

# Application of Support Vector Machine in Predicting Stock Market's Monthly Direction

<sup>1</sup>Lavi Linus Raymond <sup>2</sup>Joshua Etemi. Garba <sup>3</sup>Yusuf Musa  
Malgwi

*Modibbo Adama University of Technology, P.M.B 2076, Department of Computer Science, Adamawa State  
Yola, Nigeria*

*<sup>4</sup>Patience Agbu, Taraba State Polytechnic P.M.B 1030, Jalingo Taraba State, Nigeria*

Submitted: 10-12-2021

Revised: 20-12-2021

Accepted: 24-12-2021

**ABSTRACT**-In this paper, we explore the use of Support Vector Machines (SVM) to build prediction model for the Nigerian stock market index. Predicting the monthly direction of stock prices is a problem faced by many investors particularly during the period of Covid-19. Investors fear that the spread of coronavirus may destroy economic growth. The optimal solution was obtained from SVM to forecast the next month stock price. Confusion matrix was used to show the performance evaluation of the SVM algorithm. The output shows that our model has 86.67% accuracy for prediction using testing data with the aim of possibly achieving the highest accuracy in prediction. It was clear from the analysis

**KEYWORDS:** Machine Learning, Support Vector Machine; Dataset, Stock Market; Stock Direction Prediction.

## I. INTRODUCTION

There are two prices that are critical for any investor to know: the current price of the investment he or she owns or plans to own, and its future selling price. Despite this, investors are constantly reviewing past pricing history and using it to influence their future investment decisions. Some investors won't buy a stock or index that has risen too sharply because they assume that prices could further fluctuate, while other investors avoid a falling stock, because they fear that it will continue to deterioration.

[1].The basic tool used for predictions is spotting market trends. Technical analysis, a science of predicting future prices from historical price data has helped investors increase their accuracy regarding predictions. The algorithms used are able

to predict the right time to invest and to some extent, right quantity and right price to invest. Traders use these graphs to forecast favourable time to invest or drop stocks. The Efficient Market Hypothesis (EMH) also states that it is practically impossible to outperform the overall market. Proper and trained stock selection along with experienced and proper timing is the only way an investor can obtain higher returns if buying risky investments. However, whatever tool we use for prediction, technical or not, one supposition remains constant. We assume that trends and trajectories will repeat in the future, and that share prices will also fluctuate based on demand and supply.

[2].Stock market is a well regulated market established not only to serve as a meeting point platform for the highly liquid and insolvent (potential) investors, but to also support national economy growth and development. In line with the existing theories, the stock market thrives on information. As well structured the market is positioned, it is not insulated from or immune against arrival of information of different kinds capable of defining or redefining its values consistent with efficient market hypothesis. For instance, identify some of the major occurrences that stock market returns respond to which include disaster, sports, news, environmental issues, development in politics and outbreak of epidemics. Others include corporate governance issues, and accounting information

[3].Suffice to say therefore that the stock market thrives on heterogeneous information (endogenic or exogenic in nature) that can influence the market participants' and its stakeholders' decision about the market activities. Apparently, the outbreak of Coronavirus disease 2019 popularly

referred to as (COVID-19) has had its toll on global stock markets among other spheres of life. But, the link between COVID-19 confirmed cases (within an economy, a region and globally) and the Nigerian stock market capitalisation calls for empirical investigation and confirmation. However, apart from immediate effect recorded around the early stage of the pandemic, its short or long run impact on stock market COVID-19 Pandemic.

[4]. In case the course of the pandemic becomes less serious than anticipated, there is likelihood that the stock market will rebound as the confirmed cases continue to rise. It is indisputable that the idiosyncratic nature of the unprecedented novel Coronavirus outbreak and its implication on global economy as well as stock markets is distinct compare to previous experienced economic crises

[5]. The AIDS pandemic is the largest public health crisis in contemporary history, and the consequences of inaction range from the highly personal to problems of national security and economic development.

The sudden emergence of COVID-19 in December 2019 from Wuhan, China which was declared pandemic and Public Health Emergency of International Concern (PHEIC) in March 11, 2020 by the World Health Organisation has made history. The pandemic is characterised with quick spread, but less deadly than other viruses. The COVID-19 related first death was recorded in January 11, 2020 and still counting. Based on WHO statistics, about 5,934,936 and 100,610 cases have been confirmed globally and in Africa respectively with 367,166 and 2,554 deaths in that order as at May 31, 2020. Realities have defied past predictions regarding likely total number of the COVID-19 cases and fatality, suggesting unpredictable rise in the future days. Probable impact of such PHEIC on global and/or regional stock markets is inevitable. But, the severity might not be similar to other outbreaks. In the past decade, 2009 Swine flu, 2014 polio, 2014 Ebola outbreak, Zika virus epidemic and the current COVID-19 among other epidemic and pandemic have been declared PHEIC. In the earlier decades, HIV/AIDS was declared a pandemic.

