



Predicting the Trend of Dollar/Naira Exchange Rate Using Regression Model and Support Vector Machine

¹Abimbola, K. A., Abiola, O. A., and Akinola, S. O.

¹Kehinde.abimbola1@gmail.com

Department of Computer Science,
University of Ibadan

Abstract

The stock market is a network which provides a platform for nearly all major economic transactions in the world at a dynamic rate referred to as the stock worth that is predicated on market equilibrium. Predicting this stock value offers huge arbitrage profit opportunities that are enormous motivation for analysis during this space. Knowledge of a stock value beforehand by even a fraction of a second can result in high profits, machine learning technique has proven to be effective in the stock market prediction especially in the stock exchange traded fund and index exchange traded fund, but little effort has been extended to the prediction of the currency exchange traded fund using the linear regression and support vector machine, In this study the trend of the exchange rate of DOL/NGN was predicted from the exchange rate value generated using the linear regression and support vector Machine. The accuracy of the prediction from the Linear Regression model and support vector machine was analyzed using performance metrics.

Keywords: stock market, exchange rate, linear regression, support vector machine, Dollar, Naira

1. Introduction

Prediction of stock exchange returns is a very important issue and extremely complicated in monetary establishments. The prediction of stock value has invariably been a difficult task. It's been determined that the stock costs of any company don't essentially solely rely upon the monetary standing of the corporate however additionally depends on socio economic scenario of the country. It's no additional directly connected with the economic development of the country or explicit space. Therefore the stock worth prediction has become even tougher these days than before. Lately stock costs are affected thanks to several reasons like company connected news, political events natural disasters etc. Stock Market Prediction is an area that has driven the focus of many individuals including not only companies, but also traders, market

participants, data analysts, and even computer engineers working in the domain of Machine Learning (ML) and Artificial Intelligence (AI). It is one among the foremost necessary problems to be investigated in academic and monetary researches [1].

Regression analysis is an applied mathematical tool for finding the link between a dependent variable and more than one independent variable. The data supplied into the system is known as the training dataset and is gotten from historical data. This training dataset is employed to train the model and to estimate the unknown coefficients to be used in the regression equation. Now, the longer term stock values are foreseen using these coefficients [2].

Linear Regression can fit a best straight line that approximate the data set and make a forecast for the trend. In this the line is desired to pass through the maximum $y = a + bx$

Where y is the dependent variable and x is the independent variable a and b are the lines of

Abimbola, K. A., Abiola, O. A., and Akinola, S. O. (2022). Predicting the Trend of Dollar/Naira Exchange Rate Using Regression Model and Support Vector Machine, *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)*, Vol. 8 No. 2, pp. 51 - 59

coefficient and a least square method is used as the formula.

Figure 1 is a plot showing a linear regression while Figure 2 is a diagram showing the support vectors and hyperplane from a Support vector machine model.

Support Vector Machine solves both linear and non-linear problems. The fundamental idea behind the working principle of the

algorithm (SVM) is that it creates a line or hyperplane and later separates the data into classes [4].

The linear regression model and the support vector machine was used in the prediction of the trend of Dollar and Naira exchange rate. The accuracy of the prediction from linear regression model and support vector machine is analyzed using performance metrics. The objectives of this work are as follows:

