Stock Price Prediction

Dissertation submitted in fulfilment of the requirements for the Degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

SANJANA DEVI

11617076

Supervisor

VIRRAT DEVASER

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

(December, 2017)

LOVELY PROFESSIONAL UNIVERSITY

TOPIC APPROVAL PERFORMA

School of Computer Science and Engineering

Program : P172::M.Tech. (Computer Science and Engineering) [Full Time]

COURSE CODE : CSE548 **REGULAR/BACKLOG :** Regular **GROUP NUMBER :** CSERGD0332

Supervisor Name : Virrat Devaser **UID :** 14591 **Designation :** Assistant Professor

Qualification : MTech CSE **Research Experience :** 05 years

SR.NO.	NAME OF STUDENT	REGISTRATION NO	BATCH	SECTION	CONTACT NUMBER
1	Sanjana Devi	11617076	2016	K1637	9855321121

SPECIALIZATION AREA : Programming-I **Supervisor Signature:** 14591

PROPOSED TOPIC : Stock Price prediction using Predictive Analysis

Qualitative Assessment of Proposed Topic by PAC		
Sr.No.	Parameter	Rating (out of 10)
1	Project Novelty: Potential of the project to create new knowledge	7.00
2	Project Feasibility: Project can be timely carried out in-house with low-cost and available resources in the University by the students.	7.00
3	Project Academic Inputs: Project topic is relevant and makes extensive use of academic inputs in UG program and serves as a culminating effort for core study area of the degree program.	7.25
4	Project Supervision: Project supervisor's is technically competent to guide students, resolve any issues, and impart necessary skills.	7.50
5	Social Applicability: Project work intends to solve a practical problem.	7.25
6	Future Scope: Project has potential to become basis of future research work, publication or patent.	7.50

PAC Committee Members		
PAC Member 1 Name: Kewal Krishan	UID: 11179	Recommended (Y/N): Yes
PAC Member 2 Name: Raj Karan Singh	UID: 14307	Recommended (Y/N): NA
PAC Member 3 Name: Sawal Tandon	UID: 14770	Recommended (Y/N): NA
PAC Member 4 Name: Dr. Pooja Gupta	UID: 19580	Recommended (Y/N): Yes
PAC Member 5 Name: Kamlesh Lakhwani	UID: 20980	Recommended (Y/N): NA
PAC Member 6 Name: Dr.Priyanka Chawla	UID: 22046	Recommended (Y/N): Yes
DAA Nominee Name: Kuldeep Kumar Kushwaha	UID: 17118	Recommended (Y/N): Yes

Final Topic Approved by PAC: **Stock Price prediction using Predictive Analysis**

Overall Remarks: Approved

PAC CHAIRPERSON Name: 11024::Amandeep Nagpal **Approval Date:** 04 Nov 2017

11/28/2017 1:38:10 PM

ABSTRACT

Data Forecasting is not a simple process. Number of steps is included in it like data collection, objective identification, feature selection, data analysis predicting future outcome etc. data forecasting has a large number of application in today's era of on demand world. Data forecasting is basically used in the weather forecasting, flood and earthquake prediction, predicting future market price like stock price prediction etc... Stock market prediction is one of the favorite topics among researchers from last few decades. Number of methods and tools are used to predict the stock market price like fuzzy logic, neural network, machine learning, R programming. Various algorithms are used to implement this techniques more accurately like naïve Bayes, genetic algorithms etc... The only motive is to increase the accuracy of predictive the stock market price. Here we are going to implement some cloud based tools to predict stock market price in more accurate way.

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the Master of Technology(Computer Science & Engineering) Dissertation/dissertation proposal entitled "**STOCK MARKET PRICE PREDICTION USING PREDICTIVE SERVICES"**, submitted by **SANJANA DEVI** at **Lovely Professional University, Phagwara, India** is a confide record of her original work carried out under my supervision. This work has not been submitted elsewhere for any other degree.

Signature of Supervisor

Mr. Virrat Devaser

Date: _____

ACKNOWLEDGEMENT

All acclaim for the sake of all-powerful God, who give us in the obscurity and help in troubles. The thesis is the consequence of full semester of work whereby I have been joined and bolstered by many individuals. It is a lovely perspective to that I have the chance to offer my thanks for every one of them.

I am additionally to a great degree obliged to my guide **Mr. Virrat Devaser** (Assistant educator, Department of Computer science, Lovely Professional University, Phagwara). I am especially appreciative to **Mr. Virrat Devaser** for picking me as an understudy at the basic phase of my master degree. I warmly express gratitude toward him for his important guidance, useful feedback and his broad talks around my work.

I grow my gratitude to my loved ones who constantly kept my spirits up with their broadened love, love and support at the season of my venture work.

Finally however not the slightest, I might want to pay high respects to the creators whose work I have counseled all the time amid my undertaking work. Also, I might want to say thanks to Lovely Professional University that gave me the street to the finish of my degree in this specific field.

DECLARATION

I hereby declared that the thesis work entitled, **STOCK MARKET PRICE PREDICTION USING PREDICTIVE SERVICES** submitted for the Master of Technology(Computer Science & Engineering) Degree for Dissertation-II is altogether my unique work and all thoughts and references have been properly recognized. It doesn't contain any work for the references have been appropriately recognized. It doesn't contain any work for the honor of some other degree or confirmation.

Date: _____

Signature of Scholar

Sanjana Devi

Reg. No. 11617076

Table of Contents

List of Figures

CHAPTER-1
INTRODUCTION

Data Forecasting is the hot topic in the area of research that can be either in the stock or medical field etc. Forecasting is nothing but the process of predicting the future value by analyzing the present and past data. Forecast is only an estimation of some factor an incentive at some predefined future day and age. Data forecasting can be done in various different areas like weather forecasting, in medical field, to predict that whether a particular patient can have specified disease or not , in school colleges , to predict the student performances etc. Using historical data to predict the future trends is nothing but data forecasting. Business person, companies uses forecasting to allocate their budget and plan their future projects to maximize their profit, to analyze their performance and risks what they can face in future. Investors trying to predict that either events effecting company such as sales increases or decreases either it effect the shares of company or not. In Stock market, analyst use forecasting to predict GDP, unemployment changes, or how it changes in a year. A statistician uses forecasting in every possible situation wherever it requires forecasting.

1.1 Different Stages in Forecasting

Forecasting is not a simple procedure as we think; numbers of steps are followed in the forecasting to predict something. Various steps are included as follow:

- First we have to analyze the situation and identifying the variables on the basis of the situation that we want to predict.
- After analyzing the situation we have to select a dataset, and according to our requirement modifications are done in the dataset.
- Then that dataset is analyzed to predict the future value.
- At last, verification is done to compare the predicted value with the accurate results to validate the procedure in more accurate way.

1.2 Application of Data Forecasting

As we all know that everything cannot be predicted, but if factors that are related to what we want predict, and sufficient amount of data is there, then we can easily predict the thing. Data Forecasting has wide range of applications like as follow:

- Egan Forecasting i.e. process of controlling the heat of building by calculating the demand of heat that should be supplied to the building according to climate and building requirement.
- Forecasting is used in stock market and foreign exchange.
- In business, to predict the customer demand according to the current market analysis
- For earthquake prediction etc.
- In supply chain management, which project should be at which time and at which place, it will help to maximize the profit.
- In Banking or financing, to check the defaulter risk i.e. either there is probability to have a risk like particular person who applied for a loan can be defaulter or not.
- In Sales Forecasting, predicting whether sales will increase or decreased in future.
- Political Forecasting, predicting the results of elections.
- Weather forecasting , flood forecasting

1.3 Forecasting Methods: various forecasting methods are defined as follow:

1.3.1 Quantitative and Qualitative Methods:

Qualitative methods are subjective in nature. They are based on the judgment and opinion of the customer. If analyst doesn't having any past data, then this method is beneficial. This method is mainly used for long range decisions. Example: Historical life cycle analogy, informed judgment and opinion, market research and Delphi Method.

