International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 28 Issue 5 – MAY 2021.

STOCK MARKET PREDICTION

*K. Sai Mohith [#1], T. Mohit Venkata Naga Sai [#2], B. Ramyasruthi [#3], Dr. P. Thendral [*4]*

[#]*Student , Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Anand Nagar, Krishnankoil, India, saimohith616@gmail.com*

[#]*Student , Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Anand Nagar, Krishnankoil, India, saimohit143@gmail.com*

[#]*Student , Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Anand Nagar, Krishnankoil, India, r.msruthi1123@gmail.com*

[*]*Assistant Professor, Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Anand Nagar, Krishnankoil, India, thendral.p@klu.ac.in*

***Abstract*— Now-a-days, the important objective of the stock prediction is to evaluate the subsequent worth of money market or portfolio market of a particular specific association, company or organization. Based on the current values of the stock market which have been trained using the previous values, is now used for the purpose of advanced slope in stock market prediction using Machine Learning techniques. By doing this the predictor can get the accurate values of stock. Various algorithms and techniques can be used to make the prediction simpler, and one of the most prominent way is to use Machine Learning algorithms. The major part of this paper is based on few well known and approved algorithms for predicting namely Random Forest, LSTM, Regression and some other small techniques those who belongs to Machine Learning. These algorithms are highly used for the prediction of stock exchange. Open, close and high, low are four out of all other important factors in this project.**

Index Terms- Random Forest, LSTM, Regression

I. INTRODUCTION

A righteous prediction of the stock market can prompt a huge benefit for the predictor, the buyer and the customer. Again and again it has been proven that it is turbulent instead of being random it tends to generally be the expected and thorough with cautiously analyzing the pre historical setting for crucial and relevant stock market trading. As everyone are aware that Machine Learning is one of the most useful and effective methodology to direct some of these patterns.

To accelerate the precision using some of the various methods of Machine Learning, it is way easy to predict the market price value almost with a narrow value to the original. Many researchers, stock predictors and scholars have been extending the Machine Learning efficiency and precise accuracy for the ultimate stock market prediction. To predict stock price, it is mandatory to have the previous records that belongs to that stock. So, with the data of the past prices the buyers of that stock can prepare an algorithm and can set a estimation price that is suitable and profitable for both the buyer and customer of the stock share. Since in this project, Random forest algorithm is used the data is not needed to be concrete as the algorithm is mainly useful while some of the data is missed and even at the lose end the prediction won't get effected as the tree algorithm can pick the stock price randomly. So, there is no place for minor mistakes and there is no reason to fear about the little changes in the data. But it is advisable not to miss any data for accuracy reason. In this current project, supervised Machine Learning is used to train the dataset obtained from Infosys company through nseindia website. The trained dataset is included with five variables namely high, low, open, close and volume. Low, high, open and close are various different bid prices of the stock at different times with almost equal names. The variable Volume is the count of the shares of that particular company that moved from previous holder to present holder of the stock in that time period. And now the models will be tested on the testing data. LSTM, Random forest and Regression models are then allowed to the process individually. LSTM is to look after and remember the data along the result for further usage and Regression is to minimize the error for some variables and since Random forest is been used for higher accuracy and more randomness. And hence at last, the graphs for the randomness of the values or prices with date and period (as per Regression based model) are going to be plotted between actual and the predicted price (as it is for LSTM model). Stock price prediction for very small period windows views as a random work. Stock price have a movement for a long period of time which is generally goes a linear curve along. Customers appreciates about buying the stocks whose prices can be expected to rise in the further days. The inconsistency in the present days stock price encourages customers to buy and invest in stocks. Thus, there is always a need to predict the stock market accurately, which can be furtherly used in the present day-to-day scenario. This technique is used to foresee the stock market within incorporates a period of time to anticipate the technical analysis including machine learning and also predict the floating and uneven stock market. To predict the final value of the closing stock price, various calculations have to be considered and those calculations include depending algorithms and previous stock values of variables namely opening price, closing price and few other values must be considered. The olden time serious models were used to predict the stock price of some companies which use traditional methods to examine the previous old values. Hence the predictions calculations must be updated with latest algorithms that use different objectives and make prediction simpler. As a Regression problem, the stock market prediction is out played the performance and it is done properly only if this is served as a classification problem. Random forest plays a vital role here to examine the problem and it makes sure that the precision is correct. The executive role of this planning is to acquire the market demanding model of the stock market where Machine Learning plays a

vital role and refers the patterns of stock values with past model outcomes of the stock prices. The very executable method, Support Vector Machine (SVR) could be utilized for both the Regression and Classification. This has been seen that the Support Vector Machines are highly used in the classification situated on same issue. The SVM procedure, which has been used to plot each and every information of data components as a point in the nth dimensional space, where it has n as the count of the available data sets with a prominent shot of feature as a specified and particular coordinate and, consequently the classification is displayed by finding out the hyperplane boundaries that differentiates the two classes categorically.

