Proceedings of the 5th International Conference on Smart Systems and Inventive Technology (ICSSIT 2023)
IEEE Xplore Part Number: CFP23P17-ART; ISBN: 978-1-6654-7467-2

Stock Market Price Prediction Using Machine Learning

Dr. Shariq Mohammed,
Department of Accounting, College of Commerce and Business Administration, Dhofar University
drshariqmohammed@gmail.com

Dr Somanchi Hari Krishna
Department of Business Management Vignana Bharathi Institute of Technology
harikrishnasomanchi@gmail.com

Dr. Pralhad K. Mudalkar
Bharati Vidyapeeth (Deemed to be University) Institute of Management, Kolhapur
pralhad.mudalkar@bharatividyapeeth.edu

Narinder Verma
Department of Computer Science and IT, University of Jammu, Jammu and Kashmir
narinder733.nv@gmail.com

P.Karthikeyan
Department of Management Studies, Kongu Engineering College, Perundurai,Erode-638060 Tamilnadu.
ptp_karthi@yahoo.co.in

Dr. Ajay Singh Yadav
Department of Mathematics SRM Institute of Science and Technology Delhi-NCR Campus Ghaziabad U.P
ashokpersonal939@gmail.com

2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) | 978-1-6654-7467-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSSIT55814.2023.10061120

***Abstract*—**

The stock market is known for its high volatility, fast changes, and nonlinear behaviour; investors should be prepared for all three. It is highly challenging to precisely forecast the movements of stock prices due to the presence of multiple factors, both macro and micro, such as politics, the state of the global economy, unanticipated occurrences, and the financial success of a firm, among other factors. However, because there is such a wealth of information available, it can be challenging to draw conclusions. Because of this, researchers, analysts, and data scientists working in the financial sector are constantly looking for new analytical techniques that may be used to spot trends in the stock market. This development gave rise to the practice of algorithmic trading, which is characterized by the application of trading strategies that are pre-programmed and automated. Machine Learning models such as LSTM can accurately forecasts the prices of stocks as actual and predicted.

Keywords*— *Stock Market Price Prediction, Machine Learning, Analysis,

I. Introduction

This research mainly focuses on stock market price prediction, before proceeding further, author discusses the fundamentals of stock market price prediction, market growth and its techniques. In this section author proceeds further discussion. Fundamental analysis (FA) takes into account both the economic and financial characteristics of a security in order to establish its true value. The true value of an investment is dependent on a number of factors, including the state of the economy and the market, as well as the financial health of the corporation that issued the investment.

Fundamental analysts consider both macroeconomic and microeconomic factors when determining the value of a security. Examples of macroeconomic factors include the state of the economy and the status of businesses, while examples of microeconomic factors include the quality of management at a company.

The end result will be a number that may be utilized to decide whether or not the current price of a security is appropriate, as well as whether or not the price should be altered either higher or downwards.

Fundamental analysis is a useful tool that can be utilized when attempting to ascertain the real worth of a stock, which is also referred to as its "fair market value" in some instances.

Fundamental analysts are particularly interested in stocks that are now trading at a premium or discount relative to their own intrinsic worth.

When the present price of a stock is lower than its fair market value, investors are encouraged to buy shares of the company in order to capitalize on the lower price.

When the market price of a stock is higher than its fair market value, investors may be urged not to buy the stock or to liquidate whatever holdings they may already have. Alternatively, they may be encouraged to invest in other securities.

On the other hand, technical analysts base their forecasts on the movement of stock prices in the past in order to create predictions about the immediate future.

Knowledge of Fundamental Analysis:

Fundamental analysis is often carried out in a straight way, moving from a macro perspective to a micro one, with the goal of determining which stocks have wrong market valuations and locating those securities. The majority of the analysts read in order to:

The state of the economy as a whole, as well as the degree to which a particular industry has become well-established

How well or poorly the stock performed relative to other stocks issued by the same company.

They are able to calculate the genuine worth of the stock on the market by doing so.

While undergoing fundamental analysis, several financial measures, such as a company's revenues, earnings, future growth, return on equity, profit margins, and so on, are considered in order to arrive at an estimate of the company's true worth and growth potential.

