STOCK MARKET PREDICTION USING MACHINE LEARNING

Sai Bhavya Reddy.T[1], Manju More[2], Sanjana.S[3], Shekamoori Pavithra[4], Shwetha.V[5]

[1,3,4,5] *Student, Department of Computer Science and Engineering,*

[2] *Assistant Professor, Department of Computer Science and Engineering, REVA University.*

Abstract

Stock Market is where trading of offers occurs for openly recorded organizations. The Stocks of the organization which are otherwise called values address a negligible portion of possession in organizations. In the current money world stock exchanging has a lot of significance. Trading of offers occur through stock trade. The most seasoned stock trade is the Amsterdam Stock trade, laid out in 1602. The Network Stock trade, Chicago Board Operations trade, American Stock trade are the top trades in United States. We have the Bombay Stock Exchange (BSE) in India that was laid out in the year 1875 and it is Asia's most memorable stock trade. We likewise have the National Stock trade or NSE [1].

Stock Market Prediction is the cycle wherein the future worth of the monetary supplies of an organization are anticipated. Presently, Almost wherever Machine Learning Techniques are utilized to anticipate the worth of stocks. AI procedures make forecasts precisely taking upsides of current financial exchange lists on preparing the machine with old records. There are numerous techniques in Machine realizing which make forecasts easier [2].

1. INTRODUCTION

In present times, it is difficult to analyze and predict the value of stock markets. Dark pools, high-frequency traders, broker-dealers and the alternative trading systems always have exchanges among them. All of them interact with each other countless number of times in a fraction of second at every instant of a day [3].

It is likewise the consequence of Regulation National Market System incompletely which is otherwise called REG NMS. It is a bunch of rules passed to refine how all recorded U.S. stocks are, by the protections and trade commission (SEC) in 2005. Reg NMS made such that no single body (may be a trade), representative vendor, or any other individual could have unbalanced command over the exchanging action. The fragmentation that exists today is created by the decentralization of the market [3].

Tim Quast the pioneer and CEO of the Modern IR and Market Structure Edge has said that Regulation National Market System needed organizations that were particular, autonomous to associate and share costs and clients, and to make that framework work needs a tremendous measure of information innovation and intricacy and where all the intricacy comes from and what has emerged out of that is what the organizations that contain this environment framework have done to adjust to that.

Factors like Physical and Psychological elements, sane and silly way of behaving and so on influence upsides of financial exchange. Additionally there are factors like market instability, reliant, autonomous to choose worth of a stock on the lookout. It truly challenging to foresee the worth of stocks in light of these variables for any securities exchange expert with high exactness. But, Machine Learning helps in predicting the stock market value by understanding the stock market data through its techniques [1].

With an idea of stock and the stock marketing, let's understand why people believe that by a Machine Learning model they can predict the stock price. The predictions of values of an observation by a machine learning model are dependent on several inputs that are predictors. The value of an observation based on several inputs that are predictors. The stock market is working similarly, that is, based on several inputs, the stock price fluctuates with these factors. So machine learning has to keep all these factors included and predict the shares of the stock market.

AI at its key level requires the client to give the already existing information utilizing this past information, the AI strategies get prepared and will actually want to investigate and anticipate the future upsides of new information. In this paper, we will utilize AI procedures like straight relapse and long transient memory organization to anticipate the future upsides of stock markets [4].

Linear Regression is a regulated learning calculation to foresee the result of a constant variable. It is an exceptionally well known measurable method to settle AI problems.it can be utilized to anticipate all out income of the organization, climate expectation, Stock cost forecast etc[1].

Long Short term memory (LSTM) is a kind of repetitive brain network for learning long haul conditions. It is ordinarily utilized for handling and foreseeing based on time series information. LSTMs have a chain like structure [5].

2. LITERATURE SURVEY

In the money world stock exchanging is quite possibly the main action.In this task the expectation of financial exchange is finished by the Support Vector Machine (SVM) and Radial Basis Function (RBF) [1].