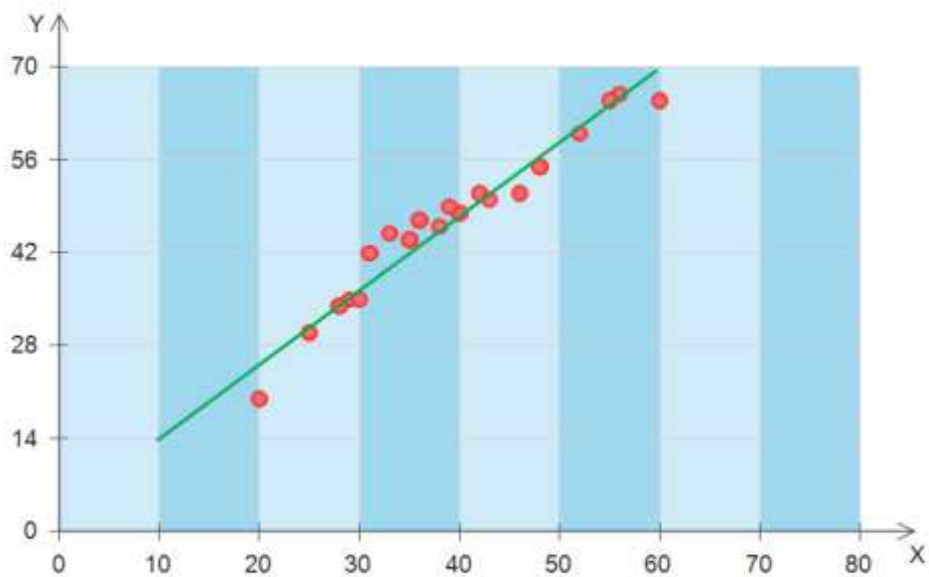


Figure 1: Plot showing linear regression model (Source: Edrawsoft [3])

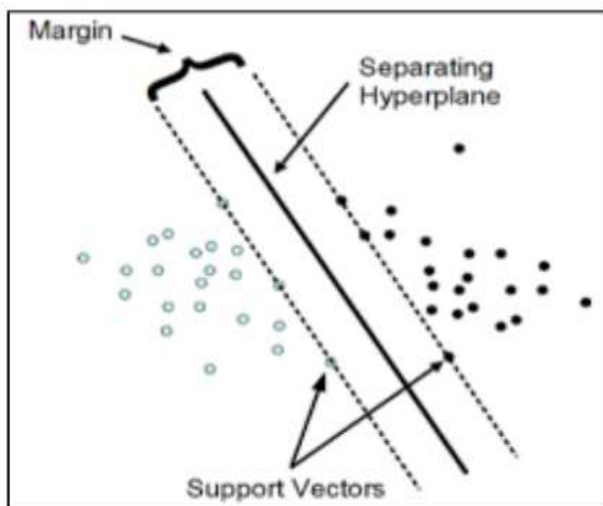


Figure 2: Graph showing hyperplane for SVM (Source: Vaisnavi, [5])

1. To develop a model to predict the trend of Dollar and Naira exchange using Linear Regression
2. To develop a model to predict the trend of Dollar and Naira exchange using Support Vector Machine (SVM)
3. To make comparison of the two aforementioned algorithms' performance on the same data
4. To employ performance measures on the predicted values to know which model gives more accurate result

This paper consists of five sections, the introductory section is just discussed, the next section highlights the literature review, the review of related works, while section 3 talks about the methodology used in furtherance of the task, In Section 4, results obtained from this study are presented and discussed. Conclusion is made in Section 5.

2. Review of Related Works

Payal Soni *et al.*, [1] explored the different techniques that are used in the prediction of share prices from traditional machine learning and deep learning methods to neural networks and graph-based approaches. It draws a detailed analysis of the techniques employed in predicting the stock prices as well as explores the challenges entailed which has been the common problem in stock market prediction.

Vaisnavi *et al.*, [5] collected a year knowledge from Coca Cola Company and perform prediction and additionally highlighted the actual fact that the stock data work well with the linear regression model and support vector regression in order that it will be a basis that future prediction will be created and also the result derived from each formula was evaluated with support vector machine acting higher than linear regression model.

Sahoo and Krishna [2] in 2015 predicted the future stock values using auto regression. If there is a linear relationship between input and output values, then it is called auto regression. The current value of output variable is predicted using the past values of the output variable and present value of input variables.

They have used Moore and Penrose technique to estimate the coefficients of regression equation and these coefficients are further used to predict the future stock values.

Sachin *et al.*, [6] applied the multiple correlation technique supported 3 variables to predict the stock exchange worth from stock exchange knowledge. During this approach monthly knowledge of past four and 0.5 years for Infosys Company is obtained from Yahoo Finance and used for building the prediction model and also the accuracy of the model is checked by exploitation same knowledge of past one and 0.5 years. This paper predicts the closing price of the stock exploitation open and close price as dependent variables. The least square method is employed to calculate the slope of the equation related to the corresponding variable quantity (X_i). The advantage of this methodology is that it reduces the full error. With the approach of multiple regression using 3 variables the accuracy of system is explicit to be eighty nine percent, thereby indicating higher accuracy as compared to regression toward the mean methodology used for prediction.