One of the consequences of globalization is that a nation (state) becomes too big to solve the small problems and too small to solve the big ones. In this age of strong global integration, arguably a declared PHEIC would have a consequential effect on stock market performance. This is owing to the fact that a declared pandemic or epidemic would always come with accompanied trade restriction orders, capable of inflicting direct and indirect effects on the capital market.

[6]. COVID-19 has affected the global market negatively, although the impact may differ from one economy, firm or sector to the other. As some companies are experiencing a boom, others are faced with retardation in operations, thereby affecting their performances.

## II. LITERATURE SURVEY

How can we predict whether the price of a particular stock will go up or down in the upcoming month? In the modern techniques, one way is to develop a predictor based on the information in the historical data. First of all we should select some major factors that may influence the performance of the stocks; we can further discover an interesting model from our dataset to predict the future performance of any stocks. To predicting the stock market status an algorithm was developed which is the Support Vector Machine (SVM) is the trending technology to predict the stock market direction. In this scenario we implement the algorithm which predicts the stock market using SVM which give an output very efficiently.

Support vector machine (SVM) is a relatively new learning algorithm that has the desirable characteristics of the control of the decision function, the use of the kernel method, and the sparsity of the solution. They present a theoretical and empirical framework to apply the Support vector machines strategy to predict the stock market.

### Importance of stock market

[7]. The stock market is one of the most important ways for companies to raise money, along with debt markets which are generally more imposing but do not trade publicly. This allows businesses to be publicly traded, and raise additional financial capital for expansion by selling shares of ownership of the company in a public market. The liquidity that an exchange affords the investors enables their holders to quickly and easily sell securities. This is an attractive feature of investing in stocks, compared to other less liquid investments such as property and other immovable assets.

Some companies actively increase liquidity by trading in their own shares. History has shown that the price of stocks and other assets is an important part of the dynamics of economic activity, and can influence or be an indicator of social mood. An economy where the stock market is on the rise is considered to be an up-and-coming economy. In fact, the stock market is often considered the primary indicator of a country's economic strength and development.

### **Covid-19 Effects on Nigerian economy**

As the world looks to eliminate COVID-19 pandemic with no result in sight. The world economic outlook is fragile, especially developing countries like Nigeria are struggling to find their fit; the global GDP growth is estimated to be 2.5 percent in 2020. Developing countries have recorded relatively lower, Nigeria currently has a 2.3 GDP growth rate before the pandemic, Nigerian had been struggling with weak recovery from the 2014 oil price shock. They are making IMF revise the 2020 GDP growth rate from 2.5 percent to 2 percent, as a result of relatively low oil prices and some low fiscal infrastructures. With that, Nigeria debt keeps increasing, the recent estimate puts the debt service-to-revenue ratio at 60 percent there is a source of concern for the policymaker, and its revenue may continue to decline with fall in the oil prices. The above factors will increase the economic effects of the COVID-19 outbreak and make it more difficult for Nigeria to get out of the recession. In Nigeria, the government has been making efforts to increase aggregate demand by increasing government spending and tax cuts for businesses. Nigeria budget increased from 8.83 trillion naira (\$24.53 billion) in 2019 to 10.59 trillion naira (\$29.42 billion) in 2020, which is 11 percent of the GDP, small enterprise receives tax relieve on company income tax, and the tax rate for small-medium enterprise comes down from 30 to 20 percent. The COVID-19 crisis affects all components of aggregate demand.

The fall in household consumption in Nigeria will stem from federal government lockdown order, thus causing consumers to spend primarily on essential goods and services, low expectations of future income, particularly workers of gig economy sector which are engaged on a short-term, and the working poor of the informal sector of the economy; and the loss of wealth and expected wealth with a decline in assets such as stocks and home equity. The state governors

have again imposed a lockdown of four weeks in all states of Nigeria including Abuja. Nigeria is more of a gig economy and a large informal sector that contributes 65 percent of economic output in Nigeria. Lockdown has reduced the consumption commodities in general and income-generating capacity of the gig and informal sector of the economy, thus reducing consumption expenditure generally. Firm's investments will reduce largely due to the uncertainties that come with lack of knowledge about the duration pandemic outbreak, the effectiveness of fiscal policy measures, and the reaction of economic agents to fiscal policy measures, and negative investor

believes causes turbulence in capital markets globally. Of note, the crisis has led to a massive decline in stock prices, as the Nigerian Stock Exchange records its worst performance since the 2008 financial crisis, which has eroded the wealth of investors.