Quantitative and Qualitative Fundamental Analysis:

The word "fundamentals" can be used to refer to almost everything that has to do with the soundness of a company's finances, which makes it challenging to pin down a particular definition for it. In addition to financial measurements such as sales and profit, these factors may also take into consideration other aspects of a firm, such as its market share or the quality of its leadership.

Every one of the basic components can be dissected into quantitative and qualitative subcategories, respectively. The financial importance of these terms is largely equivalent to what is commonly understood to be their meanings:

Quantitative: information that can be shown using numbers, figures, ratios, or formulas

Qualitative: rather than a quantity of something, it is its quality, standard, or nature

Technical Analysis:

Predictions on the price of a stock can be established using straightforward technical analysis, more specifically, the Simple MA and Exponential MA approaches. In addition, authors will develop a prediction model by making use of LSTM, which is a deep learning framework for time series, and then we will evaluate it alongside our technical evaluation.

In spite of the fact that the disclaimer may lead readers to believe differently, the purpose of this article is not to instruct readers on how to earn money trading stocks. In this study, I will be using terminology associated with trading and investments; nonetheless, it is important that you understand that I am not providing individual financial advice. Expressions such as:

The trend of stock prices during the past half-decade is referred to as medium-term movements. Trend indicators, on the other hand, are statistics that reflect the direction in which stock prices are moving.

A flowchart for the processing of stock prediction technique data is shown in figure 1.

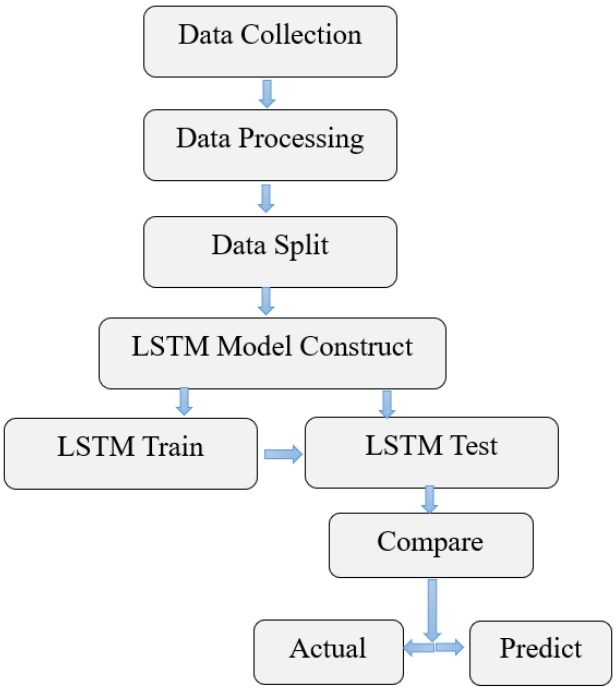

Figure.1: Flow chart of Stock Prediction Processing

II. LITERATURE REVIEW

German Capdehourat et al. proposed a machine-propelling This section discusses the article related to stock market growth and prices in each company sector,
The opinions of those around us have, and always will be, a significant factor in the majority of the decisions that we make. The Internet and the World Wide Web have made a number of things possible, one of which is the distribution of information about the opinions and experiences of people who are not well-known critics and whom we do not know directly. That is to say, we are now able to collect knowledge on persons about whom we had no prior information. On the flip side, more and more people are becoming more open with complete strangers while interacting online. What drives researchers in this field is people's interest in and response to reviews of products and services that may be found on the internet. During this process, there will be many challenges that need to be conquered before the intended results can be obtained. During our survey, we explored the common processes that are utilized in this setting as well as the steps that ought to be put into place in order to address the concerns that have arisen [1].