In Abdulsalam *et al.*, [7] Future stock exchange costs of 3 banks, First bank, Zenith, and Skye bank were foreseen by an information mining tool exploitation multivariate analysis, values of variables from the data were used to predict the longer term values of alternative variables through the utilization of time series data that utilized moving average methodology.

In 2009, Zhang and Zhang [8] analyze and forecast the stock exchange index with Markov properties, stock costs further as its state of interval visible of Markov model that provides investors with relevant reference model so as to avoid blind and irrational behavior. The most principle of victimization Markov process to predict is to make Markov model that predicts the state of an object during a bound amount of your time within the future by virtue of chance vector of the initial state and state transition probability matrix.

The review of the related work shows that machine learning technique has proven to be effective in the stock market prediction especially in the stock exchange traded fund

and index exchange traded fund, but little effort has been extended to the prediction of the currency exchange traded fund using the linear regression and support vector machine.

Hence, in this study, the trend of the exchange rate of DOL/NGN will be predicted from the exchange rate value generated using the linear regression and support vector Machine. The accuracy of the prediction from the linear regression model and support vector machine will be analyzed using performance metrics.

3. Methodology

This section talks about the methodology to be used in the work. Figure 3 is the block diagram depicting the system developed.

3.1 The Data used

The exchange rate of Dollar/Naira of 5 years, from 13/6/2016 to 26/4/2021, was downloaded from Oanda, Oanda is a company in in the United State of America that deals with foreign exchange, and Market analysis. The data is in CSV format, it consists of 255 closing price amounting to 255 weekly close price.

The first close price from Week 1 to Week 248 is selected as the training dataset which is approximately 4 and half years, while the close price from week 248 to week 255 is selected as the testing dataset which is the prediction of the next two months trend generated from the prediction of the price for the next months.

- Regressiondata = regressiondata[1:255,1:2]
- regressiondata_train = regressiondata[1:248,1:2]
- regressiondata_test = regressiondata[248:255,1:2]

3.2 Data preprocessing

Since the data is downloaded and not generated, there is need to check for missing values and other abnormalities that may exists in the CSV file downloaded, the data preprocessing operation carried out are the following:

3.2.1 Data quality Assessment:

Data Quality Assessment involves removing missing values in the dataset; this entails eliminating rows with missing data and estimating the missing values which are the methods for dealing with the missing values. Data quality assessment also involves eradicating inconsistent values and duplicate values.

3.2.2 Feature Sampling:

Feature sampling involves selecting a subset of the total data in order to avoid expensive memory and time constraints. A good sampling algorithm reduces the dataset to a subset where a good and maybe expensive machine algorithm or deep learning algorithm can be used. It is important to note that the subset gotten must have the same properties of the whole dataset in order to maintain the originality of the dataset.

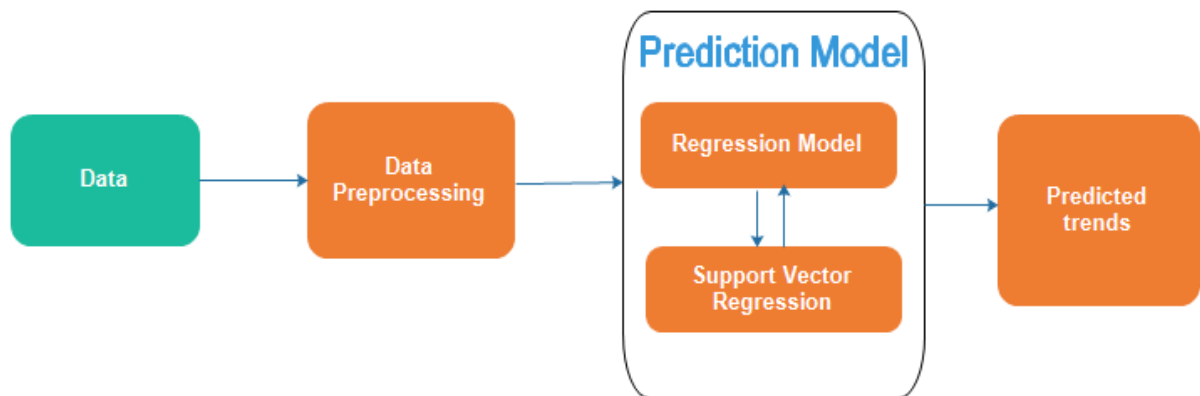


Figure 3 Block Diagram depicting the system

3.2.3 Cross validation:

Cross validation is a technique for evaluating machine model, in cross validation different model are tested on subsets of data, to see effectiveness of the model before making generalization.

