

An Comparative Study on Stock Market Prediction Using Machine Learning Algorithms

Harini L¹, Dr.C.Thiyagarajan²

¹Post Graduate – Student, Department of MCA

²Assistant Professor , Department of MCA

^{1,2} PSG College of Arts & Science, Coimbatore, India

Abstract- Stocks are possibly the most popular financial indicator assumed for building wealth and are the best part of any investment portfolio. From last few decades, there seen severe increase in the moderate person's interest for stock market. it is foremost to have a very accurate prediction of a future trend. As a result of the financial change and stored Gain, it is required to have a secure prediction of the values of the stocks. Market systems are so composite that they overcome the ability of any discrete to predict. But it is critical for the investors to identify stock market price to generate notable profit. Commodity Prices (crude oil, gold, silver), Market History, Bombay Stock exchange(BSE) and Foreign exchange rate (FEX) that impacts the stock trend, as given attributes for various machine learning [ML]models[1]. Artificial Neural Network (ANN) forms a useful tool in predicting price movement of a particular stock. pricing parallel to high degree of accuracy and be deployed to generate profits with sufficiently large amounts of data, for preference in times of low volatility and over a short time period[5] a machine learning technique called Support Vector Machine (SVM) to predict stock prices for the large and small capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies[7]. To observe the patterns in stock prices which can be helpful in the future prediction of stocks with other learning algorithms to improve the accuracy of such prediction systems.

Keywords- Stock market, Prediction, Stock price, Bombay Exchange(BSE),Foreign exchange rate(FEX),Machine Learning, Artificial Neural Network, Support Vector Machine(SVM).

I. INTRODUCTION

Prediction of stock market revert has been one of the most difficult problem. By its identity the stock market is mostly composite (non-linear) and uneasy. In the current years, increasing eminence of machine learning in other industries have informed numerous traders to apply machine learning techniques to the area of activity, and some of them have construct quite favourable results. In which there will be a dataset storing all actual stock prices and data will act

towards as training sets for the program. The main motive of the prediction is to reduce uncertainty associated to financing decision making. Stock prices are examined to be a very dynamic and unsecured.. There has been countless attempts to identify stock costs with Machine Learning. The focus of each research compared articles varies a lot in three ways [7]. In the prediction there are two types like fake and a false prediction forecasts thus which is used in stock market prediction system. In false assumption they have define some set of rules and identify the future price of shares by computing the average costs. In the real time prediction essential use of internet and saw ongoing cost of shares of the company Since in most of the pattern detection problems predicting future events (unseen data) in the continuous function that enables us to learn any of the complicated relationship between the input and thus the output of the system[5]. If stock market directs trend identified then we can avoid wastage of money. SMP (stock market Prediction)is a extract of predicting finality on the base of previous data. Prediction reduces the risk level to investors and enhances the confidence level for investment. If they predicted intends before reach then they can prevent loss of money. All these study work as SMP. On the basis of provable data directions, Thus we examine future trend that is called SMP.

II. LITERATURE REVIEW

All ANNs are being used in various fields of application includes business prognsois, claim gaining, bond capacity, business lack in prediction, medicine, pattern recognition, image processing, speech processing, computer vision and control systems[5].The use of the facts received from other global economic demands obtaining with machine learning algorithms aproprate to find such stock moves. different machine learning related models are introduced for predicting the regular trend of Market stocks. Numerical results assumes the increased efficiency.They often worked effort to capture the frequent changes and difficult relations present in stock market facts, particularly during volatile periods.While existing examines the identity of machine learning methods for stock market prediction, there is a need for all general defined surveying across various machine

learning models and datasets. Defining stock market price, therefore, becomes imperative for investors to yield a significant profit. researchers from different fields have examined and used more algorithms and different combination of attributes to recognize the market conditions. But these algorithms are all on the base of stock expense itself which has random effects. A present trend in stock market prediction paper involves the absorption of multiple data sources beyond historical price and capable data. The main aim of this study was to apply many Machine Learning Algorithms as possible dataset involving a particular domain, which the Stock Market, as opposed to coming up with a newer algorithm that is more efficient in portend the value of a stock. When applying Machine Learning to Stock Data, we are more volunteered in doing a Technical change to view if our algorithm can obviously predicts the stock price. Machine Learning shows a important part in evolving and reviewing the performance which helpful in underlying Analysis. This particular paper conveys this gap by offering an in- depth study that estimates the performance of unlike other present machine learning approaches in recognize stock prices, considering their strengths and control.

