



STOCK PRICE PREDICTION PROGRAM

A PROJECT REPORT

Submitted by

HARDEY PANDYA (166490316073)

SARVIK VAGHASIYA (166490316120)

In partial fulfilment for the award of the degree

of

DIPLOMA ENGINEERING

in

INFORMATION TECHNOLOGY

Guided by

HARDIK P. JAGAD



Sir Bhavsinhji Polytechnic Institute Bhavnagar

Gujarat Technological University, Ahmedabad

October, 2018

SIR BHAVSINHJI POLYTECHNIC INSTITUTE
BHAVNAGAR



Certificate

This is to certify that Mr. Pandya Hardey Nileshbhai from **Sir Bhavsinhji Polytechnic Institute, Bhavnagar** College having Enrollment No: 166490316073 has completed **Project Report** having title Stock Price Prediction Program in a group consisting of **2** persons under the guidance of the faculty guide Mr. Hardik Jagad.

Institute Guide-UDP

Head of Department

SIR BHAVSINHJI POLYTECHNIC INSTITUTE
BHAVNAGAR



Certificate

This is to certify that Mr./Ms. Vaghasiya Sarvik Ashokbhai from **Sir Bhavsinhji Polytechnic Institute, Bhavnagar** College having Enrollment No: 166490316120 has completed **Project Report** having title Stock Price Prediction Program, individually/ in a group consisting of **2** persons under the guidance of the faculty guide Mr. Hardik Jagad

Institute Guide-UDP

Head of Department

STUDENT INFORMATION SHEET

Name of Student	Pandya Hardey Nileshbhai	
	Surname	Name Father's Name
Enrollment Number	166490316073	
Contact Numbers	Mob: 9265595781	Land Line: -
Email ID	pandyahardey@gmail.com	
College Name	Sir Bhavsinhji Polytechnic Institute, Bhavnagar	College Code: 649
Branch	Information Technology	Semester : V
Student Team	Name	Enrollment Number
	Pandya Hardey Nileshbhai	166490316073
	Vaghasiya Sarvik Ashokbhai	166490316120
Student Signature		

STUDENT INFORMATION SHEET

Name of Student	Vaghasiya Sarvik Ashokbhai	
	Surname	Name Father's Name
Enrollment Number	166490316120	
Contact Numbers	Mob: 7622089518	Land Line: -
Email ID	sarvikvaghasiya611@gmail.com	
College Name	Sir Bhavsinhji Polytechnic Institute, Bhavnagar	College Code: 649
Branch	Information Technology	Semester : V
Student Team	Name	Enrollment Number
	Pandya Hardey Nileshbhai	166490316073
	Vaghasiya Sarvik Ashokbhai	166490316120
Student Signature		

Acknowledgement

It is our pleasure to be indebted to various people, who directly or indirectly contributed in the development of this work and who influenced our thinking, behaviour, and acts during the course of study.

We express our sincere gratitude to prof. **D.A. Dave**, worthy Principal of the college for providing us an opportunity to implement and test our project at the college.

We are thankful to **Mr. Hardik Jagad, our Project Guide**, for his support, cooperation, and motivation provided to us during the design and analysis of this.

We also extend our sincere appreciation to **Mr. G.M. Pandey, Head of Department – Information Technology** who provided his valuable suggestions and precious time in accomplishing our project report.

Lastly, We would like to thank our friends and colleagues with whom we shared our day-to-day experience and received lots of suggestions that improved our quality of work.

Sincerely,

Hardey Pandya.
Sarvik Vaghasiya.

Index

Abstract.....	i
List of Tables.....	ii
List of figures.....	ii
List of Symbols, Abbreviations.....	iii

Index	Page No
I. Introduction.....	01
i. Project Introduction / Profile + Company Profile.....	02
ii. Purpose.....	05
iii. Scope.....	07
II. System Requirement Analysis.....	08
i. Current System Study.....	09
ii. Weakness of Current System.....	10
iii. Problem Identification / Definition.....	10
iv. Requirement of New System.....	11
v. Feasibility Study.....	11
a. Technical.....	11
b. Economical.....	12
c. Operational.....	12
d. Schedule (Time Line Chart).....	13
vi. Development model (Software Process Model).....	14
vii. Requirement Validation.....	15
viii. Minimum Hardware and Software Requirements.....	16
ix. System Architecture.....	17
x. Data Flow Diagram.....	18
xi. Use –case Diagram.....	20
xii. Activity Diagram.....	23
III. System Design	
i. Database Design.....	26
a. Entity-Relationship Diagram.....	26
b. Data Dictionary.....	27
ii. GUI Design (Self Created GUI Screen shot).....	30
Reference.....	31