3.3 The prediction model:

3.3.1 The linear regression model:

Linear regression is a statistical tool that can be used to show the relationship between a dependent variable (output) and the independent variable (input) and perform investigation on the relationship to make an inference or deduction from the trained dataset to be supplied to the system.

The operation is perform using R language with the R studio environment, the in-built functions in R studio make it easier to perform the major operation. The steps involve in performing the prediction using linear regression

Step 1: The data is in CSV format downloaded from Oanda company is imported to the R studio.

Step 2: The scatter plot of the weekly closing price of DOL/NGN for the period of 13/6/2016 to 26/4/2021 is plotted using the in-built R function `plot()`.

Step 3: Then the linear regression model is perform using the function `lm()` in R studio and the coefficient and the intercept is gotten using the function `linmod()`, getting the coefficient and the slope is enough to find our dependent variable Y, in the equation : $y = a + bx$

$$y = \text{intercept} + \text{slope}(x)$$

The intercept is what is needed for prediction and the coefficient is the slope

Step 4: The DOL/NGN exchange rate of the test data is predicted, and an in-built function `predict()` is use to avoid error and to safe time.

3.3.2 The Support vector machine

Support vector machines (SVMs) are a group of affiliated supervised learning methods used for classification and regression.

Given a set of training examples, each selected as belonging to one of two categories, an SVM training algorithm come up with a model to know which of the group the selected examples will fall into. A version of the support vector machine known as support vector regression is adopted in this work, the support vector regression is a version of the support vector machine that can be extended to regression, the support vector regression can work give better result than linear regression in a situation where the relationship between the said data is nonlinear, for example in the case of a polynomial function.

The steps involve in performing the prediction using support vector machine is as follows:

Step 1: The data in CSV format downloaded from Oanda company is imported to the R studio.

Step 2: The scatter plot of the weekly closing price of DOL/NGN for the period of 13/6/2016 to 26/4/2021 is plotted using the in-built R function `plot()`.

Step 3: The next step is to compute the SVM using the function `svm()`, the SVM method will compute the support vectors on the test data, and this process essentially train the algorithm. This process simultaneously plot the hyperplane, the **hyperplane** is the best line that separate the data and it is selected by computing the maximum margin from equidistant support vectors.

Step 4: Then prediction can be made from the new value classified from the trained algorithm using the `predict()` function.

4. Implementation

4.1 The linear regression summary

Figure 4 shows the scatter plot of the exchange rate for the 255 weeks using the R Language before a linear regression model is done while the Figure 5 shows the linear regression performed on it, the rate is on the y-axis while the weeks is on x-axis.