Table 1. Summary of ML Techniques on Stock Market Predictions:

Referen ce no	Objective	Metho dused	Result	Advanta ges	Disadva ntages
[1]	Author uses stock market prediction in the proposed system with various classificati ons such asRandam Forest, SVM, Gradient Boosting, Decision Tree.	Machin e learning ,Ad boost.	Ada Boot shows the highes t accura cy of 76.79 %.	since the correlatio n factor is highest this research still includes political, and economic reforms to get more accurate results.	But it is crucial for the investors to predict stock market price to generate notable profit.

[5]	Author uses ANN pricing to high degree of accuracy deployed to large amounts of data over a short time period.	Machin e learning algorithms like Backpropagation, Artificial Neural Networks, linear Regression used.	Feedforward Backpropagation network gives better accuracy of 0.95 %	The model succeeded in prediction of the trends of stock market with 100% prediction accuracy	Thus sufficiently large amounts of data, preferably in times of low volatility and over a short time period.
[3]	Machine Learning Algorithm specially focus on Linear Regression (LR), MS Excel as best statistical tool for graph and tabular representation of prediction results.	Linear regression, Three months moving average [3MM], and Exponential smoothing prediction is used.	Exponential smoothing predictions shows better results than linear regression.	Exponential smoothing prediction results less error greater accuracy thus we considered it best stock market predictor with general trend analysis.	Prediction about future trends in any stock market cannot be considered as 100% accurate.
[7]	To predict the classification model which is based on supervised ML and DL Techniques.	Support Vector Machines and Backpropagation Neural Networks used	LS- 97.08 %, SVM - 97.07 %, RF- 88%	Various machine learning based models This numerical model results provides High efficiency and profit.	Indisputably, the stock indices is very difficult because of the market volatility that needs accurate forecast model.
[2]	This paper is used to recognize the patterns in stock prices and to improve the accuracy of prediction systems.	Support vector machine, Boosting, Linear Regression used.	SVM - 60.20 %, Boosting- 64.32 %,	The patterns in stock prices which is helpful in future prediction of stocks and improve the accuracy of prediction systems.	Thus same algorithm might give lower accuracies while predicting some other types of Stocks, say Energy Stocks.

[9]	This paper surveys the machine learning algorithms and current tools and technique suitable.	Linear Regression and Logistic Regression used	The paper predicts stock analysis and this study recommends SVM to obtain accurate results.	most accurate prediction is by learning from past instances best possible with machine learning techniques .	Since, maximum likelihood calculations are less accurate at low sample sizes in comparison to estimate coefficients correctly .
[6]	It tries to predict future stock prices using machine learning techniques which are used to get maximum accuracy .	It uses linear regression, SVM and ANN Technique is used with backpropagation.	ANN gives better results than linear regression and SVM algorithm.	This study predicts trends of market will help the regulators of the market in taking corrective measures.	Only attributes having clear relationship can be supplied as attribute Or else accuracy percentage is dropped.
[8]	It analyzed and compared based on type of data used as the input, and machine learning technique used for prediction.	Three algorithms Decision Tree, SVM, Logistic Regression were compared.	DBT - 76.9%,- SVM - 59%, Logistic Regression- 65.4%.	Inclusion of both market and textual data from online sources improve the prediction accuracies .	Difficulties are increasing in the Financial market analysis

aggregation of buyers and sellers (a loose network of economic transactions, not a physical facility or discrete entity) of stocks (also called shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange as well as those only traded privately[1]. Examples of the latter include shares of private companies which are sold to investors through equity crowd funding platforms. Stock exchanges list shares of common equity as well as other security types, e.g. corporate bonds and convertible bonds[6].

