables and one intercept or bias in the Hypothesis. Therefore, the hypothesis would look thus:

$$y = m_1x_1 + m_2x_2 + m_3x_3 + b \quad (1)$$

3.5 *Split Dataset*

The dataset collected from the NSE passed through some preprocessing activities to such as wrangling before splitting into a training set and a test set. Eighty percent (80%) of the data was used as a training set while twenty percent (20 %) was use as test set to evaluate the effectiveness of the model.

3.6 *Training the Algorithm*

Python libraries such as Numpy and Pandas were exploited in training the Algorithm with the train dataset with the aim to achieve a linear regression model. The algorithm fits the best possible straight linear to the set of data to achieve the model as it learns the pattern of the data. This is done by reducing the root mean squared error, **RMSE.**

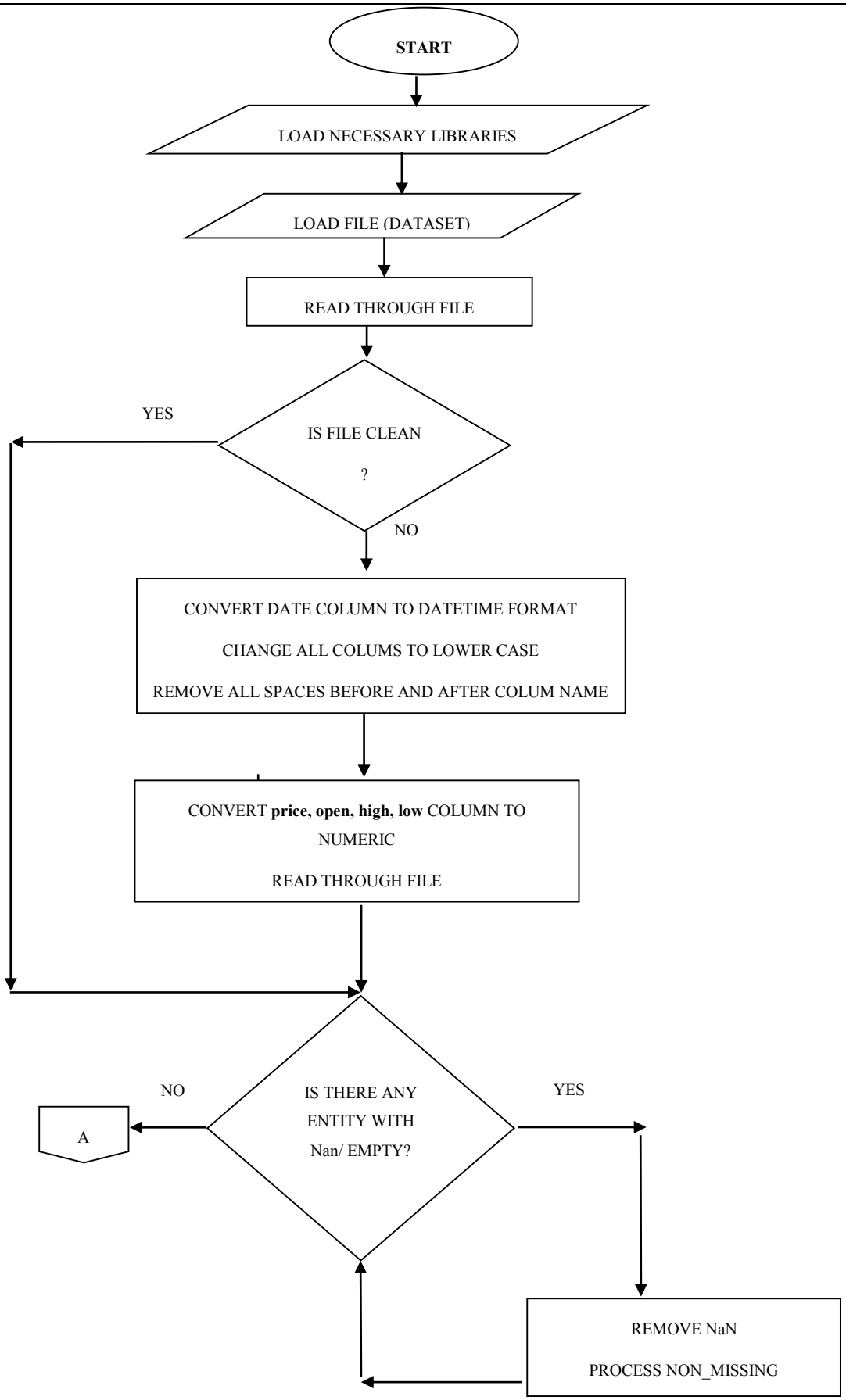

Fig. 3.3a Data Analysis Flowchart.

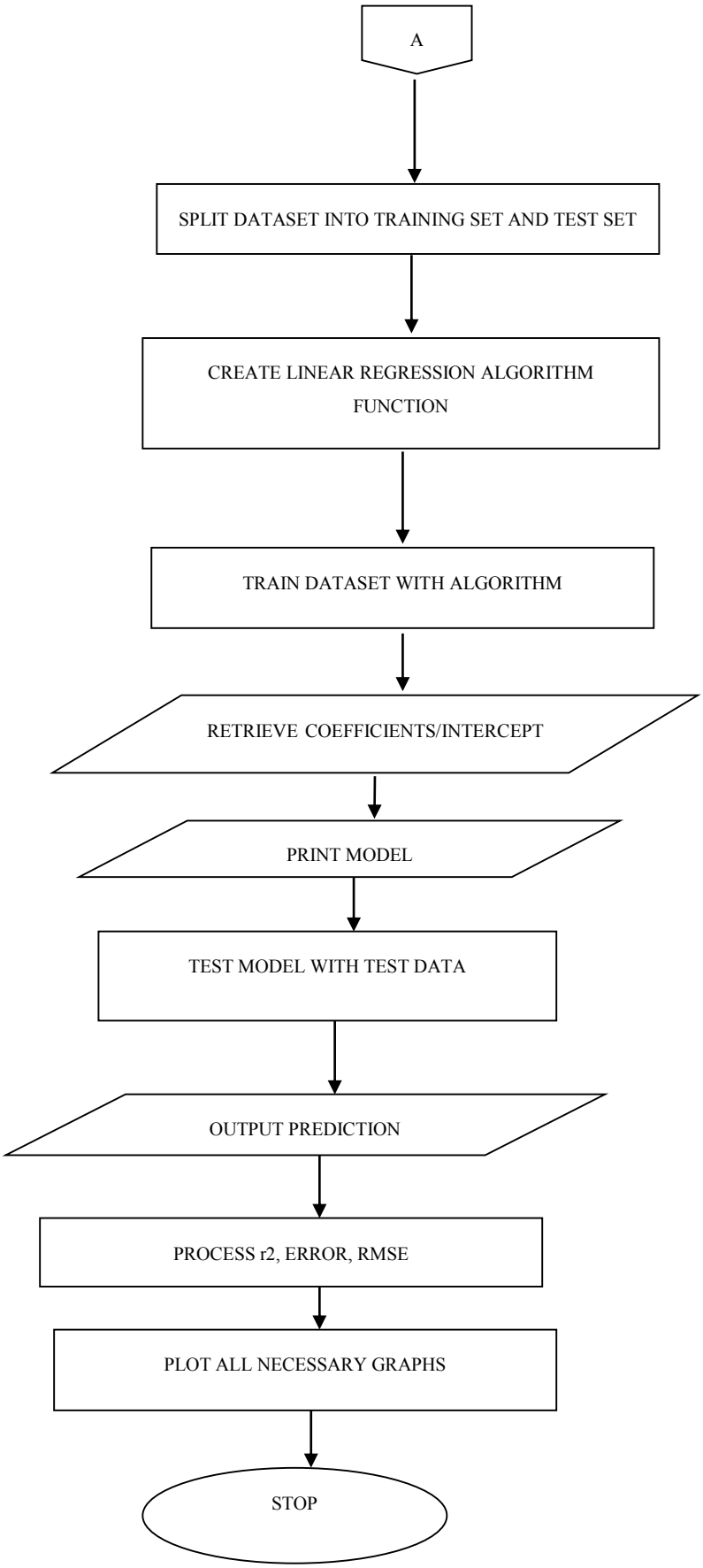

Fig. 3.3b Data Analysis Flowchart (continue).