[8]. Taking into consideration the uncertainty that is associated with the pandemic and the negative profit on possible investment; firms are likely to hold long-term investment decisions for the future. On the other hand, government purchases may increase as governments can afford to run budget deficits, using fiscal policy measures to reduce the fall in consumer spending, and in times like these, excess crude oil account is to care of the budget deficit. But, the limited fund in that account has made the government reduce the 2020 budget, implementing reduction 2020 budget will negatively affect Nigeria projected economic growth. But, governments dependent on oil sales and other commodities, the fall in the global demand for oil and other commodities due to pandemic will significantly increase their fiscal deficits. The price of oil was just over \$20 a barrel as on April 26, whereas Nigeria's budget has a beach make the price of \$57 per barrel been a mono-economy with 2.18 trillion naira (\$6.05 billion) deficit. Again, the decline in the demand for oil and oil prices associated with the pandemic will adversely affect the volume and value of Nigeria's net exports, thus, the Nigerian government will cut planned expenditure. Of note is the minister of finance announcement of a 1.5 trillion naira (\$4.17 billion) cut in non-important capital spending on the 18th March.

### **Covid-19 spillover to the Nigerian economy**

[9]. There are five main ways through which the Covid-19 pandemic spilled over into Nigeria. One, the Covid-19 pandemic affected borrowers' capacity to service loans, which gave rise to NPLs that depressed banks' earnings and eventually impaired bank soundness and stability. Subsequently, banks were reluctant to lend as more and more borrowers struggled to repay the loans granted to them before the Covid-19 outbreak. Two, there were oil demand shocks which was reflected in the sharp decline in oil price. The most visible and immediate spillover was the drop in the price of crude oil, which dropped from nearly US\$60 per barrel to as low as US\$30 per barrel in March.

During the pandemic, people were no longer travelling and this led to a sustained fall in the demand for aviation fuel and automobile fuel which affected Nigeria's net oil revenue, and eventually affected Nigeria's foreign reserve. Three,

there were supply shocks in the global supply chain as many importers shut down their factories and closed their borders particularly China. Nigeria was severely affected because Nigeria is an import-dependent country and as a result Nigeria witnessed shortage of crucial supplies like pharmaceutical supplies, spare parts, and finished goods from China. Four, the national budget was also affected. The budget was initially planned with an oil price of US\$57 per barrel. The fall in oil price to US\$30 per barrel meant that the budget became obsolete and a new budget had to be formed that was repriced with the low oil price.

Finally, the Covid-19 pandemic affected the Nigerian stock market. Major market indices in

the stock market plunged when investors pulled out their investments into so-called safe havens like US Treasury bonds. Stock market investors lost over NGN2.3 trillion (US\$5.9bn) barely three weeks after the first case of coronavirus was confirmed and announced in Nigeria on January 28, 2020. The market capitalisation of listed equities, which was valued at NGN13.657 trillion (US\$35.2bn) on Friday, February 28, 2020 depreciated by NGN2.349 trillion to NGN11.308 trillion (US\$29.1bn) on Monday 23 March 2020. The All-share index closed at 21,700.98 from 26,216.46 representing 4,515.48 points or 20.8 per cent dropped.