1.2 Stock Trade

Trade in stock markets means the transfer for money of a stock or security from a seller to a buyer[2]. This requires these two parties to agree on a price. Equities (stocks or shares) confer an ownership interest in a particular company. Participants in the stock market range from small individual stock investors to larger trader investors, who can be based anywhere in the world, and may include banks, insurance companies, pension funds and hedge funds. Their buy or sell orders may be executed on their behalf by a stock exchange trader[6].

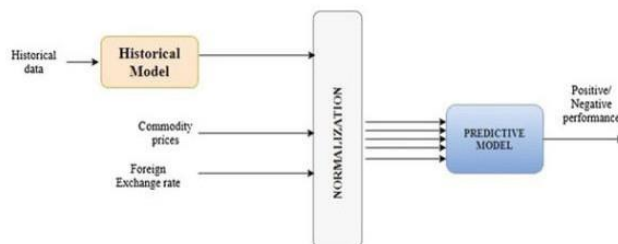
1.3 Stock Market Predictions

Stock market prediction(SMP) is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange[3]. The successful prediction of a stock's future price could yield significant profit. The efficient-market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable. Others disagree and those with this viewpoint possess myriad methods and technologies which purportedly allow them to gain future price information[6].

II. METHODOLOGY

Stock market prediction seems a complex problem because there are many factors that have yet to be addressed and it doesn't seem statistical at first. But by proper use of machine learning techniques, one can relate previous data to the current data and train the machine to learn from it and make appropriate assumptions. Machine learning as such has many models but this paper focuses on two most important of them and made the predictions using them.

Stock Prediction Using Machine Learning Algorithms

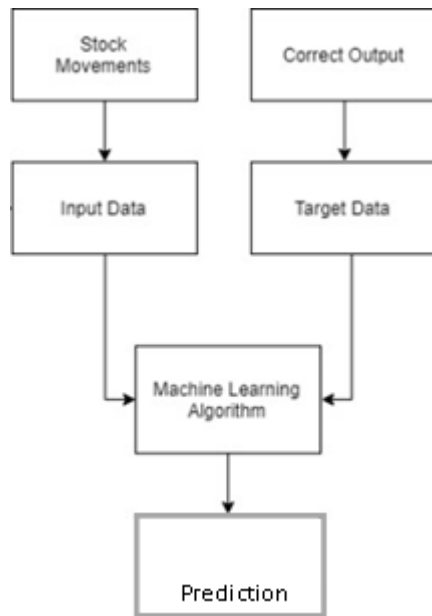


1.1 Stock Market

A stock market, equity market or share market is the

2.1 Learning Environment

In this paper I compared with previous articles that the Data Mining Environments were used for carrying out the experiments. The general setup used is as follows:



2.1 Learning Environment

In this paper I compared with previous articles that the Data Mining Environments were used for carrying out the experiments. The general setup used is as follows:



2.2 Steps for Stock Market Prediction

Step 1: This step is important for the download data from the net. We are predicting the financial market value of any stock. So that the share value up to the closing date are download from the site.

Step 2: In the next step the data value of any stock that can be converted into the CSV file (Comma Separate Value) so that it will easily load into the algorithm.

Step 3: In the next step in which GUI is open and when we click on the SVM button it will show the window from which we select the stock dataset value file.

Step 4: After selecting the stock dataset file from the folder it will show graph Stock before mapping and stock after mapping.