Quantitative methods are used to predict on the basis of past data. This method is used only if past data is available and some of pattern in which data exist in future also must be known. This method is used for short range decisions. Examples are straightforward and weighted N-Period moving midpoints, last period request, basic exponential smoothing, multiplicative occasional lists and toxic substance process show based gauging.

1.3.2 Average Approach

In this method, average of the past historical data is taken to predict the all future values. This method is applicable in any kind of data where historical data exist.

$$\hat{y}_T + h_T = \bar{y} = (y_1 + \ldots\ldots + y_t)/\mathrm{T}$$

If we want to predict the unobserved values i.e. value that is not specified in the dataset, we can predict by taking the mean of the all known values.

1.3.3 Naïve Approach

This is most efficient prediction model in terms of cost. This model is basically used for the time based data. Using this model, prediction accomplished is equivalents to the as of recently analyzed information. This technique works exceptionally well for money related and financial matters time arrangement, which by and large have design that is hard to foresee precisely and dependably.

1.3.4 Float Method

A change in naïve technique is to enable the expectation to increment or lessening with time, where the measure of variety after some time i.e. called the float is set to be the normal change in the past information

1.3.5 Time arrangement strategies

Time arrangement strategy use past information to anticipate the future results. These methods include:

- Weighted Moving average
- Extrapolation
- Exponential smoothing
- Karman filtering
- Auto Regressive Integrated Moving Average (ARIMA)
- Seasonal ARIMA(SARIMA) and ARIMARCH
- Auto Regressive Moving Average (ARMA)
- Linear prediction
- Moving average
- Growth curve (statistics)
- Trend estimation

1.3.6 Casual/Economic Forecasting Methods:

Some of the forecasting methods that are used to identifying the factors that affects the variable which is being predicted. Like change in weather can affect the sales of umbrella

etc. and sale of sweater can also be dependent on the weather. Every casual method doesn't follow the strict algorithm. Some of algorithms may be depends on the relationships between the variables of the past historical data.

1.3.7 Judgment Methods:

Judgment methods are based on the judgment and opinion estimates. This method is basically used where there is lack of historical data or some unique and new market strategies are followed. Various Judgment methods are:

- Delphi Method
- Composite method
- Cooke's method
- Scenario Building
- Technological forecasting
- Forecast by analogy
- Statistical surveys

1.3.8 Artificial Intelligence Methods:

- Support Vector Machine
- Artificial Neural Network
- Data Handling Group Method

1.4 Introduction to Stock Market

As Technology swell, there are number of techniques to solve a single problem. Stock price prediction is one of the compelling area of research from last few decades. There are multiple techniques that are used to predict stock price like Fuzzy Logic, Artificial Neural Network, Support Vector Machine, Dumpster–Shafer Theory, K-Nearest Neighbor (kNN) Algorithm, Machine Learning , using R programming etc.

Stock market sometimes known as share market or equity market is nothing but aggregation of sellers and buyers of stock or shares, which represents businesses or ownership claims including securities registered in stock exchange.

Stock market price prediction is practice to determine the value of future stock price of company or financial instrument that deals on exchange by analyzing some historical data. The successful forecast of a future stock price that could yield notable profit. Thus, the prediction of future stock price is just dependent on the historical data. We can analyze the historical data using different methods to predict the future stock price. Although there are number of different techniques to predict future stock price of a company but according to Random walk Theory, Stock price follows random movement hence it couldn't be predicted.

According to the Random Walk Theory, Burton Malkiel in 1973 said that stock market prices are well described using statistical process also known as 'Random Walk' means each and every day central value deviates in random and unpredictable way. This theory applied truly in some cases because in most of the cases the prediction technique does not provide much accurate results.

However, Stock price prediction can be done using number of techniques like using fuzzy logic we can provide a probabilistic results of upcoming outcomes, using some artificial intelligence techniques like artificial neural network where we can train the neural for analyzing the data, using SVM (Support Vector Machine), DS Theory, kNN Theory, using some programming languages also like R programming etc.

The various methods used for stock price prediction can broadly classify into three categories:

- **Fundamental Analysis Techniques:** Fundamental analyses follow the concept that for every company's growth, monetary funds are the basic requirement. If company works well, it well get more benefits or rewards in terms of monetary resources results more growth or rise in stock price. Some of the famous fundamental analysis techniques are Warren Buffett, P/E ratio etc.
- **Technical analysis (charts based):** In this, it is assumed that future price can be predicted by analyzing the trends that are followed by the past stock price.
- **Data Mining Technologies:** various data mining techniques are used in this like artificial neural network, SVM and various genetic algorithms etc.

CHAPTER-2
LITRATURE SURVEY

With the development of technology, various new theories arises and ways to solve single problems increases. With the advancement in technology and theories , stock market price prediction problem can also be solved using number of techniques. Some researchers used evaluation strategies and genetic algorithm for stock market price prediction. Some implemented Artificial neural network for price prediction and checked the effect of large amount of data on the future stock price. Neural network can be used with the different number of input layers , output layers , hidden layers to predict the future stock price. We can also implement back propagation algorithms with sigmoid functions for analyzing the prediction of stock market price.[1]

Fundamental analysis approach can also be followed some real world data analysis. There are some factors that effect changes in stock price either in government or other exchange companies. It also depends on the belief of people which can be analyzed in the form of sentiments that are expressed in social media sites and news articles etc.so sometimes sentimental analysis can be used with other techniques like neural network to predict the stock market price.

Stock price prediction can be done using various new techniques that can be based on artificial intelligence like fuzzy, neural network , support vector machine , machine learning etc., using some programming languages like R programming etc. and can be using some of the enterprise based tools. Special algorithms can be implemented like kNN , Dumpster Shafer theory[2] etc..

2.1 **Stock price prediction using Neural Network and Sentimental Analysis:**

As we know that in stock market, one can invest in number of ways. Some can invest by buying stock for long term and selling it to the others after some years or months.Some will go for intraday investment deals where transactions occurs within a day or for a day. Intraday trading is much riskier than long term investment because during a day price fluctuations can be high and not easy to predict. Lot of time and effort is required to analyze the movements among the prices in a day.

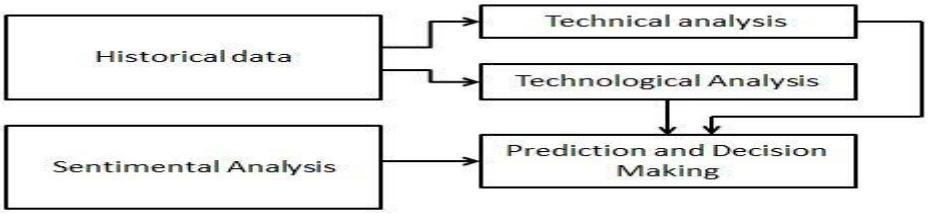

Figure 1 Stock prediction using neural network and sentimental analysis

Even there are number of methods and technical indicators than provide good estimation but still there is a chance of lack in prices.