In securities exchange expectation the point is to introduce the future worth of the monetary loads of a company.AI as such has many models however this paper centers around two generally significant of them like Regression Based model and LSTM(Long Short Term Memory Network Based Model)[2].

There has been a developing revenue in securities exchange forecast advances by means of the utilization of machine learning.This is finished by taking the ongoing upsides of the market subsequent to accepting past stock qualities as the preparation data.In this paper the AI procedures including single layer perception (SLP), Multi-Layer Perception,Radial Basis Function(RBF) and Support Vector Machine(SVM) are compared[4].

Accurate expectation of financial exchange returns is an exceptionally moving assignment because of unstable and non-direct nature of the monetary markets.In this paper it is finished by utilizing the Regression and LSVM based algorithm [5].

Stock value forecast is a significant issue in monetary application.Prediction is generally founded on the specialized investigation of the information.In this paper Stock Index is anticipated utilizing Regression and Neural Network Models under Non Normal Conditions [8].

3. METHODOLOGY

It is truly challenging and complex to anticipate stock market as there are such countless elements that it relies upon. In this paper, we use machine learning strategies called linear regression algorithm and long short term memory network as examined above by relating the past information to the ongoing information. we train the machine with a bunch of information so it will actually want to investigate and foresee any given test data[2].

3.1 Regression Based Model

Regression Based Model is used to predict the outcome of a continuous variable. It depends on the utilization of relapse calculation for anticipating right qualities. The factors that are considered here for regression are date, open, high, low, close, adjusted closing price. Here, date is the cost at which the stock began exchanging when the market opened on a specific date close is the cost of a singular stock when the stock trade shut the market for the afternoon. It addresses the last purchase sell request executed between two dealers. High is the greatest cost at which a stock exchanged during a period. Low is the most minimal cost of the period. Volume here is the aggregate sum of exchanging movement that occurred during specific timeframe. Changed shutting cost is computation change made to the stocks shutting cost, more intricate and exact than the end cost. The changes made to the end cost portrays the genuine cost of the stock on the grounds that the external elements would have adjusted the genuine price [2]

The work is completed on csv organization of information through libraries like pandas, numpy and matplotlib and so forth in jupyter journal.Linear regression is made on the data and predictions are made. Regression utilizes a linear function recently given for foreseeing continuous values [2]

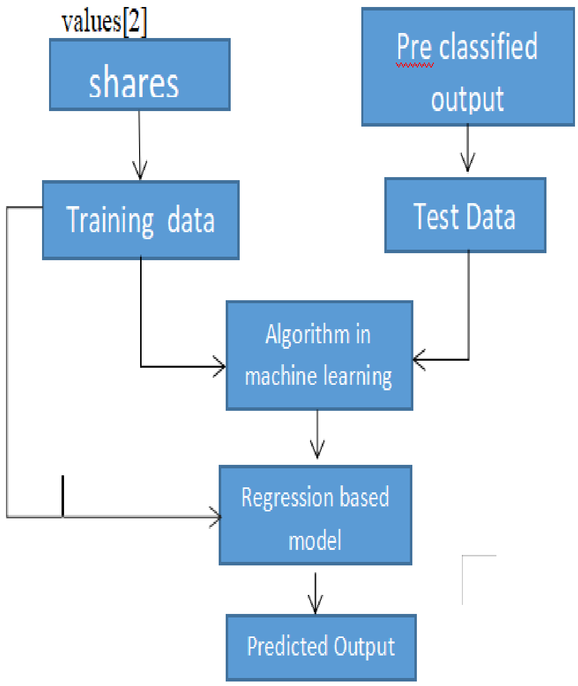

Figure1: Regression Based Model

Regression based model is generally large used to anticipate the continuous values utilizing a few given independent values.Figure-1 minimizes the error function and predicts the given values. Regression generally uses the given linear function to predict the continuous values.

$$V = a + bK + error$$

In the above function, V represents a continuous value; K addressing known independent values and a, b are the coefficients [2].