IV. Results and Discussion

The dataset was then trained with the Algorithm which gave Linear Regression time to study the pattern of the train dataset after which the coefficients derived by the algorithm was retrieved along with the intercept as follows;
M_i = [-0.0257 0.5358 0.4893], b = 0.0354

4.1 *The Model*

Y Represent our target variable which is Price being what we are to predict.

M_i Represent the coefficients of the equation.

X_i Represent the features Open, High, Low.

b Represent the Intercept of the model.

Hence, the Model is represented as

$$Price = -0.0257 * Open + 0.5358 * High + 0.4893 * Low + 0.0354 \quad (2)$$

4.2 *Testing and Evaluation*

After retrieving coefficients and intercept from the Linear regression algorithm at the end of the learning process the model is then put to test. Using the test dataset X_test to implement the model against the "Predicted Price" shows the effectiveness of the model and it returned a minimal difference in prediction.

The Root Mean Squared Error (RMSE) was then used to evaluate the difference and found to be 4.1. below is the combined plot of the observed price movement, predicted price movement Error.

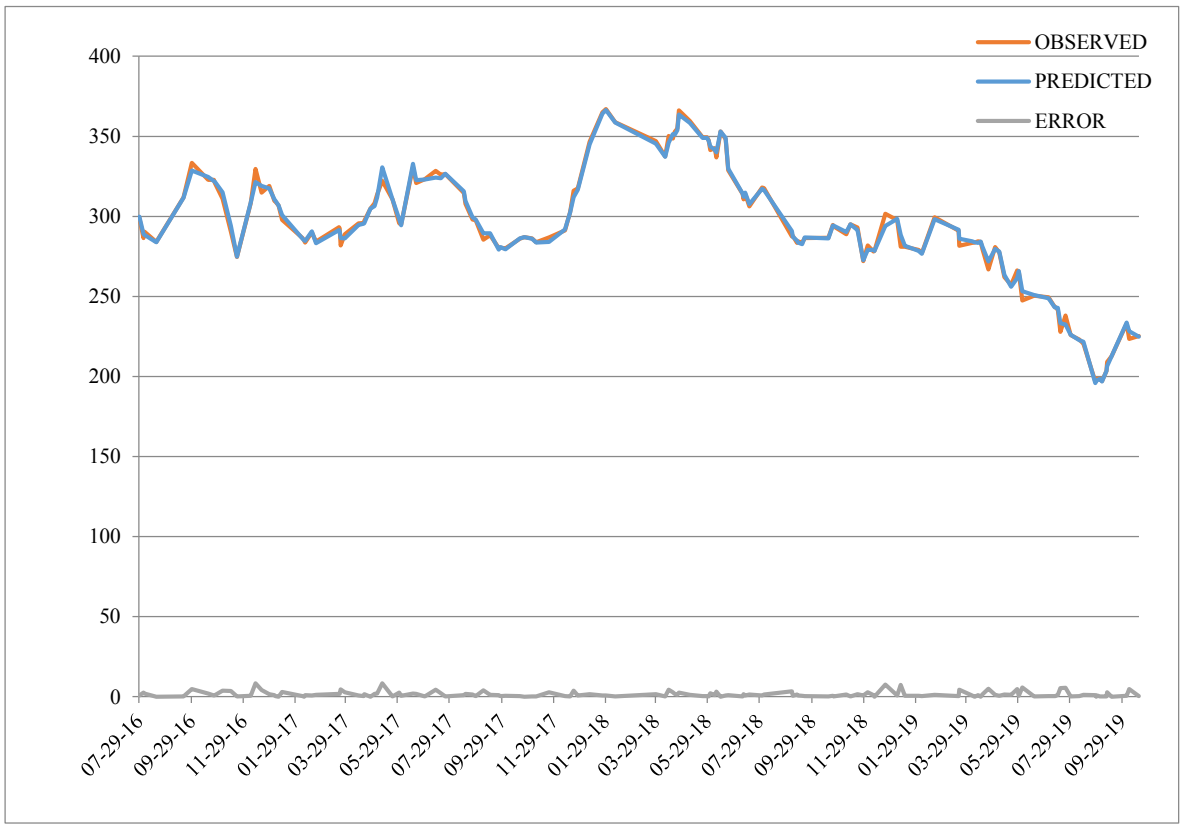

Fig. 4. Trend of Prediction, Observed data and Error.

V. CONCLUSION

The purpose of this research is develop a model to predict the movement of stock based on a historical stock data collected from the Nigerian Stock Exchange (NSE). Multivariate Linear Regression algorithm was used to train a section of the historical dataset after which coefficients and intercepts where retrieved to fit into the hypothesis to develop the model. After evaluation it is observed that the result is not far from the actual. This model is a first step in the development of a real time software for the prediction of the Nigerian Stock Market.

REFERENCES

[1] www.towardsdatascience.com (retrieved, August 6, 2019. Unpublished work)
[2] Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies*, *7*(2). https://doi.org/10.3390/ijfs7020026.
[3] Okaro, S.C. (2002). The Nigerian Stock Exchange and the Challenge of Economic Growth and Development in Nigeria.pdf. *The Nigerian Accountant*, Vol. 35, pp. 18–21.