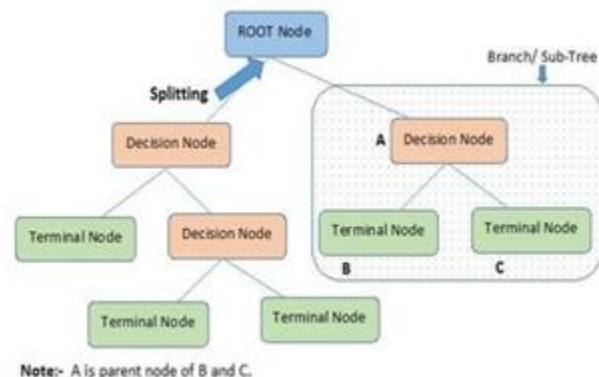
Step 5: The next step algorithm calculated the log2c and log2g value for minimizing error. So, it will predict the graph for the dataset value efficiently.

Step 6: In final step algorithm display the predicted value graph of select stock which shows the original value and predicted value of the stock.

III.VARIATIONS OF MACHINE LEARNING ALGORITHMS

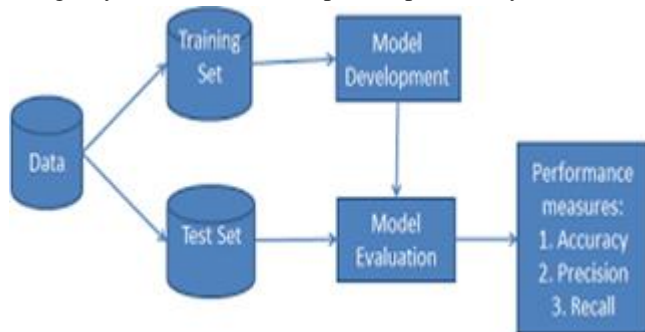
a. Decision Tree:

Decision Tree is a supervised learning method, which is used for resolve classification problems. The goal of the method is forecast the class value of get target variable. This gives decision tree a lead of choosing the most compatible hypothesis among the training dataset. Input: training data set **Output:** decision model (tree structure) [2].



Naive bayes:

Naive Bayesian method takes and improvised the dataset as input, performs analysis and implies the class label using Bayes Theorem. It computes a probability of



Formula

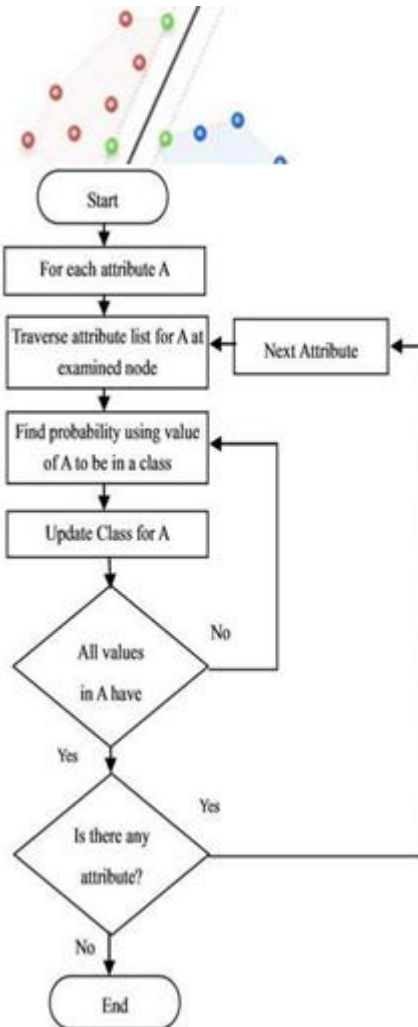
$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

- p (c |x) is the eventual probability of class (target) given the (attribute)
- P(c) is the superior probability of class.
- P (x | c) is the liability which is the probability of forecast -ting given class.
- P(x) is the leading probability which forecast Support vec -tor machine.

b. Support Vector Machine (SVM):