### III. METHODOLOGY

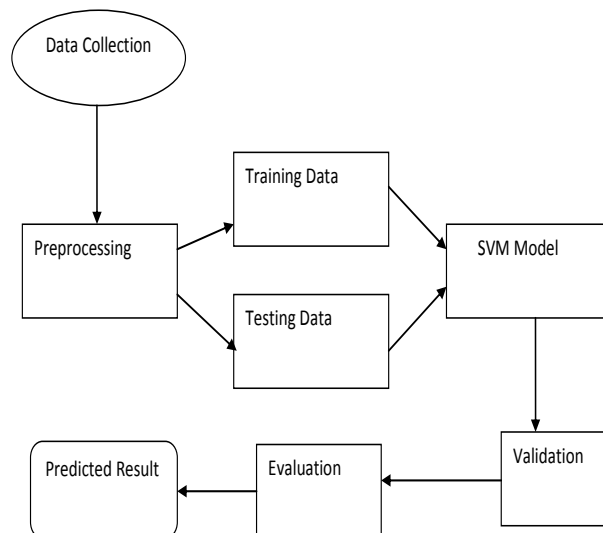


Figure 3.1 model development processes for stock prediction

#### Data collection

Data collection is a very basic module and the initial step towards the project. It generally deals with the collection of the right dataset. The dataset that is to be used in the market prediction has to be filtered based on various aspects. Data collection also complements to enhance the dataset by adding more data that are external.

In order to undertake the experiments and evaluate the results from the experiments, Nigerian stock market was selected as our research domain.

#### Data preprocessing

[10].Data pre-processing is a part of data mining, which involves transforming raw data into a more coherent format. Raw data is usually inconsistent or incomplete and usually contains many errors. The data pre-processing involves checking out for missing values, looking for

categorical values, splitting the dataset into training and test set and finally do a feature scaling to limit the range of variables so that they can be compared on common environs.

The whole dataset is divided into a training set (70% of the data) and a testing set (30% of the data). The training sample is used for learning of the network WHILE the test sample is used for measuring the accuracy of the model. The advantage of the transformation is that the distribution of the transformed data will be more symmetrical and close to a normal distribution.

#### Training Data

[11].Training the machine is similar to feeding the data to the algorithm to touch up the test data. The training sets are used to tune and fit the models. The training set is the largest and is used to train the SVM model to learn data patterns or used

for constructing the desired model. Training data is nothing but enriched or labeled data used to train the machine learning algorithms in order to increase prediction accuracy. You might just need to collect more of it to sharpen your model accuracy. But the chances of using your data are pretty low because, as you build a great model you need great training data at scale

### Testing Data

Next, a smaller testing set is used to test the generalization ability of the newly trained network. It is actually the input given to a software program. It represents data that affects or is affected by the execution of the specific module. Some data may be

used for positive testing, typically to verify that a given set of input to a given function produces an expected result. During the test phase of machine learning, this data is used to estimate how well your model has been trained and to estimate model properties.

### SVM model classifier

SVM classifier is a type of discriminative classifier. The SVM uses supervised learning that is, a labeled training data. The outputs are hyperplanes which categorizes the new dataset. They are supervised learning models that uses associated learning algorithm for classification and as well as regression.

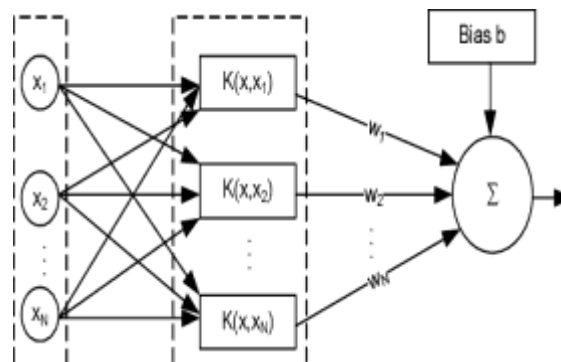


Figure 3.2 Support Vector Machines (SVM) model.

### SVM Parameters

The tuning parameters of SVM classifier are kernel parameter, gamma parameter and regularization parameter.

**Kernels** can be categorized as linear and polynomial kernels calculate the prediction line. In linear kernels prediction for a new input is calculated by the dot product between the input and the support vector.

$$K(X_i, X_j) = \begin{cases} X_i \cdot X_j & \text{Linear} \\ (\gamma X_i \cdot X_j + C)^d & \text{Polynomial} \\ \text{Exp}(-\gamma |X_i - X_j|^2) & \text{RBF} \\ \tanh(-\gamma |X_i - X_j + C|) & \text{Sigmoid} \end{cases}$$

Where  $K(X_i, X_j) = \phi(X_i) \cdot \phi(X_j)$  that is, the kernel function, represents a dot product of input data points mapped into the higher dimensional feature space by transformation  $\phi$

**C parameter** is known as the regularization parameter; it determines whether the accuracy of model increases or decreases. The default value of  $c=10$ . Lower regularization value leads to misclassification.

**Gamma parameter** measures the influence of a single training on the model. Low values signify far

from the plausible margin and high values signify closeness from the plausible margin.