[4] Kumar, D. Ashok, and S. Morugan. "Performance analysis of Indian stock market using neural network time series model." In pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on, pp. 72-78. IEEE, 2013.
[5] Mbeledogu, N.N., Chiemeke, S.C., Imianvan, A.A., & Ekuobase, G.O. (2017). *Soft Computing Stock Market Price Prediction for the Nigerian Stock Exchange* (August).
[6] Nivetha, R.Y., & Dhaya, C. (2017). Developing a Prediction Model for Stock Analysis. *Proceedings - 2017 International Conference on Technical Advancements in Computers and Communication, ICTACC 2017*, *2017-Octob*, 1–3. https://doi.org/10.1109/ICTACC.2017.11.

AUTHOR'S PROFILE

First Author
Dr. Burdilion Omijeh O is the Direcetor, Center for Information and Telecommunication Engineering (CITE), University of Port Harcourt, rivers State. Holds a B. Eng. Degree in Electrical/Electronic Engineering and M. Eng. Degree in Electronics/ Telecommunications Engineering from Ambrose Alli University (A.A.U), Ekpoma and University of Port Harcourt respectively and a PhD Electronics/ Telecommunications Engineering. His research areas include: Artificial Intelligence, Robotics, Embedded Systems Design, Modeling and Simulation of Dynamic systems, Intelligent Metering Systems, Automated Controls, Telecommunication and ICT. He is currently a Lecturer in the department of Electrical / Electronic Engineering, University of Port Harcourt, Rivers State, Nigeria. He is a corporate member of Nigerian Society of Engineers (NSE), Council for the Regulation of Engineering in Nigeria (COREN); and also, a registered Consultant to Companies and Institutions.
email id: omijehb@yahoo.comm

Second Author
Stanley Noah Tams is a Post graduate student at the Center for Information and Telecommunication Engineering (CITE), University of Port Harcourt with a BSc. in Computer Science at the Rivers State University of Science and Technology, Port Harcourt all in Rivers State. He is a Principal Program Analyst at the Power Equipment and Electrical Machinery Development Institute PEEMADI, Okene, Kogi State, Nigeria.