3.2. Long Short Term Memory Model

Long Short Term Memory Model is a sort of repetitive brain organization to learn long haul conditions. Long Short Term Memory Model is usually used to process and foresee based on time series information. LSTMs have a chain like structure.[2]

By and large, in this strategy initial step is utilized to settle on the data that will be prohibited from the cell in that specific time. Sigmoid capacity helps in choosing it. It thinks about the past state and current info state and registers the capacity. There are two sections in the subsequent advance, sigmoid capacity and the tanh work. The sigmoid capacity settles on which values to let through. In the third step, the last result is chosen. A sigmoid layer is executed, which concludes the pieces of the cell express those come to the result and afterward we put the cell state to the digression work and increase it by sigmoid [2].

Long Short Term Memory (LSTM) is a well versed version of recurrent neutral networks (RNN). Long Short Term Memory and Recurrent neutral networks are not same as LSTM is the advanced version or different version of RNN. Long Short Term Memory (LSTM) involve long term dependencies. Recurrent neural networks (RNN) work to find the relationship between the current and recent information [2].

The issues of customary repetitive brain organization (RNN) which can be the evaporating and detonating slopes can be given an answer for these issues as Long Short Term Memory (LSTM) Long Short-Term Memory is the knowledgeable adaptation of (RNN) design. It is utilized intended to demonstrate sequential arrangements and their long-range conditions more definitively than traditional RNNs. Generally speaking a reference to specific data put away a seriously quite a while in the past is expected to anticipate the ongoing result. Yet, "long haul conditions" can't be dealt with by RNNs. Here there could be no better command over what piece of the setting should be conveyed forward and the amount of the past should be 'neglected'. Accordingly here the LSTM is utilized where the disappearing angle issue is totally taken out, while it is left unaltered to prepare model. Long delays in specific issues are addressed utilizing LSTMs where they additionally handle commotion, disseminated portrayals, and constant qualities. The essential contrast between the RNNs and LSTMs is in the design where the secret layer of LSTM is a gated cell. LSTMs consists of 3 logistic sigmoid gates and one tanh layer. These gates have been introduced to limit the information that is passed through the cell [2].

LSTMs likewise utilize input and neglect entryway rather than two separate doors that aided in pursuing both the choices at the same time. Profound LSTM will include various LSTM in the middle between the info and result. Profound LSTM with a Recurrent Projection Layer will have various LSTM layers where each layer has its own projection layer. For the situation where the memory size is too enormous expanded profundity is very useful. It has specific layers like -

Layer = lstmLayer(numHiddenUnits) makes a LSTM layer and Num stowed away property.

Layer = lstmLayer(numHiddenUnits, Name, Value)sets extra Output exercises, boundaries and introduction, name properties utilizing at least one name-esteem pair contentions. You can determine different name-esteem pair contentions, by encasing every property name in statements.

Here at first, the output of an LSTM at a particular point of time is dependent on three things –

1. The neglect door is the initial phase simultaneously. Here we will conclude what pieces of the cell state (long haul memory of the organization).
2. The new memory organization and the info door is engaged with this progression. The objective of this progression is to figure out what new data ought to be added to the organizations long haul memory.
3. One of the person might think that we can just output the updated cell state; but however; this would be comparable to someone else who is unloading everything they learned about the stock market when only asked if the, they think it will go up or down tomorrow. To prevent this from happening we create a filter, the output gate, exactly as we did in the forget gate network.

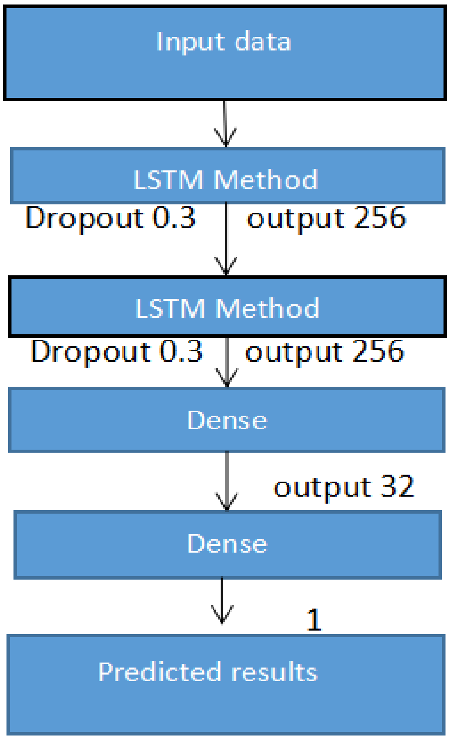

Figure 2: LSTM Layers

Figure 2 indicates that the interval of information is relatively smaller compared to LSTM. The main reason to use this model in stock market prediction is due to that the predictions depend on large amounts of data and are generally dependent on the history of the stock market. [1]