II. LITERATURE REVIEW

According to this overview of the literature review, the main observation is that the various applications of machine learning methods to stock market price prediction was to attempt it completely all over the world. Machine Learning methods have been proving that it is equipping to be a much quick and accurate, while differentiating the up-to-date projection methods. A remarkable research in this field was done by many people throughout the world. A result to the work for M. Usmani, S. Adil, K. Raza, A. Ali, and by K. Raza separately, a survey for the applications of machine learning techniques which has been presented for the present furtherance in this respective field. Many more scholars contributed furtherly to this field of stock prediction by conducting experiments and simulations to survey the probability by utilizing Deep Learning methodology to predict the stock market price precisely and this work was done by K. A. Althelaya, E. M. El-Alfy, S. Mohammed. Prediction of stock market price applying neural network techniques and presenting them to the latest advancement in the current filed is done by H. Gunduz, Z. Cataltepe, Y. Yaslans. Essentially in the same way, few others namely M. Billah, S. Waheed, A. Hanifa have recommended additional improvements for stock price prediction by applying the same neural networks using a trained calculated algorithm all alone. Liu, G. Liao and Y. Ding has conducted and led the similar research to design their model with LSTM for stock price prediction including more width for enhancements to predict accurately. Some have suggested a brilliant method to handle very difficult situations that may have often emerge while working with the system and mislead the user or stock predictors to mistaken the prediction. This work has been done by K. V. Sujatha and S. M. Sundaram.

Usually, predicting the stock precisely is not an easy task. It is a challenging task to complete, but today's modern web has already proven that it is the most valuable and useful tool for predicting the stock easily. Since the entire format of the data and now that this made the predictor to extract the data using some sentiments and thus doing it simpler to demonstrate the interrelations among various variables and hardly scope out an example of expenditure. Financing design including different firms gives the signature of equality and the best thing is that the key value for predicting the stock market price successfully is to exploit the very same steadiness among the included data sets. The process that stock market price data with information of the values can be estimated successfully by utilizing not only the technical and historical data including with minor methods like using the sentimental analysis to derive the major connection but also to analyze the people emotions and how the emotions have been influencing the investment done for particular stocks. Another further significant segment of the estimating method is finding out and extracting the most important prices of specific events for web news to picturize how the extracting benefits the stock price. The far most and vast majority of the stock users specializes the fundamental involving the time series analysis during the utilized prediction. Even the users used the fascinating and unique techniques, the stock predictors cannot believe it completely, so there emerges the purpose to provide strong procedure to financial stock exchange prediction. To track down the best version of accurate outcome, the procedure to select the upgraded machine learning and artificial intelligence techniques along with regulated supervised classifier. And finally, the results will be executed on the binary characterized using SVM classifier for the different alternative data set with new feature list. The most considerable part of Machine Learning approach is that the prediction takes the role to led the business further and controls its issues for their benefit over material methods which did not include Artificial Intelligence, regardless the major fact that there would be an individual process for particular explicit issues. The Surge for intelligence improvement technique named Cuckoo search is one of the simplest method to accommodate the boundaries of Support Vector Machine. The proposed hybrid CS-SVM methodology showed the performance to pursue progressively extract the results of the variation with ANN. In the same way, the Regression-SVM shows the better performance in gauging the stock market prediction. Also. Predicting the stock market price using defined records to the process that includes computing the prediction, send it over some predefined ways to the predictor, and also finding a automatic way to perform some tasks like buying and selling the part of the shares with automation concept. And some have used Naive Bayes Algorithm for the same process.