Abstract

Stock Market Prediction is the act of trying to determine the future value of a company stock. The successful prediction of stock's future price will maximize the investor's gains. Investors regularly check the prices, analyse them and according to that they invest in particular share of a company. But it is full of uncertainty. There are no specific rules for estimating stock prices and therefore a lot of approaches has been used with their own pros and cons. We consider the traditional historical time series analysis, Technical Analysis and Fundamental Analysis.

We analysed some research papers and performed some experiments with Data Mining tool WEKA and found that Machine learning model which use Combination of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) are most efficient. Financial News, Business Articles are also taken into consideration to perform Textual Analysis. Here we take data of 10 companies registered in NSE India, which are always under the eye of traders. The project is aimed to aid regular shareholders to invest in right area. We consider the traditional historical time series analysis, Technical Analysis and Fundamental Analysis.

Keywords: Machine Learning, Technical Analysis, Time Series Analysis, Artificial Neural Networks, Support Vector Machine, WEKA, Textual Analysis.

:- List of Figures :-

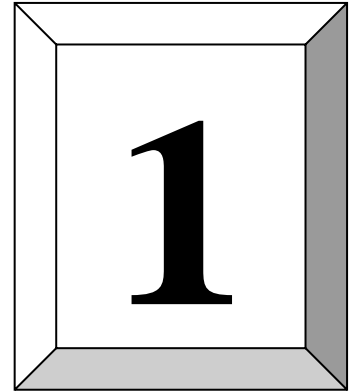
Figure	Page
1. Efficient Market Hypothesis.....	2
2. Time Line Chart.....	12
3. Software Process Model (Incremental Model).....	14
4. System Architecture (Repository Architecture).....	16
5. DFD Symbols.....	17
6. Context Level Diagram.....	18
7. Level-1 DFD.....	18
8. Use Case Diagram.....	19
9. Activity Diagram.....	23
10. ER Diagram.....	25
11. System GUI.....	29

:- List of Tables :-

Table	Page
1. Comparative Study of Prediction Techniques.....	4
2. Data Dictionary: Company Data.....	12
3. Data Dictionary: Predicted Data.....	14
4. Data Dictionary: Technical Indicators.....	16

:- List of Abbreviations :-

1. ANN – Artificial Neural Network
2. RNN – Recurrent Neural Network
3. EMH – Efficient Market Hypothesis
4. SVM – Support Vector Machine
5. NSE – National Stock Exchange (India)
6. BSE – Bombay Stock Exchange
7. CNN – Convolutional Neural Network
8. DFD – Data Flow Diagram
9. ERD – Entity-Relationship Diagram
10. RSI – Relative Strength Index
11. CCI – Commodity Channel Index
12. SMA – Simple Moving Average
13. MD – Mean Deviation



Chapter # 1: Introduction



1.1 Project Introduction

1.2 Purpose

1.3 Scope

1. Introduction

1.1 Project Introduction:

Stock Market prediction and analysis is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. Stock market is the important part of economy of the country and plays a vital role in growth of the industry and commerce of the country. Both investors and industry are involved in stock market and wants to know whether some stock will rise or fall over certain period of time.

The main motivation behind choosing the project:

- There is large amount of relevant financial data available on the internet which is increasing day by day.
- Large number of Computer Science disciplines including software engineering, databases, distributed systems and machine learning have increased possibility to apply skills.
- The opportunity to expand our knowledge in finance and investing, as we had only little prior exposure to these fields.
- It possesses many theoretical and experimental challenges.

Stock market is very difficult to understand. It is considered too much uncertain to predict due to huge fluctuations in the market. Stock market prediction task is interesting and it divides researchers into two schools: one who believes market can be modelled into some algorithms and subsequently can be predicted and those who believe in EMH.