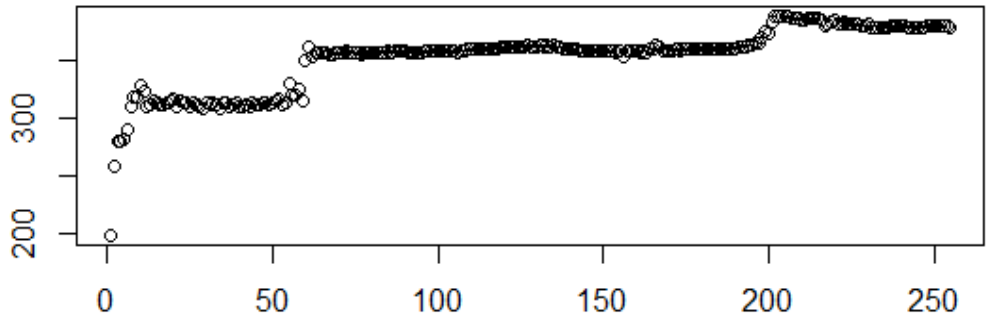


Figure 4: scatter plot of the exchange rate

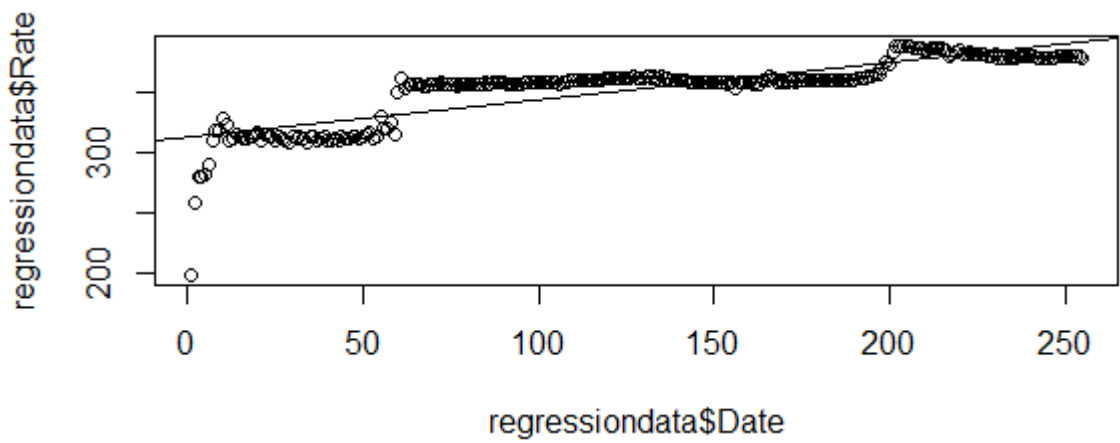


Figure 5: Linear regression fit on scatter plot

After the linear regression operation on the scatter plot, the predicted value of the testing data set is shown in the Table 1.

Table 1: The LR predicted value for the testing dataset

Week	Rate	LR Predicted value
248	379.8334	390.4839
249	380.5251	390.7983
250	379.5461	391.1128
251	380.0508	391.4272
252	380.5363	391.7416
253	380.1950	392.0561
254	380.4258	392.3705
255	379.3991	392.6849

4.2 The Support Vector Machine summary

SVM parameters:

SVM-Type: eps-regression

SVM-Kernel: radial

cost: 1

gamma: 1

epsilon: 0.1

Number of Support Vectors: 135

The SVM model is carried out on the dataset using the R-Studio IDE and the resultant graph showing the trend generated is shown in Figure 6.

After the SVM operation, the predicted value of the testing data set is shown in the Table 2