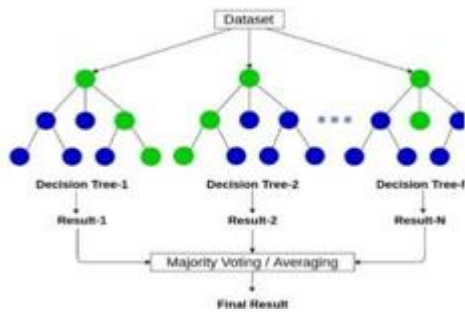
SVM is a supervised learning, differential classification technique. This technique can be used for both regression and classification. The SVM training algorithm constructs a model that conserves new samples to one of the classes [2]. SVM can distinguish both continuous and discrete data as it automatically normalizes the data before they are modelled. It was actually developed for solving the binary classification problems. Its usability has now extended to make it suitable to support multi class data and regression problems. However, primarily it is used for classification problems in machine learning. SVM can also be used in the KNN classifier. It becomes difficult to imagine when the number of features exceeds more. SVM have their unique way of implementation as compared to other machine Learning Algorithms. Lately, they are extremely popular because of their ability to handle multiple continuous and categorical variables.



c. Random forest:

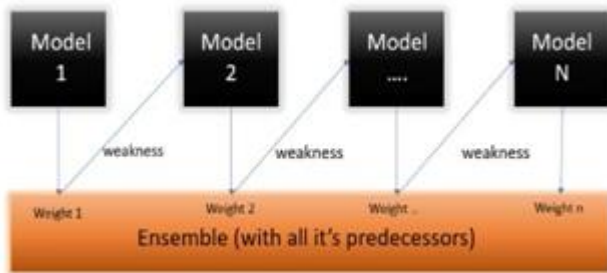
Random Forest is a supervised learning which is used for both classification and Regression. The random forest is the proceeding process of ruling the root node and separating the component node will run at any rate. The Steps given below are:

- a. Load the data where it consists of “m” features emphasis characterizing the attribute of the dataset.
- b. The training algorithm of random forest is also called as bootstrap algorithm or extending technique to select n component at any rate from m features(i.e.) to create arbitrary samples, this model trains the new sample to out of bag sample (1/3rd of the data) used to decide the impartial OOB error.
- c. Compute the node d using the best split.Divide the node into sub-nodes.
- d. Repeat the steps, until n number of trees.
- e. Compute the total number of votes of each tree for the forecasting target. The highest majority class is the final projection of the random forest.



d. Adaboost

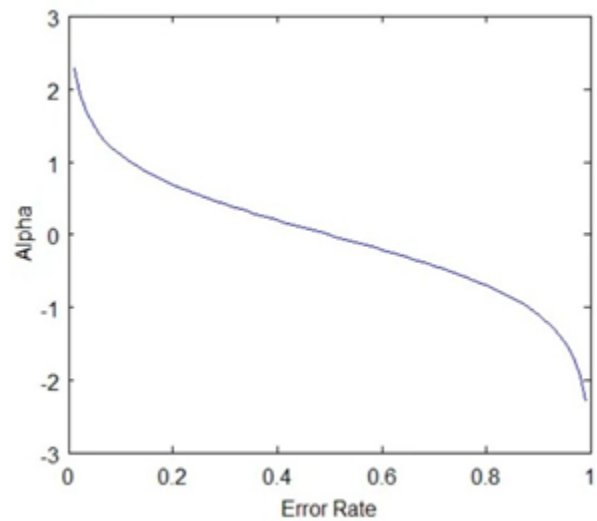
AdaBoost is a boosting algorithm that also works on the principle of the stagewise addition method where multiple weak learners are used for getting strong learners. The value of the alpha parameter, in this case, will be indirectly proportional to the error of the weak learner, Unlike Gradient Boosting in XGBoost, the alpha parameter calculated is related to the errors of the weak learner, here the value of the alpha parameter will be indirectly proportional to the error of the weak learner.