4. RESULTS

The above frameworks are prepared and tried over the dataset taken from Tesla for first strategy and google for second method. It is parted into to preparing set and testing set individually. The following results are yielded upon passing through the different models.

4.1 Regression Based Model Results

The accompanying diagram is acquired on applying the straight relapse calculation on the dataset of tesla from 29-06-2010 and 15-03-2019 to foresee shifting costs.

Figure 3: Stock Prices of Tesla

The actual stock prices are shown in blue and the predicted ones are shown by red line

4.2 LSTM Based Model Results

The following graph is obtained on applying the linear regression algorithm on the dataset of Google from 13-08-2018 and 13-08-2019 to predict varying prices.

Figure 4: Google Stock Price Prediction

Red indicates actual stock price and green indicates predicted stock price.

5. IMPLEMENTATION

Following are some pictures of implementation with inputs of Tesla and Google companies and respective outputs

5.1 Regression Based Model

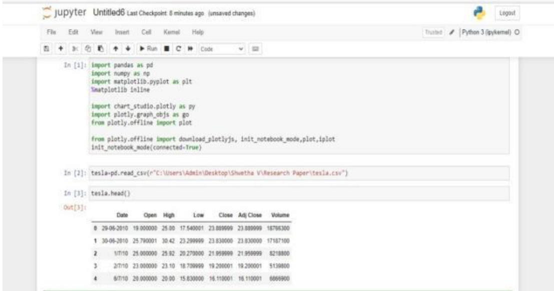

Figure 5: Top 5 rows from the dataset

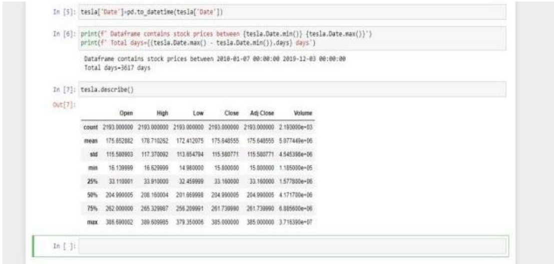

Figure 6: Some summary statistics for the data

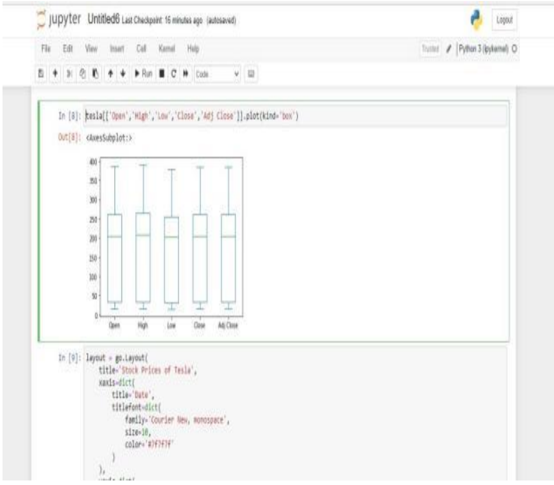

Figure 7: Boxplot

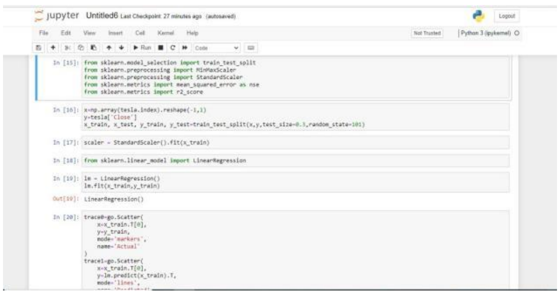

Figure8: This part gives the graph which contains actual and predicted stock values using iplot function.