Here three methods are combined basically i.e. technical analysis, technological method like neural network and sentimental analysis for stock market price predication and to build a trading system which will help to decide that how and when the buys and sells calls for next day will be executed. As we know, intraday trading is much riskier than long term because it requires continuously user monitoring of the stock price. An automated system not only helps the user to analyze the profits but also save a lot of time. Combination of technical analysis with technological methods helps to give the beneficial results with good performance, sentimental analysis along with system works well to provide effective results as it give analysis of day to day activities occurring in the market. System works in the way that it first trying to predict whether there is rise or fall in the stock price and on the basis of this analysis it will help to decide either to buy or sell the shares.[3]

2.2 Fuzzy Based Stock Market Prediction

Stock price prediction can be done using fuzzy logic . Fuzzy logic is basically used to analyze the long term scenario where trading occur for a long time. In this method TS fuzzy model is used i.e. Takagi-Sugeno fuzzy modeling. Basically this model is useful to analyze the prediction on the basis of data analysis. This process is based on the two model i.e. scenario model and prediction model, on the basis of which particular variables are analyzed. No prior background knowledge related to market is required. This model proved to be very successful in controlling and system modeling.

In this model we used first two models i.e. scenario model and prediction model where Scenario model represents the market behavior it helps to find the answer of questions like

how the market scenario changes with the change in price or inflation while all the other variables affecting market remain same and prediction model helps to predict the future behavior on the basis of the historical data that we are used to analyze.

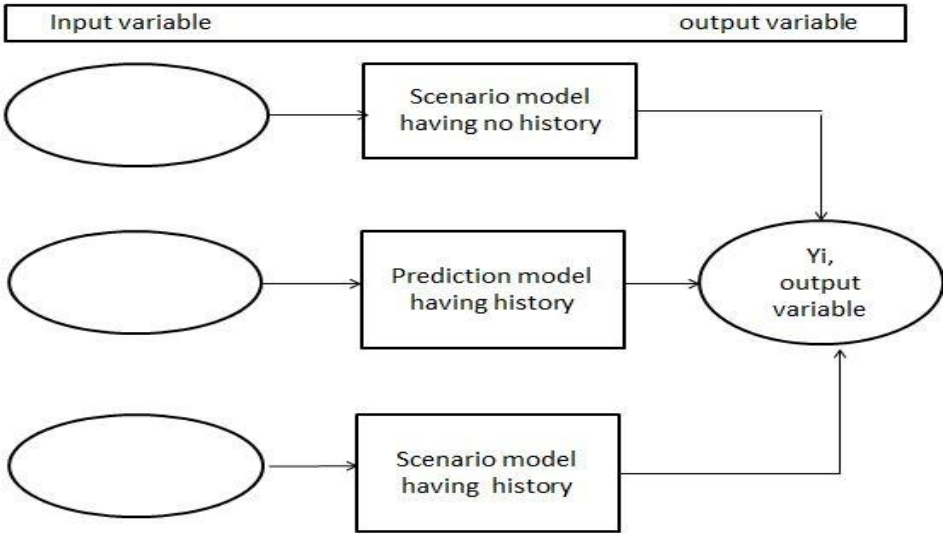

Figure 2 Stock price prediction using TS fuzzy modeling

In this way we can analyze the TS model is proved to be valuable in the case of these two model to analyze the future criteria. However we know the accurate results are very difficult to predict but still this helps a lot to give a beneficial results. For further work in this area, preprocessing can be done by calculating the monthly based log value for each variable which will help to work in more accurate way.[4]

Fuzzy model is absolutely applicable in field of stock price prediction. Fuzzy logic is nothing but a model that is based on improper and vague values. If we don't have any specified historical data in fixed format we can go with the fuzzy logic. Fuzzy logic is mainly used to work with improper and vague data to provide beneficial and good results in the probabilistic format.

In case of stock market analysis, what we want a model which can:

- Able to store and analyze the knowledge to predict the future outcome for the stock market

- Work with the vague or improper knowledge and understand the vagueness of the improper dataset
- Provide proper, declarative and effective results which help to analyze the future outcome.

A fuzzy logic will able to fulfill all the three objectives. Fuzzy logic will able to work will with the functionality of the input function where we can specify the input parameters. Second thing fuzzy model does not follow any well specified mathematical model thus it can be able to work with the vague or improper knowledge. By using different range of the membership function it will help to analyze the parameters and provide good results using the output function.

Here on the basis of three factors we are trying to analyze the stock market prediction. These three factors are business cycle, interest rates and inflation. By suing these three factors we can analyze the market return which is related to the fundamental information about the market excess return which we can get by calculation the difference between the market return and risk free rate of return. Thus in this way using fuzzy logic with these three factors as an input function we can calculate the output function which will help to analyze the value of the market excess return.[5]

This will help to predict the stock market price on the basis of the important factors that mainly affect the rise and fall in the price in the stock market. This will also help to analyze the how much probability that there is a particular rise or fall is the price or not and at this particular movement what will be the output. This method we can implement using the tool named Mat Lab i.e. matrix laboratory where we can implement artificial intelligence technique to analyze the stock market. We can also implement some automated system using fuzzy logic. Cloud based tools can also using the fuzzy logic for price prediction.

2.3 Stock Price Prediction Using The Hybrid Decision Tree- Neuro-Fuzzy System

As different researchers works a lot in stock market to predict the future outcome with the more accuracy and correctness. Stock market is the way which provides potential to get more benefit through less investment in short duration of time. This is the reason that everyone

who invests in stock market trying to predict the future outcome so that they can get to know how and when they have to invest. If we want to design an automated system for trading, then there are some aspects that must be considered like extraction of features that are required for the stock market, selection of features that provide most appropriate analysis, technical analysis method used to analyze the data for future prediction.

Here feature extraction can be done using the technical analysis, and for feature selection, decision tree is used. Features are selected in such a way that it will provide high accuracy in the prediction. Selected features are used along with the dataset for further analysis using the neuro-fuzzy system for the prediction. Using this neuro fuzzy system next day prediction can be done. That prediction must be considered with higher accuracy. This method is tested for different scenarios and it was proved that hybrid system provide more accuracy as compare to the individual models. Stock price prediction can be done using the technical analysis but it provides only limited accuracy. There are some artificial intelligence techniques that are used like neural network and fuzzy logic etc.. but all these methods results some limited accuracy. To get the higher degree of accuracy , the combination of neural network, fuzzy logic and decision tree is used to be considered as the hybrid automated system for stock trading.[6]

For this, the very first step is identifying some features and extraction that useful features by using technical analysis which are also known as technical indicators. A useful set to these technical indicators must be selected to get the accurate results. These indicators can be selected using decision tree. Basically number of sets is selected using the decision tree. The set of indicators which gives high accuracy will be selected and prediction method is implemented on that particular set of indicators.

A Neuro-fuzzy system basically is a model where artificial neural network and fuzzy logic is implemented. Fuzzy logic is used to implement IF-THEN rules and on the basis of these IF-THEN rules implementation analysis done and interconnection are done using the neural network where there is a input layer, output layer and number of hidden layers are also present.

The combination of neural network and fuzzy logic is considered in such a way that neural

network and Fuzzy Inference System worked together to give accurate results. The technical indicators that we select are considered as the input parameter for this system, which will help to make interconnection and this neural network algorithm is help to identifying the various membership functions of the FIS (Fuzzy Inference System). With the help of this member ship functions finally result are calculated to predict the future outcome. The main aim to use the decision tree is to minimize the number of features and for neuro-fuzzy system, the main target is to maximize the accuracy of predicting the next day's close price which help to analyze the stock that weather stock will rise or fall in next day on the basis of the historical data, high low, open, and average price of the day.[7] Hybrid System can also be implemented using some heuristic techniques with the neural network for predicting stock price to predict the result more accurately and minimizing the threshold.