$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$$

Alpha_t is weight assigned to classifier. Alpha_t is calculated as follows:

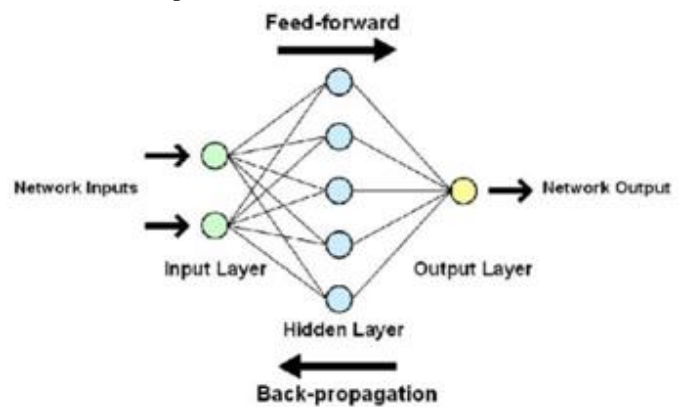
Alpha_t = 0.5 * ln((1 - E)/E) : weight of classifier is straight forward, it is based on the error rate E. Initially, all the input training example has equal weightage.



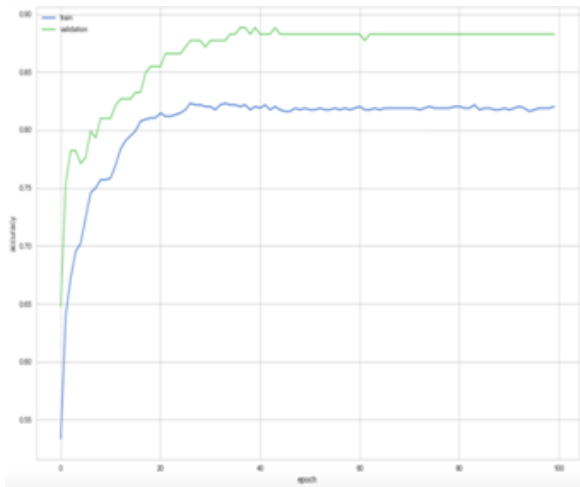
e. Artificial Neural Network(ANN):

Artificial Neural Networks are a special type of machine learning algorithms that are modeled after the human brain. That is, just like how the neurons in our nervous system are able to learn from the past data, similarly, the ANN is able to learn from the data and provide responses in the form of predictions or classifications.

ANNs are nonlinear statistical models which display a complex relationship between the inputs and outputs to discover a new pattern.



The succession of the network is directly proportional to the chosen instances, and if the event can't appear to the network in all its aspects, it can produce false output.



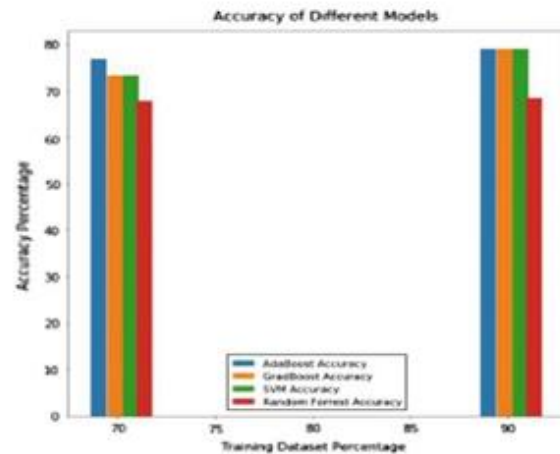
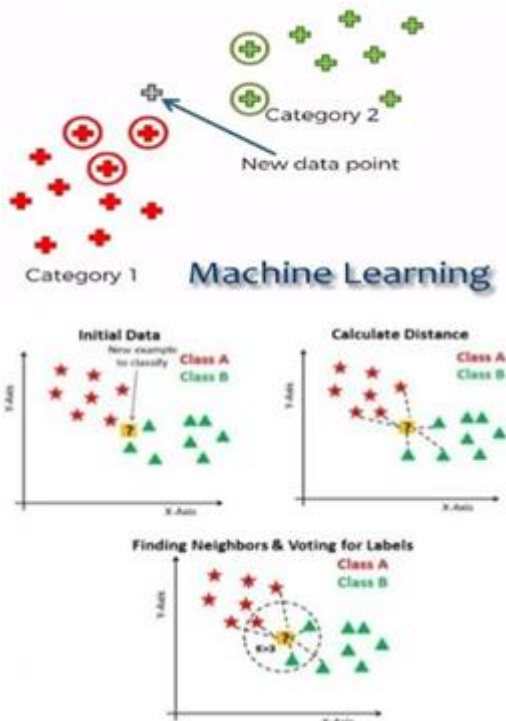
IV. RESULTS AND DISCUSSIONS