This research was carried out by V Kranthi Sai Reddy[1-2], who is affiliated with the Sreenidhi Institute of Science and Technology in Hyderabad, India. He studied Electronics and Communication Management for his undergraduate degree. The buying and selling of stocks is an extremely important part of the functioning of the international economy. A good example of a stock market prediction would be to speculate on the potential value of a stock or other security that would be traded on the stock market in the near future. In this research, we evaluate whether or not it is possible to make accurate predictions about the stock market by utilising machine learning. The majority of stock brokers rely on either fundamental or technical research, as well as time series analysis, in order to forecast the movements that will occur in stock values in the future. Python is used for the majority of the work involved in making accurate stock market forecasts using machine learning. In this paper, we describe a Machine Learning (ML) method for acquiring insight into historical data relating to the stock market and applying that understanding to future forecasts based on that previous data. This research incorporates stock price projections for the large-cap, medium-cap, and small-cap markets respectively. It does this by utilising data from both daily and minute-by-minute price changes, together with the machine learning strategy known as Support Vector Machine (SVM).

Forecasting the Stock Market Index:

This was the focus of the investigation that was conducted by Lufuno Ronald Marwala[3-5]. In fulfilment of the degree requirements for the Master of Science in Engineering, a thesis that was handed in to the Faculty of Engineering and the Built Environment at the University of the Witwatersrand in Johannesburg. The Efficient Market Hypothesis (EMH) asserts, at its most fundamental level, that it is impossible to effectively anticipate future asset values using historical asset prices. This is one of its central tenets. Because of this, the market acts in a way that is very similar to a random walk, which makes it impossible to make precise predictions. The intricacy of the international monetary system also makes it difficult to formulate economic forecasts that can be relied upon. This study was conducted with the intention of constructing a model with the help of AI techniques and predicting the movement of a stock market index (AI). In order to generate forecasts for

future index values, several methods of data analysis such as neural networks (NN), support vector machines (SVM), and neuro-fuzzy systems are used to the available historical information. The ability to predict financial time series is made possible by technologies powered by artificial intelligence that take into account the intricate nature of the financial industry. Both linear modelling techniques, such as autoregressive moving average (ARMA), and non-parametric modelling techniques, such as random walk, are used in the comparison of several alternative AI methods (RW). The RW method is seen as having a less formal approach. The information that was gathered for the study was obtained from the Johannesburg Stock Exchange. In order to collect this information, historical All-Share Index closing prices were analysed and taken into consideration. Each of the three approaches was able to exhibit an accurate capacity to forecast the price of the index. The linear model performed exceptionally well compared to all of the different AI approaches that were tested. Nevertheless, the strategy of taking a random walk produced the most encouraging results. If we put these strategies into practise, we will be able to demonstrate that it is indeed possible to forecast future prices. Because buying and selling on the market costs money, it is unfortunately difficult to demonstrate that any of these three ways can disprove the ineffective form of market efficiency. This is because of the ineffective form of market efficiency. The data indicate that the ranking of performance is dependent on the accuracy metric that was selected. [Citation needed] There is a high level of performance across the board with multilayer perceptron neural networks, neuro-fuzzy systems, and support vector machines.

Indian stock market prediction using artificial neural networks on tick data:

Dharmaraja Selvamuthu, Vineet Kumar, and Abhishek Mishra are the writers of this work. It was published by the Mathematics Department of the Indian Institute of Technology Delhi in Hauz Khas, New Delhi 110016, India[6-7]. On the stock market, a corporation's stocks and derivatives can be purchased or sold at a price that has been established in advance. The amount of money that is exchanged in stock transactions is what powers it. It is reasonable to assert that the stock market is one of the industries that is expanding at the quickest rate worldwide. These days, a significant number of people are involved in the industry in some capacity. As a result, it is essential to have a current awareness of the state of the market. As a result of this, stock price forecasting has developed into a well-liked hobby in tandem with the growth of the stock market. Because of the fluid nature of the stock market and the rapidity with which stock values can shift, it can be difficult to forecast stock prices. m It has already been demonstrated that Stock Learning event representations that are capable of capturing syntactic and semantic information over the entirety of a text corpus are a feasible research direction. One of the many ways in which these strategies have demonstrated their value is through the prediction of script events. However, the events that are gleaned from raw texts lack the common-sense information, such as the intentions and feelings of the individuals, that can be utilised to discern between events that just have surface variances. This research suggests that in order to find a solution to the problem at hand, one may make use of outside common sense knowledge about the event's goals as well as the responses it drew from the participants.