2.4 Stock Price Prediction Using Support Vector Machine

As technology increases, new method arises to get more accuracy in the prediction of stock price. Support Vector Machine is the method to be used with the existing scenario to get more accurate and comfortable results.

By using SVM, a system can be proposed that mainly consist two parts:

- In first step, it deals with the feature selection where SVM filler is used that is based on the correlation, which is used to provide ranking to the various features. According to this ranking the most suitable feature subset is selected.
- Second step deals with the processing of these selected features, which deals with the prediction model. A quasi linear SVM model is used that work with the best selected features subset using quasi linear kernel function to predict the next day outcome.

This system not only helps to predict the stock price but also helps to control the over fitting problem in the stock market prediction. This system has various advantages like best feature subset is selected, the subset which contain only those features which are highly related to the output function so that it help to increase the accuracy. As we eliminate the least significant features which help to get most significant data and to remove the noisy data easily. Ranking helped to easily picked up the best features.[8]

The following procedure is followed for stock market price prediction:

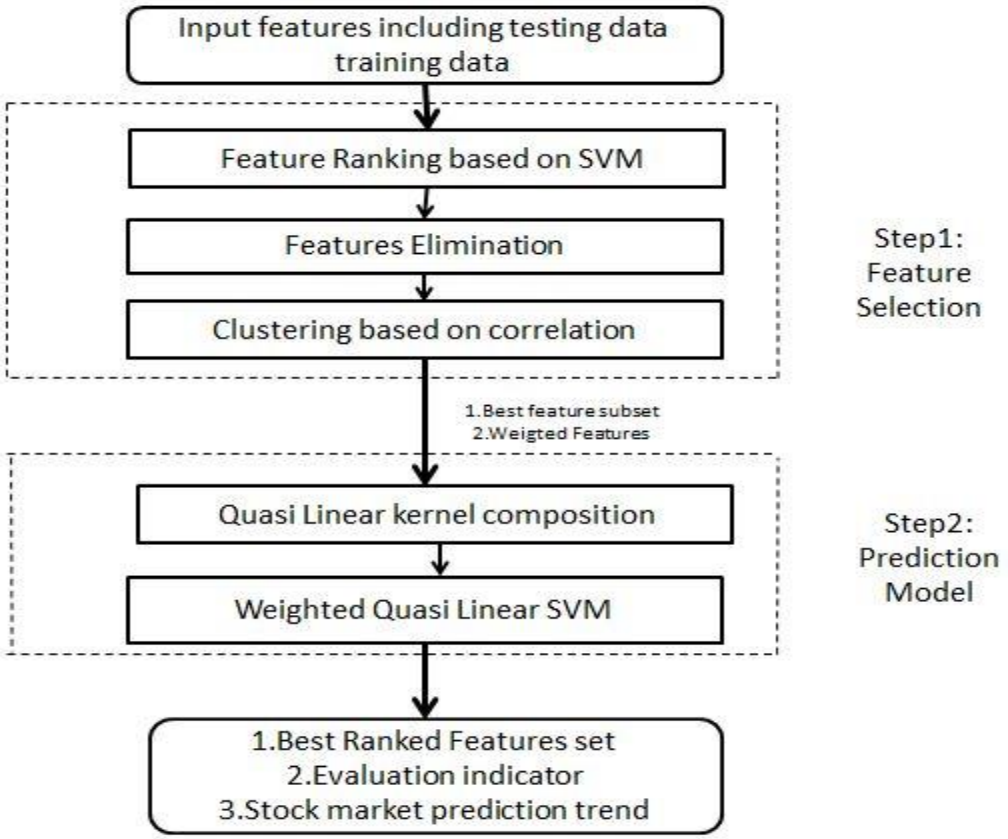

Figure 3 SVM Based Stock market price prediction

2.5 Stock Market Analysis Using Machine Learning

Machine Learning technique is used to implement artificial intelligence concept like if we want to analyze some data, if we want to predict the future outcome or if we want to train machine on the basis of past history. Machine Learning technique uses the same concept. It helps to train the machine or model on the basis of past history to analyze how model will react on the basis of different changes that may occur in future or already occurred in past. On the basis of training the result are performed. Two operations are performed mainly i.e.

- Training period, here model is being trained on the basis of past history so that if same changes occur in future then model will have to respond in same manner.
- Testing period, after the training period completed successfully it is assumed that

model I well trained and its testing is performed by providing some data for analyzing and predicting some future outcome and check if it provide true results or not.

There are various algorithms that can be used in machine learning to predict the stock price. The main aim is to train the model so that it will provide most accurate results. We not only want to predict the movement of stock price but also trying to check the ratio of movement with the fixed amount of time. For example 10% rise in stock price. A methodology named Equity stock price movement is tried to predict.[9] Equity price movement is nothing but movement in stock price in equal way. This methodology I considered as a classification techniques where if there is movement in stock market price is 10% or higher then it will be considered as good else it will categorized in bad category.

Similarly if there are large amount of stocks are there, there may be some stocks that are not more useful. So we can categorized the stock again in two categories one is good and other is bad. Some basic parameters or useful indicators are selected. These indicators can be selected on the basis of feature selection procedure. That useful or best suited features are considered as the base features and on the basis of the feature subset, various stocks are categorized. We have to categorized in good or bad category. If stock contain all essential feature or does not contain any leas significant feature we can accept that stock and put it in the category of good stock. there may be some stock that doesn't even contain single useful indicator we simply discard that stocks even without even categorized them. The stock which may contain best suited features or some unessential features may be categorized as bad stock. We can analyze that bad stock and by some modification can convert them in the category of good stock. Feature extraction can be done on the basis of rise in stock they provide. If any features provide higher rise it will selected as best as compare to other and on the basis of that features stock market price prediction is done.

2.6 Stock Price Prediction Using Naïve Method

Naïve method also known as Naïve Bayes method can also be used to predict the future stock price. Here we proposed a model which mainly includes two operation one is automation and another is prediction. Using Automation, one can fix the prices for the shares and according to it, shares can automatically buy or sell. In prediction it can be of two types, one is dummy and another is real time. In dummy prediction we can predict the value of future stock where

in case of real we can find the value of current stock price.

The conceivable market expectation objective can be the future stock cost or the instability of the costs or market drift. In the expectation there are two sorts like Dummy and Real time forecast utilized as a part of securities exchange System. In Dummy forecast we characterize a few principles and anticipate the future cost of offers by figuring normal cost. In the Real time forecast mandatory utilize web and see current cost of offers of organizations.

Dummy Prediction: in Dummy expectation partner select date and friends name and take choice rely upon benefit or misfortune. At the point when there are benefit then partner deal shares. What's more, if misfortune then partner buys shares?

Real/Ongoing Prediction: continuously forecast partner select organization name and select date to see current offers cost of chose organization.

In Automation partner perform two operations like buy shares and deal the shares. In this framework these operations are perform naturally. At the point when partner set one fix cost to deal the shares, after that framework coordinate cost of shares & consequently perform deal operation. Like that buy operation is likewise performed. In that set one fix cost to buy the shares, after that framework coordinate cost of shares& consequently perform buy shares.