With respect to applications of machine learning and data mining methods the present study was reviewed in stock price research. Thus, each article was accordingly categorized and analyzed. Article comparison was carried out based on the previous retrieved articles. The above techniques are analyzed which gives an insight of various ML Algorithms. In contrast to other classification models Random Forest, Naïve bayes, Support vector machine (SVM) which results and performs the better accuracy of overall analyzed papers. The top performing was support vector machine which it generates prediction of stock market price thus it is acknowledged that they are not statistically significant with price. According to some surveys which have been published. The authors described only the studies related to Decision Tree, Support Vector Machine, Artificial Neural Network and some DL techniques. It also surveyed the main well-known classification techniques to predict Stock price.

f. K -nearest neighbour (KNN):

KNN is a classification method which classifies the new sample based on closeness measure or distance measure. The steps for KNN are given below:

1. Training aspect of the algorithm consists of only conserving the feature sample and class label of training sample.
2. Classification aspect: the user has to describe a “k” number of the class labels, so the unlikable sample can be classified into the determined class based on the feature comparison.
3. Mass of voting classification occurs for unlikable class. The value of the k can be selected by various takings like heuristic method.



	VarName1	VarName2	VarName3	VarName4	VarName5	VarName6
251	132.00999	193	191.2	192.14999	3878600	187.30965
252	189.8	192.81	189.74001	192.27	5419700	187.42664
253	187.25	190.7	187.00999	190.00999	11255500	185.22356
254	198.05	196.71001	195	196.30999	8527300	191.45259
255	195.98	197.41	195.42	197.02	5351300	192.05696
256	196.24001	198.24001	195.88	197.77	5431100	192.78809
257	195.03999	197	194.27	195.19	4835000	190.27307
258	196.06	199.21001	195.66	195.67999	8417900	190.75072
259	193.88	197	193.80001	196.64	5480100	191.68654
260	193.89	194.17	192.89	193.28999	5125000	188.42093
261	191.72	195.31	191.57001	194.52	6740500	189.61995
262	193.12	193.97	191.28	191.77	6089900	186.93922
263	193.2	193.38	192.35001	192.69	4091100	187.83604
264	194.24001	194.5	192.49001	193.55	4924100	188.67439
265	193.12	195.13	192.78	194.5	5394100	189.60545

Stock dataset for IBM Inc. in CSV file

V. CONCLUSION AND FUTURE SCOPE

In this paper we come to analysis that they have trained various ensemble The Correlation based feature method improvised the performance of the Model to determine the best and most stock Prediction Algorithm, a variety of various Algorithms and combinations of algorithms can be examined by the existing analysed Models. We can use Artificial intelligence in financial markets and addressing concerns related to algorithmic trading and market manipulation. Deep Learning and Reinforcement learning for future enhancements to predict stock price of larger datasets to examine higher accuracy. This Overview helps to provide a clear-cut view of stock market prediction and helps to frame better. More Over we have analysed and evaluated different Schemas for Optimal performance and results. This may be capable to predict the chances of price of stock market and gives the first- class getting to know set of rules with better accuracy comparatively. The core objective of future is to enhance the accuracy of predictive model Incorporating alternative data sources such as satellite imagery, social media sentiment analysis, and real-time news data can provide richer inputs for prediction models. This accuracy can be increase by improving the performance of the data, the algorithms or even by algorithm tuning. This is a limitation of this research.