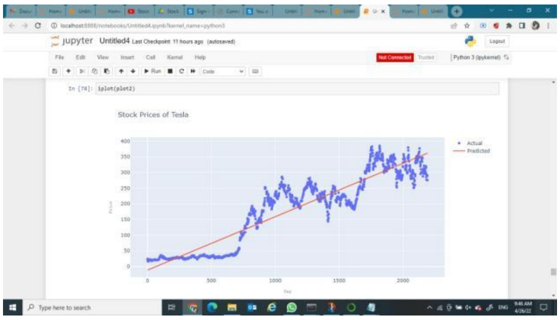

Figure 9: Actual and Predicted Stock Prices of Tesla

5.2 LSTM Based Model

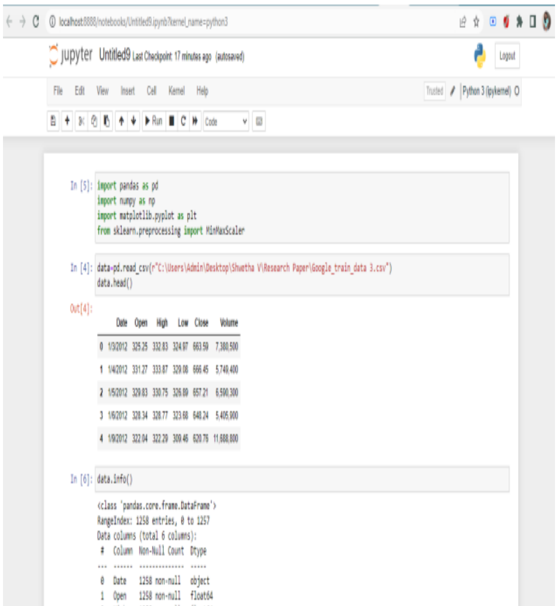

Figure 10: Top 5 rows from the dataset

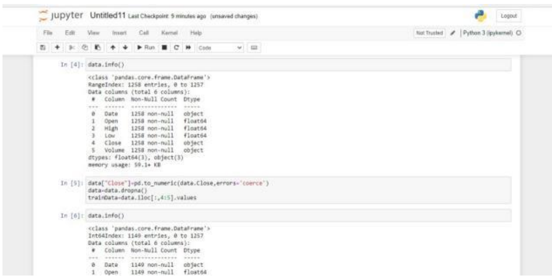

Figure 11: Total no of rows and columns in the dataset and the data types of each variable

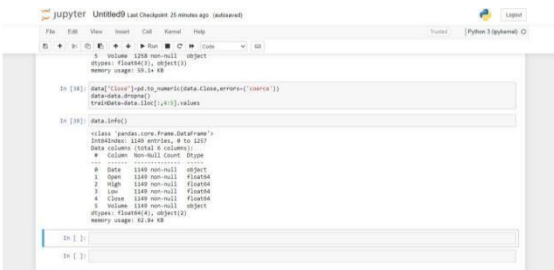

Figure 12: Change of datatype of closed column from object to float

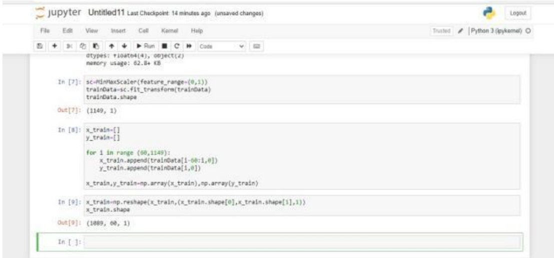

Figure 13: Reskilling the data between 0 and 1 for better performances

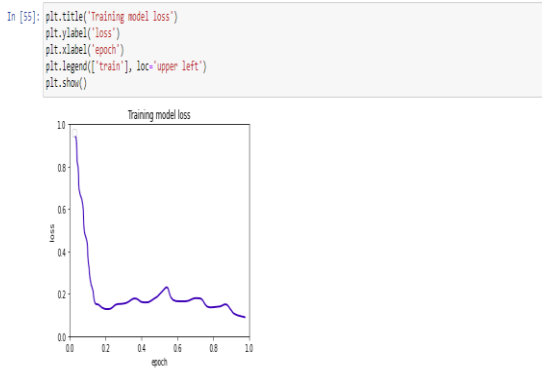

Figure 14: Visualization of loss that occurred during training process for each epoch