Naïve Bayes calculation is a characterization system which produces Bayesian Networks for a given dataset in view of Bayes hypothesis. It expects that the given dataset contains a specific element in a class which is irrelevant to some other element. For instance, a protest is thought to be an on account of a few highlights. These highlights nearness may rely upon each other or on different highlights however the majority of the highlights nearness autonomously adds to the likelihood that this protest is A. what's more, that is the reason it is known as “Naïve”. Focal points of Naïve Bayes calculation are it is anything but difficult to assemble and helpful for extensive datasets and even known to outflank exceedingly refined arrangement systems. The following steps must be considered into the naïve Bayes algorithm:

1. The given dataset is to be changed over into a recurrence table. Ascertain probabilities of the occasions and utilizing the probabilities make Likelihood table.

2. Utilizing the Naive Bayesian condition, ascertain the back likelihood for all classes.
3. The class with the most noteworthy back likelihood is the result of forecast.

This paper provide not only the way to predict the future outcome using naïve Bayes method but also help to predict the current stock prediction and also help to check whether it will provide true results or not. We can use automated system also to sell and purchase the shares. No need to notify the stock market all the time.[10]

2.7 Using Fundamental and Technical Analysis

The Stock market process is brimming with vulnerability and is influenced by many components. The Stock market forecast is one of the critical efforts in fund and business. There are two sorts of examination feasible for expectation, specialized and essential. Specialized investigation is finished utilizing chronicled information of stock costs by applying machine learning and essential investigation is finished utilizing web-based social networking information by applying slant investigation. Web-based social networking information has high effect today than ever, it would aide be able to in anticipating the pattern of the stock exchange. The technique includes gathering news and online networking information and removing assessments communicated by person. At that point the relationship between the assessments and the stock esteems is dissected. The educated model would then be able to be utilized to make future expectations about stock esteems.

When building up a stock expectation demonstrate, ideas to be considered are

- Random walk hypothesis: This theory follow the phenomenon that stock generally works on the random method, it means drift and growth of stock price follow the random method that it is not easy to predict. It is given by the formula

$$u(t) = u(t-1) + c(t)$$

$$\Delta u(t) = \frac{u(t) + d(t) - u(t-1)}{u(t)}$$

$$u(t): stock\ price\ at\ time\ t$$

$$u(t-1): stock\ price\ at\ time\ t-1$$

,

$$\Delta u(t): change\ in\ stock\ at\ time\ t$$

$$d(t): dividend\ at\ time\ t$$

$$c(t): adjustment\ factor\ at\ time\ t$$

Efficient Market Hypothesis: It expresses that market value reflects the absorption of all the data accessible. As the new data enters the market the framework quickly enters the unequal state furthermore, anticipated right change is disposed of by the new cost. Thus given the data it isn't conceivable to foresee the future cost of the stock. Be that as it may in view of the data used to anticipate what's to come value EMH has three structures.

- Weak shape: Only the past data is considered.
- Semi solid shape: All openly accessible data is utilized.
- Strong frame: All the data openly and secretly accessible data is utilized.

Securities exchange costs could be displayed utilizing two methodologies

1) Technical: Statistical examination of the stock costs.
2) Fundamental: Considers everything about and conduct of financial operators that may influence cost. It is performed on authentic and display information, however with the goal of making monetary conjectures.

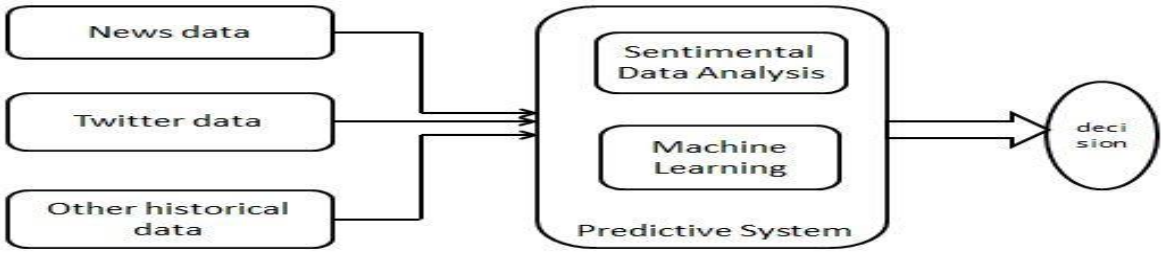

Figure 4 Stock price prediction using big data approach

A prediction model has been manufactured that utilizations enormous information scientific abilities, online networking investigation and machine figuring out how to intermittently

anticipate the drift about securities exchanges. Display demonstrates that estimation examination of the social information supplements demonstrated specialized examination strategies, for example, regression analysis. It demonstrates that unpredictability of the business sectors and the future execution of the framework is influenced by the financial and political news and impact of the online networking.

Machine learning is used to analyze the historical data and to train the model on the basis of this past data. This method is basically used to predict the future outcome.

Exploiting online networking information notwithstanding numeric information expands the nature of the information and gives made strides expectations. The associate of enormous information innovation permits forecasts at ongoing.

However the calculation utilized for assessment investigation employments summative appraisal of the slants in a specific news article or tweet, this could be enhanced for better assessment figuring, which would enhance the exactness of the expectation. In this way this will help to increase the accuracy and correctness of the model and also help to predict the efficient future outcome .here technical analysis i.e. machine learning help to train the model for the prediction of future outcome and fundamental analysis i.e. sentimental analysis of the social data which is twitter data, news related data help to increase the accuracy in the prediction of the future data. This will effect a lot because market or stock price increase or decrease with the market scenario which is mainly effected by the society or social data. In this way social sites data help to increase the correctness in the prediction which is done by technical analysis i.e. machine learning.[11]

Another Scenario we can consider using the same criteria i.e. fundamental and technical analysis. Here data can be collected from the companies such as Google and Yahoo. Expects to consolidate the regular time arrangement investigation system with data from the Google slant site and the Yahoo back site to anticipate week by week changes in stock cost. Vital news/occasions identified with a chose stock over a five-year traverse are recorded and the week after week Google drift file esteems on this stock is utilized to give a measure of the greatness of these occasions. Fundamental data is collected from news and internet whereas technical data is collected from the past historical data. The Technique used for analysis is

Text Mining. So for analysis data is collected from yahoo finance and Google trend data. In this way a correlation between stock data and sentimental data is analyzed.

2.8 Stock Price Prediction Using R Programming Language

R is a programming dialect and condition for factual preparing and designs. The R lingo is for the most part used among examiners and information excavators for measurable programming and information investigation. R dialect was made by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand and is directly kept up by the R improvement center group. In the midst of the latest decade, the vitality beginning from both the academic world and industry has lifted the R programming lingo to transform into the most basic instrument for computational bits of knowledge, recognition and information science. Around the globe, countless and data specialists utilize R dialect to deal with their most troublesome issues in the fields going from computational science to quantitative advancing.[12]

R-Studio provides effective, free and user friendly coordinated environment for R language. R-Studio enables the client to run R contents in a more easy to use condition. R-Studio can access by two ways: one is R Studio Server which is cloud based, enables user to run this application on web based platform. It can be accessed by using Linux server remotely; another is Desktop based R Studio which enables user to run this functionality by using simple desktop; R-Studio is composed in the C++ programming dialect and utilizations the system for its graphical UI. R-Studio has a general comfort where we can sort charges and see yield, it has a supervisor that backings coordinate code execution and has highlights of featuring the language structure, investigating and dealing with the workspace. R-Studio additionally has a workspace tab which incorporates all the dynamic protests, a history tab which demonstrates a rundown of orders utilized up until now, and the records tab that demonstrates every one of the documents and indexes in the default workspace.