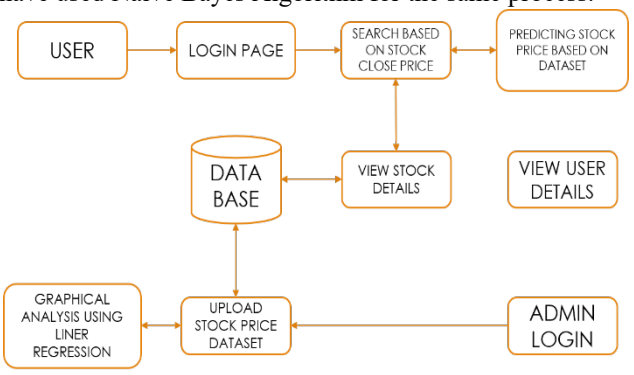

Figure 1: Flow diagram of our proposed system

It seems that Stock market prediction is a very complex problem because it appears to be a problem with no statistical strategy at first and it have more than a single factor which has to be addressed first. And yet with a proper use of Machine Learning algorithms and techniques, it is easy to compare the previous data with the present data. It is also possible to train the data sets using Machine Learning and get

some appropriate results. For this project the dataset used for prediction analysis is considered from nseindia website. These data sets were consisted of around 500 price records which are required for the analysis. Since it is a prediction algorithm it is not correct to train the datasets by dividing them. The entire data records must be trained with all the forest algorithms. The considered data is between two particular time intervals of the same year excluding the holidays. As already discussed, the various sections considered are Volume, open, close, high, low respectively. For perfect simulation and Analysis purpose, the data set from only one company is inspected. The entire data set is saved in a csv file and after uploading the data, the records are then changed into data-frames and this work is done by Pandas library in Python. By doing this process the data belongs to a stock market company is now converted as a segregated data using symbol field. Succeeding this, the regularization process of the required data is done by performing the sklearn library from Python, the data is further trained and sent for prediction. Even though Machine Learning have many algorithms, the two most prominent algorithms are used for this paper.

III. METHODOLOGY

Support Vector Machine

Support vector machine (SVM) is quite possibly the most strong and accurate techniques in all machine learning algorithms. It essentially includes support vector classification (SVC) and support vector regression (SVR). The SVC depends on the idea of decision boundaries. A decision boundary will divide a set of instances which are having different class values between two groups. SVC has the ability to support both the multi-class and binary classification. The support vector is the nearest point to the partition hyperplane, which will decide the optical separation hyperplane. In classification interaction, the mapping input vectors situated on the separation hyperplane side of the component space fall into one class, and position fall on the other class on the opposite side of the plane. We can't linearly separable in the case of data points, SVM utilizes suitable kernel functions to map them into higher dimensional spaces so that they become detachable in those spaces.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

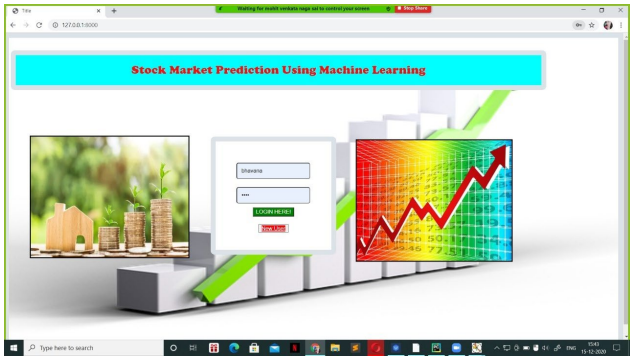

Figure 2: Home Page

Data Upload Module

Upload module all stock market details from most recent long-term dataset, there will be the volume, high, low, close, open, delivery rate. Close, open, high, low are particular offered costs for the stock at autonomous events with practically various names. The volume is the number of shares that passed starting with one proprietor then onto the next during the time span.

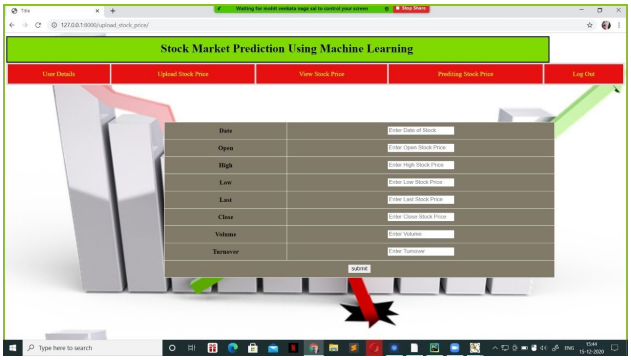

Figure 3: Data Upload Module

Regression Module

Linear regression as addressed by the above condition is performed on the information or data set and then the significant assumptions are made. The components that are considered for the regression were high, low, close, open, delivery volume, volume. The R-square assurance test was used to decide the certainty score and the stock market prices vs time predictions results were plotted to show. Best analysis of the framework.