Our model obtains significantly better event embeddings for the tasks, with improvements of 78% on the hard similarity task, better predictions of what will happen next in given contexts, and better predictions of stock market volatility, as shown by experiments on three event-related tasks: event similarity, script event prediction, and stock market prediction. The market is frequently characterised as being predictable, noisy, nonparametric, and chaotic (Ahangar et al. 2010). As a result of advancements in technology, market participants are moving away from using fundamental analysis to predict stock prices and instead relying on Intelligent Trading Systems. They will be able to make quick decisions on their finances by using this knowledge. The basic goal of a trader is to forecast movements in stock prices so that they can profit from those movements by selling low and purchasing high. According to the efficient market hypothesis, the behaviour of stock prices is consistent with that of a random walk, making it impossible to forecast their movements. It would appear that using today's technology to forecast the price of a stock with the same degree of precision as an experienced trader would be challenging. However, with the availability of vast amounts of data and the progression of technology, it may be possible to construct an intelligent prediction algorithm, the results of which can assist traders or investment organisations in earning more money. As a result, the advantages of utilising an algorithm are directly related to the effectiveness of the method.

.

III. PROPOSED LEARNING TECHNIQUES

Let's use the current situation of the globe as an illustration of a pattern that is not typical, in which a large number of companies and other industries are suffering severe setbacks. A good example of an extraordinary event that nobody could have anticipated is the epidemic that spread all over the world. When anything like this happens as a result of a natural disaster or some other event, it will have an effect not only on trading prices but also on stock price charts, businesses, and companies.

Data Set:

In this article, authors considered the price of one share of Apple Inc. stock (represented by the symbol AAPL) every year for the past 21 years (1999-11-01 to 2021-07-09). Alpha Vantage provides its users with free access to both historical and real-time stock market data through its API, which will be used to load data for the research. A free API key can be obtained by enrolling on Alpha Vantage, which will allow access to the company's information. This article provides a comprehensive description of the steps required to obtain one. Once you have access to the API, all you need to know about a stock is its ticker symbol in order to analyze it.

The first 80% of the data will be used to train the model, and 20% will be used to validate it. The data set was taken from Kaggle.

3.1 Classification of Machine Learning Techniques

LSTM:

Long-Term Short-Term Memory (LSTM), often known as the technique, is an extraordinarily efficient approach to time series. It is able to recognize trends over a period of time and make exceptionally accurate projections of future values. The Cell State (Ct), which reflects both the short-term and long-term memories that are stored internally to the cell, is the most essential component of an LSTM model.

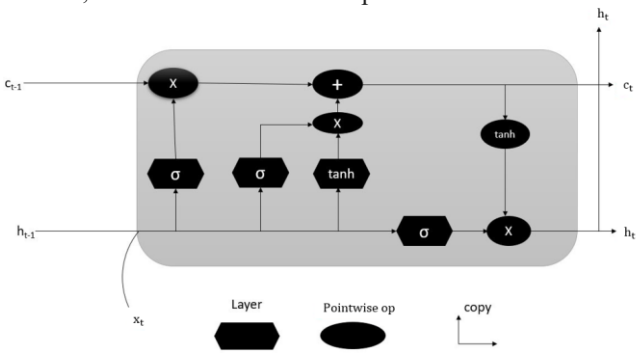

Figure 3.1 LSTM Structure

RNN - "Recurrent Neural Networks," are a specialized form of neural network that may take into consideration the results that have occurred in the past. The phrase "recurrent neural networks" is what this abbreviation refers to, and RNN is the acronym for that phrase. In a forthcoming piece, additional information on the intricate workings of RNNs as well as their level of complexity will be presented. LSTM design will continue to be the primary focus of this area's research in the future, with an emphasis on explaining and clarifying LSTM design through the medium of instruction. Because of the potential for gradient explosion and data loss, it can be difficult for RNNs to transmit long-term data. These problems can be solved using the recurrent neural network (RNN) architecture, which is also referred to as the Long short-term memory (LSTM) model. It is utilised for the deep learning applications that present the greatest amount of difficulty. When utilising deep learning to make predictions regarding stock values, it has been demonstrated that LSTMs are extremely helpful.