Here a prediction model is proposed for the time arrangement securities exchange information. This model will robotize the procedure of progress of stock value files in view of specialized investigation and gives help to budgetary masters to pick the better planning for obtaining and offering stocks. Information mining methods are utilized to build up the expectation model and R programming dialect is utilized for representation of results.

Information mining can be interpreted as a learning revelation process. Information mining procedures are formulated to address the issues by giving a solid model information mining highlights. To build a model that researches the stock examples by using the past stock trade patterns; we utilize the auto-backward incorporated moving normal (ARIMA) model.

Frame Architecture is a model that characterizes the conduct of a framework in the applied model. The enormous frameworks are deteriorated into subordinate frameworks to give comparative arrangement of administrations. The starting format procedure of seeing these sub-frameworks and working up a structure for sub-frameworks control and collaboration is called engineering plan. This Architecture as shown in figure consist seven steps explained as below:

A. Objective Identification

The initial phase in building up a venture is to comprehend the target which includes a comprehension of the aim and fundamentals of a framework. This cognizance is utilized as an issue depiction and a preliminary framework to achieve the desires. The goal of our venture is neither to manufacture a framework that makes billions nor to squander billions as well. However, the goal is to build up a framework that finds the course of progress of stock value files in view of the co-relations between stock costs and help the financial specialists in money markets in taking a choice whether to purchase/offer/hold a stock by giving the outcomes regarding representations.

B. Data Storing

Once the comprehension of the goal is finished, the following stage is to gather the information. Information accumulation includes the comprehension of introductory perceptions of the information to distinguish the valuable subsets from speculations of the concealed data. Here we utilize R content to gather the information from Google back.

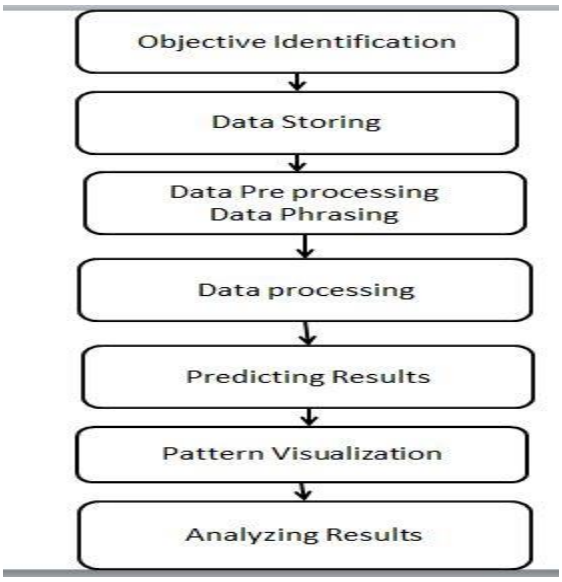

Figure 5 Flow control for the process of stock price prediction using R

C. Data Pre-processing: Data Phrasing

The information pre-handling stage includes every one of the exercises to set up the last dataset from the preliminary crude data. The information readiness undertakings can be played out a few times as there is no particular request. These errands incorporate the determination of a record, table, property and cleaning of information for displaying instruments. In our technique, the info information will be changed over into a joined esteem vector list or separated esteem vector list.

D. Data Processing

To process the information we utilize ARIMA (p, d, q) display. In specialized investigation financial specialists utilize the auto backward and moving normal models to estimate the stock patterns. Real advances required here are distinguishing proof, parameter estimation and gauging. These means are rehashed until the point that a proper model is distinguished for forecast. R gives auto.arima () technique to gauge the time arrangement information as indicated by ARIMA (p,d,q).

Order of ARIMA: The request of an ARIMA demonstrate is by and large spoken to as ARIMA (p,d,q), where-

- p = request of the autoregressive part
- d = level of first differencing included
- q = request of the moving normal part

Here if d=0, at that point the model moves toward becoming ARMA which is straight stationary model.

E. Predicting Results

The way toward making expectations without bounds by depending upon the various times information is known as gauging. Different forecast systems are utilized by the stock experts to assess the future stock patterns esteem. Forecast likewise offers a huge standard for associations that have a long haul view of activities. We utilize 'gauge' bundle for anticipating the future stock patterns in light of the examination of past patterns. This 'gauge' bundle gives various determining capacities for showing the time arrangement forecasts alongside exponential smoothing and space models.

F. Pattern Visualizations

Information perception is a graphical portrayal of the numerical information. In our philosophy, in the wake of gauging the share trading system patterns we envision the outcomes for here and now venture help with terms of line outlines, candles diagrams, bar graphs, and histograms. Here x-hub demonstrates the day and age as far as year/months/days and y-pivot demonstrates the stock value esteems.

G. Analyzing Results

Once in the wake of plotting the outcomes as far as perceptions we can discover the connections to get the transient forecasts. In the following area we give a portion of the screen shots by which the financial specialist can break down and anticipate the future stock patterns of a specific organization at a particular day and age. So the financial specialists in the stock exchange can utilize this as help to offer/purchase/hold an offer.

A prediction model was developed for estimating the share trading system patterns in view of the specialized investigation utilizing chronicled time arrangement securities exchange

information and information mining procedures. The trial comes about got exhibited the capability of ARIMA model to foresee the stock value lists on here and now premise. This could manage the financial specialists in the share trading system to settle on productive speculation choices whether to purchase/offer/hold an offer. With the outcomes got ARIMA model can contend sensibly well with rising anticipating systems in here and now expectation.[13]

2.9 Cloud Based Stock Market Prediction Using Genetic Algorithms

Stock exchanges are turbulent in nature and it based on fractal market Hypothesis. Fractal market speculation takes after five guidelines [9]:

- FMH1. The market is comprised of numerous people with countless venture skylines.
- FMH2. Data differently affects distinctive venture skylines.
- FMH3. The strength of the market is to a great extent a matter of liquidity (adjusting of free market activity). Liquidity is accessible when the market is made out of numerous financial specialists with a wide range of venture skylines.
- FMH4. Costs mirror a mix of here and now specialized exchanging and long haul central valuation.
- FMH5. In the event that a security has no attach to the monetary cycle and after that there will be no long haul drift. Exchanging, liquidity, and here and now data will rule.

Fractal Market theory calls attention to some additionally fascinating raw numbers e.g. it is more worries about interest of the share trading system players and financial specialists, confused conduct of money markets and so forth.[14]

2.10 Stock Price Prediction Using Outlier Data Mining Algorithms

The exchanging volume may follow some arbitrary distribution since in the effective market theory the market continuously follows an arbitrary walk. In this way, it is expected that if the volume isn't so irregular any longer that there are a few peculiarities in the dissemination. By then the market isn't effective and this implies the stock cost isn't an irregular walk any longer so a long term foreseeing technique is conceivable. Here, the strategy is executed to check in the case of utilizing the distinguished oddities from authentic financial time

arrangement information can anticipate stock pattern successfully or not.

The following steps are included:

- A Trading Volume distribution is proposed using anomalies to predict stock price's higher trend.
- Utilize tick-by-tick information rather than time arrangement information on stock cost in a novel exception mining calculation.
- Select 200 stocks haphazardly in our test. The outcome demonstrates that utilizing peculiarities can foresee the upward pattern of stock costs adequately.