Figure 15: Testing of model with a new dataset

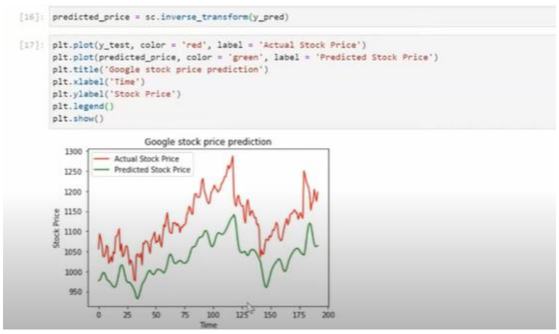

Figure 16: Actual and Predicted Stock values of Google. [14]

6. CONCLUSION

Investors commonly predict the stock prices to know the amount they get back. Generally, stock prices were predicted in traditional way by Brokers and Technical Analysts based on previous prices, volumes, price patterns and the basic trends. But, present it has become very difficult and complex task to predict stock markets because now stock markets are dependent on many other factors like social and economic conditions of the country ,natural disasters and political atmosphere etc. Sometimes, the returns that investors get back is very uncertain and disastrous and is difficult to predict using traditional methods. So, a lot of research is made on finding the methods that can predict the accurate values of stocks .Some of them are machine learning techniques like LSTM AND Regression Based Model. These kind of methods help Stock Brokers and Finance Institutions in getting good return [6].

6.1 Future Work

In the future, a lot greater dataset can be utilized than the one being utilized presently accordingly expanding the financial exchange forecast framework. This prompts expansion in the exactness of our forecast models. Further, to foresee the upsides of stock costs different models of Machine Learning can likewise be studied [2].

7. REFERENCES

[1] Rekha KB, Gowda NC, "A framework for sentiment analysis in customer product reviews using machine learning", International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 267-271, Oct 2022.

[2] Shalini L, Manvi SS, Gowda NC, Manasa KN, "Detection of Phishing Emails using Machine Learning and Deep Learning", 7th International Conference on Communication and Electronics Systems (ICCES), pp. 1237-1243, Jun 2022.

[3] M.Usmani, S.H. Adil, K.Raza and S.S.A. Ali,"Stock market prediction using machine learning techniques",2016 3rd International Conference on Computer and Information Sciences (ICCOINS), KualaLumpur,2016,pp.322-327.

[4] "LSTM and Regression Methods for Stock Market Prediction"-Nishanth Vaidya,Nikhil Bharadwaj-Students,Department of CSE-Sambhram Institute of Technology,Bangalore,India.

[5] "Stock Closing Price Prediction using Machine Learning Techniques"-Mehar Vijh,Deeksha Chandola,Vinay Anand Tikkiwal,Arun Kumar-Jaypee Institute Of Information Technology,India.

[6] H.L. siew and M.J. Nordin,"Regressioon techniques for the prediction of stock price trend ",2012 international conference on statistics in science,Business and Engineering(ICSSBE),Langkawi,2012,pp. 1-5

[7] Basha, S. M., Poluru, R. K., & Ahmed, S. T. (2022, April). A Comprehensive Study on Learning Strategies of Optimization Algorithms and its Applications. In 2022 8th International Conference on Smart Structures and Systems (ICSSS) (pp. 1-4). IEEE.

[8] "Stock Market Prediction Using Machine Learning Algorithm" Sayyed Aman,Sayyed Gulfam,Shaikh Naba-UG Students-Mumbai University-Mumbai,India.

[9] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.