REFERENCES

- [1] https://scholar.google.co.in/scholar?q=related:8UOCFr00C20J:scholar.google.com/&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694500550548&u=%23p%3D8UOCFr00C20J
- [2] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Machine+Learning+Techniques+for+Stock+Prediction&btnG=#d=gs_qabs&t=1694500830128&u=%23p%3D_GImEyqZXccJ
- [3] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Market+Prediction+Using+Machine++Learning%28ML%29Algorithms+M+Umer+Ghania+%2C+M+Awaissa++and+Muhamma&btnG=#d=gs_qabs&t=1694500772815&u=%23p%3DCKj_J1kud4sJ
- [4] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Market+Prediction+Using+Machine+Learning&btnG=#d=gs_qabs&t=1694501059007&u=%23p%3DFmBqZt4QgRUJ
- [5] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Market+Prediction+Using+Feed-forward+Artificial++Neural+Network&btnG=#d=gs_qabs&t=1694501109795&u=%23p%3DMYYA-ZEECfMJ
- [6] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=STOCK+MARKET+PREDICTION+USING+ANN&btnG=#d=gs_qabs&t=1694501153097&u=%23p%3Dv6G_biyPMxwJ
- [7] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Market+Prediction+Using+Machine+Learning&btnG=#d=gs_qabs&t=1694501212059&u=%23p%3DIAtz1xI1s_UJ
- [8] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Market+Prediction+Using+Machine+Learning+Techniques%3A+A+Decade+Survey+on+Methodologies%2C+Recent+Developments%2C+and+Future+Directions&btnG=#d=gs_qabs&t=1694500698984&u=%23p%3D6ASAYsLaouYJ
- [9] https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&scioq=stock+prediction+using+machine+learning+algorithms+research+articles&q=Stock+Prediction+using+Machine+Learning+a+Review++Paper&btnG=#d=gs_qabs&t=1694500985629&u=%23p%3DEPVP9B6lJmKJ
- [10] https://scholar.google.co.in/scholar?q=Stock+market+prediction+using+machine+learning+research+articles&hl=en&as_sdt=0&as_vis=1&oi=scholar#d=gs_qabs&t=1694526952150&u=%23p%3Dt1eWMERrflkJ
- [11] <https://www.ijraset.com/research-paper/stock-price-prediction-using-machine-learning>
- [12] <https://www.hindawi.com/journals/sp/2022/4808088/>
- [13] <https://revistas.usal.es/cinco/index.php/2255-2863/article/view/ADCAIJ20198497116>
- [14] <https://www.internationaljournalsrsg.org/IJCMS/paper-details?Id=53>
- [15] <https://www.sciencedirect.com/science/article/pii/S1877050920307924>
- [16] https://scholar.google.co.in/scholar?start=10&q=stock+market+prediction&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694530974997&u=%23p%3Dt1eWMERrflkJ
- [17] https://scholar.google.co.in/scholar?start=10&q=stock+market+prediction&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694531076372&u=%23p%3DcVAumxt7bYIJ
- [18] https://scholar.google.co.in/scholar?start=10&q=stock+market+prediction&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694531191789&u=%23p%3DwLzyVvLxew0J
- [19] https://scholar.google.co.in/scholar?start=20&q=stock+market+prediction&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694531244589&u=%23p%3D3jxjtWW7egJ

- [20] <https://www.sciencedirect.com/science/article/pii/S1877050921021128>
- [21] https://scholar.google.co.in/scholar?q=stock+market+prediction+using+ml&hl=en&as_sdt=0&as_vis=1&oi=scholar#d=gs_qabs&t=1694531799147&u=%23p%3DDI1xT0VdXoAJ
- [22] https://scholar.google.co.in/scholar?q=related:CKj_J1kud4sJ:scholar.google.com/&scioq=stock+market+prediction+using+ml&hl=en&as_sdt=0,5&as_vis=1#d=gs_qabs&t=1694531900046&u=%23p%3DJ9gUC-fag4MJ