Here is the case where k mean algorithm is used for clustering which is considered to design the clusters. This algorithm proved to be more effective as compare to the other algorithms because in the case of other algorithms there must be normal data with the anomaly data exist. It means that prediction of stock price is not possible if there is no normal data exist before the anomaly data. But in the case of this algorithm there is no such criteria exist. Anomaly data should exist, no matter where it is. In this way the algorithm is more effective.[15]

2.11 Dealing with Mispricing in Stock Price Prediction

There is always a link between the stock predicting price and actual price. Stock price what we predict is actually quite deviate from the actual price. This deviation between the predicted price and actual price can be considered as the threshold. Traditionally these functions could be considered as the function of cost. But now these functions may be considered as changed with various factors. These certain factor that can affect the value of threshold are liquidity, unpredictability and short-deals limitations utilizing information on single stock fates from the National Stock Exchange (NSE) of India.

Deviations in costs between securities furthermore, its fates contract takes after a multi-administration model. Specially, mispricing is accepted to fall under one of three administrations:

(1) an administration mispricing is beneath a lower edge;
(2) administration where mispricing is over an upper limit;

(3) a halfway administration where mispricing is between these limits.

This method basically contains three processes. To start with, first look at the effect of liquidity on misestimating limits. Oehmke (2009) proposes that liquidity is a key determinant of the speed with which capital moves to arbitrage openings. He demonstrates that for a given level of com- request, arbitrageurs would exchange less forcefully when liquidity is low. Further, the level of rivalry itself may diminish with expanding illiquidity. On the off chance that arbitrageurs are reluctant to exchange illiquid resources, the degree of mispricing should increment with a lessening in liquidity.

Second, in the wake of controlling for these liquidity effects, look at how the mispricing limits differ with unpredictability. It is intriguing to break down this issue, as hypothetical direction is blended. Higher instability builds execution hazard and prompts more noteworthy edge calls. Henceforth, arbitrageurs may be less ready to execute these exchanges the substance of expanding vulnerability. This thus could prompt more extensive limits. Then again, Brennan and Schwartz (1990) hypothesize that arbitrageurs have a profitable early leave choice. Regardless of the possibility that current mispricing is lower than exchange costs, dealers may start a position planning to sell it early furthermore, book profits from interval value developments.[16] They likewise find that the estimation of this choice is represented by the unpredictability of evaluating mistakes. Instinctively, instability in mispricing would be corresponded with unpredictability of the stock; Hence, the early leave choice is more significant for a stock with high instability than for a stock with low unpredictability.

At last, we explore the effect of short-deal requirements. These market contacts have a direct influence on our prior forecasts as the early leave choice may be applicable as it were at the point when fates are overrated. When they are underpriced, a dealer won't not have the capacity to put the underlying exchanges - long fates, short stock - given the limitations on short deals. Subsequently, the effect of unpredictability on mispricing by means of the early leave alternative may be exceptionally uneven. Further, short-deal requirements have impacts on mispricing that are more principal. For case, the lower limits would be higher in extent contrasted with the upper limits. In any case, if a greater part of arbitrage exchanges are executed by foundations that can offer stocks.[17]

The liquidity of the prospects showcase largely affects the span of the mispricing window contrasted with that of the spot showcase. Subsequent to controlling for liquidity effects, the span of the mispricing window is found to increment with an expansion in instability. Higher instability is related with the lower bound winding up more negative and the upper positive relationship amongst unpredictability and the span of mispricing. These findings propose that higher instability is related with more prominent underpricing in prospects contracts. Swinging to the more central effects of these showcase gratings, the mean lower bound of mispricing is higher in size than the mean upper bound. This recommends short-deal requirements likewise don't have a symmetric effect on mispricing limits.

An extensive examination of the connection between liquidity, unpredictability and way reliance of mispricing in single stock prospects. The information is used from the National Stock Exchange (NSE) of India, which is universally positioned second as far as exchanges stock fates. The high liquidity in these business sectors grants us to look at a number of intriguing theories. Initially, the measure of the mispricing window increments with a diminishing in liquidity. Second, the liquidity of the prospects advertises has a bigger effect contrasted with that of the spot showcase. Third, even subsequent to controlling for these liquidity effects, the measure of the mispricing window increments with an expansion in instability. These worries over edge calls and execution deficits command the early exit options.

Higher instability is related with the lower bound winding up more negative and the upper bound getting to be more negative. In any case, the previous commands the last prompting prior finding that an expansion in unpredictability is related with an increment in mispricing groups. It was guessed that this outcome is driven by short-deal imperatives. At the point when the fates are underpriced, an arbitrageur would react by starting a long position in the prospects contract and a short position in the money advertise. Be that as it may, if short offering is compelled, the broker won't not be ready to at the same time execute these exchanges. While he could start an exposed long position, the worries about edge calls would be high. Subsequently, higher unpredictability pushes the lower bound considerably additionally down.[18].

CHAPTER -3
SCOPE OF THE STUDY

Data forecasting is really convenient topic of research from last few decades and may remain active topic in upcoming years also. Stock price prediction also used data forecasting is basically is one of favorite topic among researchers. A lot of research is already done in this field like Stock price prediction can be using neural, fuzzy, machine learning, R programming and so on. Here we want to predict future stock price using the some predictive services. This will provide help to get more accurate results for predicting stock price prediction. In future we can analyze this stock market historical data in some other way to find more accurate results. We can deal with the not only finding of future stock price prediction but also tried to reduce mismatch value i.e. difference between actual price and predicted price. Threshold value can be reduced to move toward more accurate value.

CHAPTER -4
OBJECTIVES OF THE STUDY

Objectives of the study are defined as follow:

- The main objective of this study is to predict the future stock price by analyzing the past historical data that we were going to collect from the National Stock Exchange.
- Predicting the Stock market cost in such a way that it will provide most accurate results.
- Stock market price forecasting should be done in such a way that predicted price should minimize the threshold value (difference between actual value and predicted value also known as mispricing) and close enough to the actual value.
- Process of analyzing the historical data should be simple and easy to understand. For this feature identification can done intelligently to provide most accurate results.
- To increase the efficiency of the data analysis technique by using some cloud based tools.
- To analyze the performance and comparing proposed algorithm with the existing algorithms in terms of predicted price accuracy, close price predicted and accurate close price etc.

CHAPTER-5

RESEARCH METHODOLOGY

This idea of research we got from the "Stock Market Prediction and Analysis Using Naïve Bayes"[10] where author want to predict the data using the Naïve Bayes Theorem. He followed an approach of predicting stock price by mainly including two steps i.e. Automation and prediction. Data is accessed from the yahoo finance site with having the real time data access. The main two processes are followed. In first process which is automation an automated procedure for buy and sells the shares is followed. At the point when user set one fix cost to deal the offers, after that framework coordinate cost of shares& consequently perform deal operation. Like that buy operation is likewise performed. In that set one fix cost to buy the offers, after that framework coordinate cost of shares& consequently perform buy operation.

The conceivable market forecast objective can be the future stock cost or the instability of the costs or market incline. In the forecast there are two sorts like Dummy and Real time expectation. In sham forecast we characterize a few guidelines and foresee the future cost of offers. In the ongoing forecast utilize web and see current cost of offers.

This Method is trying to predict the future outcome in best possible way with most accurate results.

This Method we can follow using some cloud based tools. We can use cloud based tools to analyze the stock future price. This idea come from "Cloud based Financial Market Prediction through Genetic Algorithms"[14] where author tried to differentiate the different technique we used to predict the future stock price. According to this paper, there are so many technique that we used to predict the cost of future stock and each time we tried our best to get more accurate results but there are some limitations that every technique have like fuzzy logic is able to provide some of the probabilistic results which are not enough in the case of the predicting future outcome. Neural network we used most of the time to predict the future outcome also have some limitations. Similarly data mining technique etc.