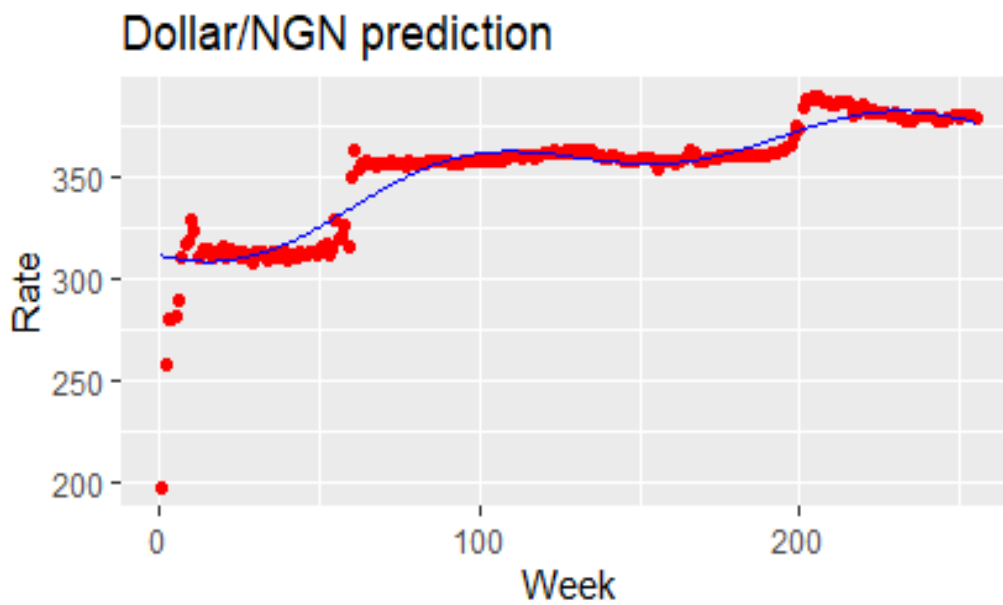


Figure 6: SVM model plot

Table 2: The SVM predicted value for the testing dataset

Week	Rate	Predicted value
248	379.8334	379.5294
249	380.5251	379.1863
250	379.5461	378.8278
251	380.0508	378.4545
252	380.5363	378.0670
253	380.1950	377.6660
254	380.4258	377.2523
255	379.3991	376.8264

4.3 Performance Analysis & Comparison of the Predicted Test_Data From LR and SVM

There is need to compare the accuracy of the two algorithm to know which one perform better, the following performance measure was used to evaluate the result from the LR and SVM operation to know which one perform better.

1. The root mean square error (RMSE)
2. The mean absolute error (MAE)
3. The mean absolute percentage error (MAPE)
4. The R-squared

4.4 Result and Discussion:

The work shows that the two machine learning algorithm, LR & SVM that has proven to work with stock exchange traded fund and index exchange traded fund prediction can also be extended to the currency exchange traded fund prediction taking the prediction of the trend of DOL/NGN exchange rate as a case study, with the SVM model performing better than the LR model and this is justified from the result of the performance metrics done to evaluate the predicted values derived from the LR & SVM model.

The lower the RMSE, MAE and MAPE value the better the model and the higher the R-squared value the better the model, Table 3 shows the RMSE, MAE, MAPE and R-squared value of the LR and SVM model.

Table 3: The comparison of performance metrics done on LR & SVM model

Performance Analysis Method	LR_Model	SVM Model
RMSE	11.55215	2.065548
MAE	11.52047	1.837737
MAPE	3.031327	0.48344
R-squared	0.005388226	0.007829579

5.0 Conclusion

In this work, prediction was done using the linear regression model and the support vector machine and it was done successfully.

The prediction was done on a dataset which is the exchange rate of Dollar/Naira of 5 years, from 13/6/2016 to 26/4/2021, the data was downloaded from Oanda, Oanda is a company in in the United State of America that deals with foreign exchange, and Market analysis.

From the study, it can be concluded that the two machine learning algorithms, Linear regression and Support Vector Machine that has proven to work with stock exchange traded fund and index exchange traded fund can also be extended to the prediction of currency exchange traded fund taking the prediction of the trend of DOL/NGN exchange rate as a case study, with the SVM model performing better than the LR model.

References:

- [1] Payal Soni et al (2022). Machine Learning Approaches in Stock Price Prediction: A Systematic Review” J. Phys.: Conf. Ser. 2161 012065.
- [2] Sahoo P. K and Krishna Charlapally (2015). Stock Price Prediction Using Regression Analysis", International Journal of Scientific & Engineering Research, Volume 6, Issue 3, March-2015 ISSN 2229-5518
- [3] EdrawSoft, 2014 version 7.9, released date Dec 15, 2014.
- [4] Rohith Gandhi (2018). Support Vector Machine- Introduction to Machine Learning Algorithms, <https://towardsdatascience.com/>, Accessed August, 2021
- [5] Vaishnavi Gururaj, Shriya V. R. and Ashwini K. (2019) “Stock Market Prediction using linear regression and support vector machine”, International Journal of Applied Engineering Research, Volume 14, pp. 1931-1934.
- [6] Sachin K, Shailesh J and Thakur R. S. (2013). Multiple Regression: A data mining approach for predicting the stock market trends based on open, close and high price of the month, International Journal of Computer Science Engineering and

Information Technology Research
(IJCSSEITR) ISSN 2249-6831 vol. 3, issue 4.

- [7] Abdulsalam S. Olaniyi, Adewole, Kayode S., Jimoh, R. G. (2011). Stock Trend Prediction Using Regression Analysis – A Data Mining Approach, ARPN Journal of Systems and Software, Volume 1 No. 4
- [8] Zhang D, Zhang X (2009), “Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method” in International Journal of Business and Management, Vol 4, No. 6,