An LSTM architecture is a helpful tool for handling complex problems and enhancing the accuracy of predictions. This is because the memory cells of an LSTM architecture function so well. The ability to know when it is appropriate to forget something and when it is appropriate to remember it is a skill that is formed by both the long-term and short-term memory systems. Demonstrating the operation and structure of an LSTM may primarily be done using one of three different methods. The cellular level is where researchers start their investigation because here is where the most fundamental information is kept. The second phase is known as the "concealed state," and it has three gates: the "forget gate," the "input gate," and the "output gate." At some time in the future, we are going to talk about these fences, locks, and gates. Last but not least, a looping stage has been included so that the data components that need to be computed after each time step can be integrated.

When all of the target states have been modified, it is an indication that the one-time process has been finished. Let's investigate the idea of gates, and then find out what each type of gate does.

The forget gate is considered to be one of the most important gates in LSTMs. The sigmoid function is applied to the data in order to establish which values should be kept and which ones should be thrown away. In most cases, we will keep the numbers that are nearer to 1, and we will get rid of the ones that are nearer to 0.

The function of the input gate is to supply the LSTM cells with newly acquired information. Both the tanh and sigmoid functions have been incorporated into the system. The tanh functions are responsible for compressing the range from 1 to 1, making certain that neither extreme of the range is excessively large.

This component, known as the Output Gate, is responsible for determining which of the subsequent set of concealed states will be brought into visibility. This encapsulates the LSTM architecture in its entirety as well.

The mathematical theory that underpins LSTMs can be analysed from a number of different perspectives. These two pictures show two different ways that LSTMs approach the computation of mathematical expressions. The second illustration shows a potential outcome in which the bias function is left out of the equation. In any event, we won't get into specifics on the first picture that was shown. You may learn more about the photograph by following the link that is provided in the image caption and clicking on the source of the image. The initial LSTM image is capable of being analyzed with rather accurate mathematical equations, which may then be used in computing.

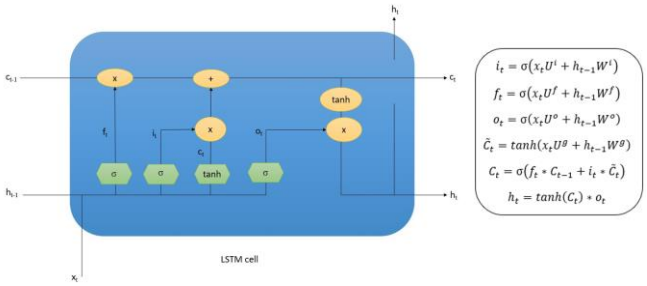

Figure 3.2 LSTM – Mathematical equation

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$
$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$$
$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$$
$$\tilde{c}_t = \tanh_c(W_c x_t + U_c h_{t-1} + b_c)$$
$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t$$
$$h_t = o_t \circ \sigma_h(c_t)$$

Where:

X(t): input vector

H(t): hidden state vector, also known as output vector of the LSTM unit

č(t): cell input activation vector

C(t): cell state vector

F(t): forget gate's activation vector

I(t): input/update gate's activation vector

O(t): output gate's activation vector
WW, UU,and BB: weight matrices and bias vector parameters which need to be learned during training
In the analytical patterns, the values stay the same:

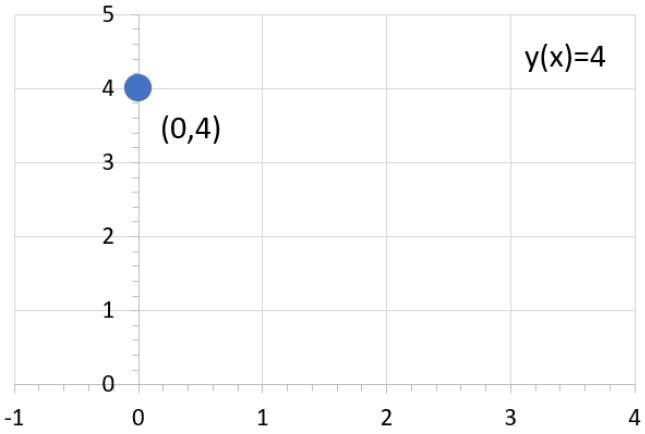

Figure 3.3 Stock pattern

3.2 *Results and Discussion:*

Time series forecasting is of no use when the values being predicted do not change. No matter what the external factors are, the value you obtain will remain consistent. Time series analysis is thus unnecessary in these cases. The preceding diagram is an illustration of a constant graph, in which the values always stay four regardless of the other constraints. Assume, for the sake of argument, that the price of a certain chocolate has been stable for quite some time.

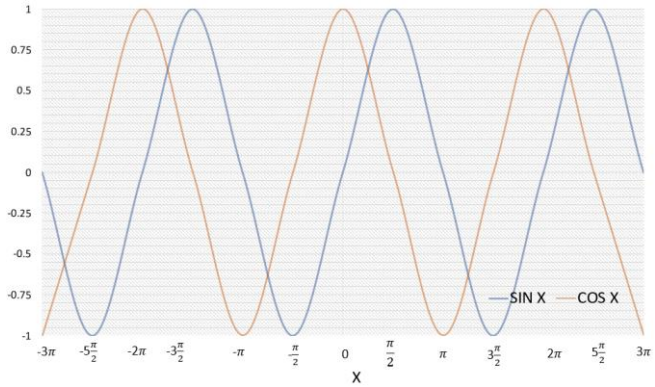

Figure 3.3. The analytic pattern's data were represented as functions.

Stochastic modelling loses value in the same way that time series forecasting does when a constant function exists that provides the required values at all desired times. By presenting the time series analysis as a function, mathematics makes it possible to use a calculator to calculate the recurring pattern. Graphs of two types of such waves—a sine wave and a cosine wave—follow. When Y(x) = X, the majority of functions are easy to evaluate. Artificial intelligence is unnecessary here.