Limitations of different techniques used for prediction:

i. Neural Network comes about are insecure. The neural system capacities are Black Box work. The principles of operations are totally obscure.
ii. Back proliferation systems can be set aside long opportunity to prepare the substantial measure of information.
iii. Data mining methods are utilized to file the most noteworthy certainty up to 70% as it were.
iv. Unlike regression demonstrate, ARIMA show isn't support the stationary time arrangement information.

These limitations can be override by implementing hybrid method. By using the method to implement its best functionality only likes to overcome the limitation of neural network.
We can integrate some heuristic techniques with the neural network to work in stable environment and to understand its working[19]. We can use the some cloud based tools to demonstrate the stock price prediction which are very much efficient to give accurate results. Stock price prediction using cloud based tools is expected to be more helpful in case of analyzing huge data in little time interval by using such efficient algorithms will not only provide accurate results to predict stock market price but also help to understand the complete process.

5.1 Stock Price Prediction Using Cloud Based Tools

As from last some years, cloud is really an interesting area of use. Cloud is used in basically in every field now. Some automated systems are also designed in cloud even for stock market prediction. The major benefit of cloud based tools is performance increases. By using some efficient algorithms we can implement stock price prediction in cloud also which mainly help to analyze the data easily.

We are going to implement stock market prediction using SAP HANA predictive services. This predictive service we can access using SAP HANA on demand interface. We can elaborate this process by using the SAP HANA studio where we can access data and analysis and prediction based on historical data can be done on the basis of the predictive services.

SAP HANA is basically is a column oriented , in memory relational database management

system which is managed and developed by the SAP SE(System, Application & products in data processing) where SE stands for Software Enterprise.

HANA is basically stands for high-performance analytic appliance, is an application based on in memory database where data resides in memory i.e. in RAM instead of hard disk drive. That's why this may have fastest data accessing and processing capabilities.

We can use this as on premise or in cloud based. Cloud based tool can have some services that we are not able to use in studio. For Stock price prediction we are going to use predictive analysis service which provides functionality for predicting stock market price by processing the historical data. Historical data that we are going to analyze will be taken from National Stock Exchange. We can analyze the data by using the workbench or can be using the SAP HANA. if we are going to use HANA studio, we have to first load the data into the SAP HANA studio and then we have to connect SAP HANA with the predictive service workbench where we can analyze the data.

CHAPTER-6
EXPECTED OUTCOMES OF THIS STUDY

Some outcomes that we expect from our System is described as follow:

- This is expected the this system will help to analyze the past historical data in the better way and provide more accurate results as compare to the previously used techniques for predicting future stock price.
- This is also expected that this technique will help to reduce threshold value i.e. Mispricing value and predicted result may be closed enough to its actual value.
- The technique used provides efficient result as compare to previous techniques.

CHAPTER-7
SUMMARY AND CONCLUSION

As mentioned in the previous chapter, Stock price prediction is the active area of research from few decades. There are multiple methods used in the stock price prediction for data analysis and forecasting. Hybrid methods can be used in which we can combine two or more techniques to predict results in more accurate way like neural network can be combined with the fuzzy logic for technical analysis where feature identification can be done using the decision tree. This method can be known as hybrid decision tree-neuro-fuzzy system. Data that we used for the analysis can also be collected from two or more sites like in stock we can collect data from National Stock Exchange along with Bombay Stock Exchange to compare the price variation between these two market. Stock price can be predicted more accurately if we include sentimental data analysis along with historical data that we can collect from the various social sites like twitter, Facebook, Google Trend etc. Multiple techniques arises to get more accurate results like Text Mining, Genetic algorithms, Naïve Bayes Method, R programming, Outlier based data Mining Algorithms, Cloud Based Enterprise Tools and so on. All these methods are used just to predict the future outcomes more accurately so that the threshold value (difference between actual value and predicted value also known as Mispricing) can be reduced.

CHAPTER-8
REFERENCES

[1] G. Iuhasz, M. Tirea, and V. Negru, "Neural network predictions of stock price fluctuations," *Proc. - 14th Int. Symp. Symb. Numer. Algorithms Sci. Comput. SYNASC 2012*, pp. 505–512, 2012.

[2] L. I. Hong-dong, Z. Jing, X. Lin, L. I. Hai-ping, and F. Yi, "Application of DS Evidence Theory in Combined Price Forecasting," no. April, pp. 1025–1029, 2008.

[3] A. A. Bhat and S. S. Kamath, "Automated stock price prediction and trading framework for Nifty intraday trading," *2013 4th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2013*, 2013.

[4] M. Setnes and J. H. Van Drempt, "Fuzzy modeling in stock-market analysis," pp. 250–258.

[5] Hiemstra, "A stock market forecasting support system based on fuzzy logic," *1994 Proc. Twenty-Seventh Hawaii Int. Conf. Syst. Sci.*, vol. 3, pp. 281–287, 1994.

[6] B. B. Nair, M. Minuvarthini, B. Sujithra, and V. P. Mohandas, "Stock market prediction using a hybrid neuro-fuzzy system," pp. 243–247, 2010.

[7] B. B. Nair, N. M. Dharini, and V. P. Mohandas, "A stock market trend prediction system using a hybrid decision tree-neuro-fuzzy system," *Proc. - 2nd Int. Conf. Adv. Recent Technol. Commun. Comput. ARTCom 2010*, pp. 381–385, 2010.

[8] Y. Lin, H. Guo, and J. Hu, "An SVM-based approach for stock market trend prediction," *Proc. Int. Jt. Conf. Neural Networks*, 2013.

[9] N. Milosevic, "Equity forecast: Predicting long term stock price movement using machine learning," *J. Econ. Libr.*, vol. 3, no. 2, p. 8, 2016.

[10] 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.

[11] S. Yue and X. Uc, "Stock Price Forecasting Using Information from Yahoo Finance and Google Trend."

[12] K. Kelley, K. Lai, and P.-J. Wu, "Using R for data analysis: A best practice for research," *Best Pract. Quant. methods*, pp. 535–572, 2008.

[13] M. C. Angadi, "International Journal of Advanced Research in Computer Science RESEARCH PAPER Available Online at www.ijarcs.info Time Series Data Analysis for Stock Market Prediction using Data Mining Techniques with R," vol. 6, no. 6, pp. 104–108, 2015.

[14] N. Soni and T. Kumar, "Cloud based Financial Market Prediction through Genetic Algorithms: A Review," *Int. J. Comput. Appl.*, vol. 123, no. 8, pp. 18–20, 2015.

[15] L. Zhao and L. Wang, "Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm," *Proc. - 2015 IEEE 5th Int. Conf. Big Data Cloud Comput. BDCloud 2015*, pp. 93–98, 2015.

[16] K. L. V Zrun *et al.*, "No Title," vol. 3.

[17] H. Berkman and V. R. Eleswarapu, "Short-term traders and liquidity : a test using Bombay Stock Exchange data," vol. 47, pp. 339–355, 1998.

[18] R. L. Shankar, "Mispricing in single stock futures : Empirical examination of Indian markets Mispricing in single stock futures : Empirical examination of Indian markets," no. June, pp. 1–36, 2015.

[19] I. Metaheuristics, S. P. Prediction, M. Ozc, M. Ozc, and I. Metaheuristics, "PT US CR," 2015.