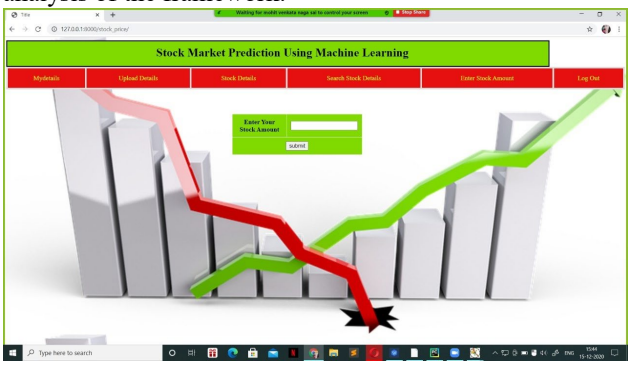

Figure 4: Regression Module

Long Short-Term Memory Based Module

LSTM is the significant level variation of Recurrent-Neural-Networks (RNN) where the information having a spot with past state continues. They are not quite the same as RNNs as they include long term dependencies or conditions and RNNs will discover the connection between the new and current data. exhibits that the interval of information is relatively smaller than LSTM. The main prepose this model in stock market prediction is that the prediction depends on the huge amount of data and dependent on the long-term history of the market. So, LSTM with control mistake by giving a manual for the RNNs through

holding information for older stages for the forecast exactly. In this manner itself it is more accurate compared to another methods.

Figure 5: Long Short-Term Memory Based Module

Stock Price Prediction Module

Since the data used for this process is more than a year, the accuracy is further improved and it can also be upgraded by using higher number of data sets than which has been used currently so that the efficiency may increase more.

Figure 6: Stock Price Prediction Module

This procedure is to try whether it is possible to determine the upcoming prices of stock market with more accuracy and constancy using the best Machine Learning techniques. Being the implementation of LSTM model as a source of estimating the stock price it became the primary contribution for the stock predicting researcher. Both the algorithms have their unique way in predicting the stock market price and they can improve the accuracy of prediction and can give beneficial outcomes using LSTM algorithms showing more functional. It can be concluded that the stock price prediction may be more accurate and efficient if the predictor is using Machine Learning techniques and can view the results as quite promising. As the size and count of the data sets are increased compared to the current scenario, in future it will very easy to find the stock price can be improved with more accuracy. As large as the data sets are the more the efficiency goes. The accuracy may also get increased by using a prominent and suitable Machine Learning algorithm with a significant approach in predicting studies, and thus there will an increment in results also. There are some other factors that change the stock price namely, the news about the company, the stock buyer's opinion and various unalike elements like Sentiment Analysis. One should also keep an eye in these areas for higher prediction numbers. Few Deep Learning techniques and the those based models should also have their place at predicting process. Basically, Random Forest is mainly used in prediction because it allows us to predict the stock accurately even though some values are missed in between the whole data. This made the prediction simpler and fear free of losing some data.

REFERENCES

[1] SachinSampatPatil, Prof. Kailash Patidar, Asst. Prof. Megha Jain, "A Survey on Stock Market Prediction Using SVM", IJCTET 2016.
[2] H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
[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), Kuala Lumpur, 2016, pp. 322-327.
[4] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1-5.
[5] S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modelling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2787-2790.
[6] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.
[7] T. Gao, Y. Chai and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.
[8] K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156.
[9] S. Qamar, and S. H. Adil, Comparative analysis of data mining techniques for financial data using parallel processing, In Proceedings of the7th International Conference on Frontiers of Information Technology,2009.
[10] Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017.
[11] Xi Zhang1, Siyu Qu1, Jieyun Huang1, Binxing Fang1, Philip Yu2, "Stock Market Prediction via Multi-Source Multiple Instance Learning." IEEE 2018.
[12] Hakob GRIGORYAN, "A Stock Market Prediction Method Based on Support Vector Machines (SVM) and Independent Component Analysis (ICA)", DSJ 2016.
[13] VivekKanade, BhausahebDevikar, SayaliPhadatare, PranaliMunde, ShubhangiSonone. "Stock Market Prediction: Using Historical Data Analysis", IJARCSSE 2017.
[14] 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.
[15] RautSushrut Deepak, ShindeIshaUday, Dr. D. Malathi, "Machine Learning Approach In Stock Market 9. Prediction", IJPAM 2017.
[16] Pei-Yuan Zhou , Keith C.C. Chan, Member, IEEE, and Carol XiaojuanOu, "Corporate Communication Network and Stock Price Movements: Insights From Data Mining", IEEE 2018.
[17] K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.
[18] S. H. Adil, and S. Qamar. Implementation of association rule mining using CUDA, International Conference on Emerging Technologies (ICET),2009.
[19] Loke.K.S. "Impact Of Financial Ratios And Technical Analysis On Stock Price Prediction Using Random Forests", IEEE, 2017.