### Validation

[12]. This is defined as the model's quality response as the training process is completed. The model should be validated with an earlier prepared set of the input and output data for that purpose (15% of data) in order to compare the model response with the measured data. The model response based on the validation data set has to be evaluated graphically and by applying numerical quality measures which are going to be appraised according to each used numerical model quality criteria. Finally, the validation set is used to double check the performance of the network and to measure the performance of the trained model. To properly divide the data, some points must be followed regarding each set. The training set should be the largest to give the network enough data to learn, and the validation set should be chosen as to balance between having enough of a sample size to evaluate the network and enough remaining data for the other two sets.

[13]. Recommend a 10% to 30% portion for testing data, and also recommend 90% to 70% for training data.

### Evaluation

The last group of the methodology steps refers to the model evaluation process, which is defined as the model's response quality on the dataset that is not used in the training or validation process. The model should be evaluated with an earlier prepared set of the input and output data for that purpose (15% of data) in order to compare the model response with the measured data. The process of the model evaluation is similar to the process of the validation; the only difference is the number of numerical quality measures that are used in that process.

### Prediction

New data is fed to the classifiers to predict the next stock market change.

The result of the SVM is a projection that the stock market will move up or down, rise or fall, (increase or decrease) the following month compared to the previous stock prices. The predicted result concludes if the machine learning used has led to profits or losses.

## IV. CONCLUSION

In this paper, we proposed the use of data collected from different global financial markets with machine learning algorithms to predict the stock index movements.

The main objective of the work is to develop a more adaptive and effective market prediction system by applying machine learning techniques. However, in this work I have analyzed various machine learning algorithms for stock market prediction and I found that the most suitable algorithm for predicting the market price of a stock based on various data points from the historical data is the Support Vector Machine algorithm. It is very encouraging as a data classifier because of its simple structure and less feature space and can classify numerical data as well non numerical data such as text, images, patterns etc. SVM is basically a two class classifier.

## REFERENCES

- [1]. Lekhani, R., 2017. Stock Prediction using Support Vector Regression and Neural Networks. *International Journal of Advance Research, Ideas and Innovations in Technology*, 3(6).
- [2]. Al-Awadhi, A. M., Al-Saifi, K., & Al Hammadi, S., 2020. Death and contagious infectious diseases: Impact of the COVID-19 virus on stock market returns. *Journal of Behavioral and Experimental Finance*, 100326.
- [3]. Burdekin, Richard C. K., & Tao, T. 2021. The golden hedge: From global financial crisis to global pandemic. *Economic Modelling* 95:170–80.
- [4]. Alfaro, L., Chari, A., Greenland, A., & Schott, P. 2020. Aggregate and firm-level stock returns during pandemics, in real time, SSRN Discussion Paper Series.
- [5]. Davis, G. F., & Anderson, P. J., 2008. Social movements and failed institutionalization: Corporate (non) response to the AIDS epidemic. *The Sage handbook of organizational institutionalism*, 371-388.
- [6]. Tashanova, D., Sekerbay, A., Chen, D., Luo, Y., Zhao, S., & Zhang, T., 2020. Investment opportunities and strategies in an Era of Coronavirus pandemic. Available at SSRN 3567445.
- [7]. Sachin Sampat Patil1, Kailash Patidar, Megha J., 2016. "A Survey on Stock Market Prediction Using SVM". *International Journal of Current Trends in Engineering & Technology*, 1(2), Issue: 0, 1-7.
- [8]. Adetunji, D. A., 2020. Coronavirus Facts and Control: What Multiplier Effects for Nigeria Economy.
- [9]. Ozili, P. K., & Arun, T., 2020. Spillover of COVID-19: impact on the Global Economy. Available at SSRN 3562570. Pandemic: Lessons from the 'Spanish Flu' for the Coronavirus's Potential Effects on Mortality and Economic Activity, NBER Working Paper 26866.
- [10]. Tay, F. E., & Cao, L., 2001. Application of support vector machines in financial time series forecasting. *omega*, 29(4), 309-317.
- [11]. Sadia, K. H., Sharma, A., Paul, A., Padhi, S., & Sanyal, S., 2019. Stock Market Prediction Using Machine Learning Algorithms. *IJEAT*.
- [12]. Demuth, H., Beale, M., & Hagan, M., 2004. *Neural Network Toolbox User's Guide*, the Mathworks Inc. South Natick. MA.
- [13]. Kaastra, I., & Boyd, M., 1996. Designing a neural network for forecasting financial. *Neurocomputing*, 10, 215-236.