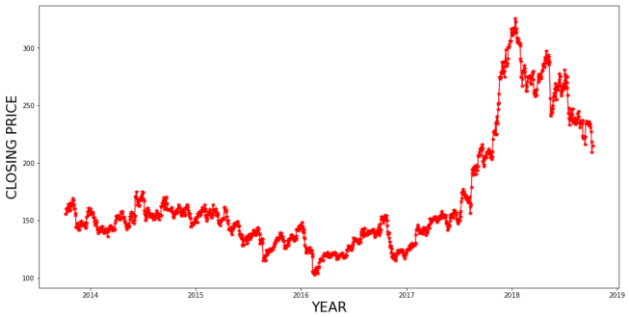

Figure 3.4. Stock price plot from 2014 to 2019

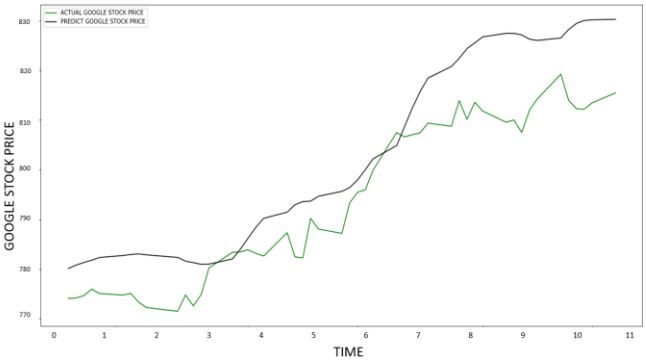

Figure 3.5. Google stock value prediction

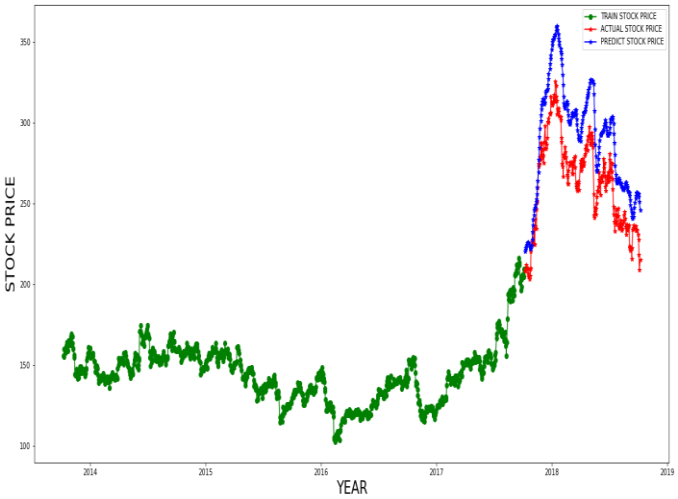

Figure 3.6. Comparison of actual and predicted stock price

Form figure 3.6 it indicates that, once completing the above mentioned steps with these processes using LSTM and will be able to use matplotlib to display the difference between the predicted stock price and the actual stock price.

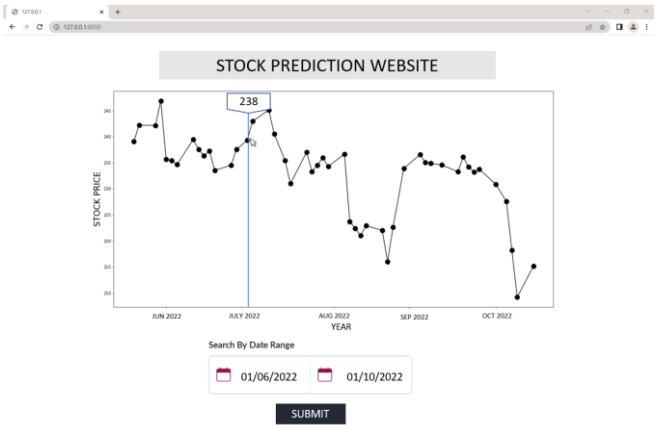

Figure 3.7. Working of developed dashboard

IV. CONCLUSION

Machine learning is quite different from random selection, which only has a 90 percent chance of producing accurate results. In spite of having less access to test data, LSTM has a better performance. This is still the case Bi-LSTM shall be implemented it may have a chance of increasing its accuracy in future. The results of classifying features utilizing only graphical figures, while the exact price points from predicted price weren't always close to the actual price, this proposed LSTM model did still indicate overall trends such as going up or down. This proposed LSTM indicates that, it is somewhat effective in times series forecasting and can be observed from figure 3.7.

References

[1] N. Babu and B. E. Reddy, "Selected Indian stock predictions using a hybrid ARIMA-GARCH model," 2014 Int. Conf. Adv. Electron. Comput. Commun. ICAECC 2014, 2015.

[2] M. D. Shubhrata, D. Kaveri, T. Pranit, and S. Bhavana, "Stock Market Prediction and Analysis Using Naïve Bayes," Int. J. Recent Innov. Trends Comput. Commun., vol. 4, no.11, pp. 121–124, 2016.

[3] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep Learning for Event-Driven Stock Prediction," no. Ijcai, pp. 2327–2333, 2015.

[4] K. A. Smith and J. N. D. Gupta, "Neural networks in business: Techniques and applications for the operations researcher," Comput. Oper. Res., vol. 27, no. 11–12, pp.1023–1044, 2000.

[5] M. S. Babu, N. Geethanjali, and P. B. Satyanarayana, "Clustering Approach to Stock Market Prediction," vol. 1291, pp. 1281– 1291, 2012.

[6] A. A, Adebiyi, A. K, Charles, A. O, Marion, and O. O, Sunday, "Stock Price Prediction using Neural Network with Hybridized Market Indicators," J. Emerg. Trends Comput. Inf. Sci., vol. 3, no. 1, pp. 1–9, 2012.

[7] S. Kayode, "Stock Trend Prediction Using Regression Analysis – A Data Mining Approach," ARPN J. Syst. Softw., vol. 1, no. 4, pp. 154–157, 2011.M.A.H. Farquad, Vadlamani Ravi, S. Bapi Raju "Churn prediction using comprehensible support vector machine: An analytical CRM application", Applied Soft Computing 19 (2014) 31–40.