Most of the trading in the Indian stock market takes place on its two stock exchanges: the BSE and the NSE. The BSE has been in existence since 1875. The NSE, on the other hand, was founded in 1992 and started trading in 1994. However, both exchanges follow the same trading mechanism, trading hours, settlement process, etc. At the last count, the BSE had about 4,700 listed firms, whereas the rival NSE had about 1,200.^[1]

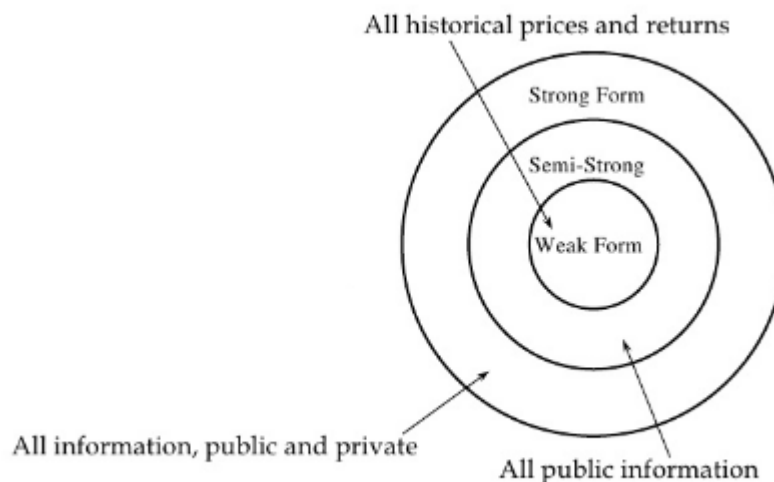
Here is the basic terminology related to prediction of stock market:

Efficient Market Hypothesis (EMH) :-

Weak-form Efficient Market Hypothesis: - The weak form of the hypothesis says that no one can profit from the stock market by looking at trends and patterns within the price of a product itself. It is important to note that this does not rule out profiting from predictions of the price of a product based on data external to the price.

Semi-Strong Efficient Market Hypothesis: - The semi-strong form rules out all methods of prediction, except for insider trading. This means that if we are only to use public domain information in our prediction attempt, the semi-strong form says that we will be unsuccessful.

Strong form Efficient Market Hypothesis: - The strong form says that no one can profit from predicting the market, not even insider traders.



(Efficient Market Hypothesis)

In order to clarify the goal of the project, following are the dominant schools of thought on investing must first be introduced.

Time Series analysis: A time series is a sequence of numerical data points in successive order. In investing, a time series tracks the movement of the chosen data points, such as a security's price, over a specified period of time with data points recorded at regular

intervals. There is no minimum or maximum amount of time that must be included, allowing the data to be gathered in a way that provides the information being sought by the investor or analyst examining the activity.

Fundamental analysis: - This approach is to analyze fundamental attributes in order to identify promising companies. This includes characteristics such as financial results, company's assets, liabilities, and stock and growth forecasts. It's very important to understand that this type of analysis is not static; newly released financial information, corporate announcements and other news can influence the fundamental outlook of a company. Fundamental analysis requires expertise in a particular sector and is often conducted by professional analysts. Their recommended investments are regularly published and updated.

Technical analysis :- In contrast to fundamental analysis, technical analysis does not try to gain deep insight into a company's business. It assumes the available public information does not offer a competitive trading advantage. Instead, it focuses on studying a company's historical share price and on identifying patterns in the chart. The intention is to recognize trends in advance and to capitalize on them.

Within the technical analysis community there exist several schools with different techniques, but they all have in common that they use price and volume history. A basic thought is that it takes time before the market reacts upon new information and that pattern often occurs in price behavior which makes forecasting possible.

There are several factors that explain why technical analysis works:

1. Most speculators on the market act upon fundamental analysis, so that kind of facts influence stock prices strongly. But all operators do not get this information at the same time. When there is positive news of a company, those acting immediately can buy shares for a lower price than those getting the news later.
2. Large investors such as mutual funds and banks are often not placing their whole block orders at the same time when they are buying larger quantities of securities because this would risk triggering an unnecessary high price advance. Instead, the orders are spread over a period that can last several weeks. The resulting increased purchase pressure may result in a steady advancing trend under the period the purchases continue.

3. It is more psychological stressing to go against the trend than to follow it. People are herding animals and like to do as others are doing. This is why a rising stock price is a signal in itself that the price will advance even more. Of course, one has to be careful with stocks that have been rocketing, because they will often recoil.

Comparative Study of Prediction Techniques :-

Criteria	Technical Analysis	Fundamental Analysis	Traditional Time Series Analysis
Data Used	Price, volume, highest, lowest prices.	Growth, dividend payment, sales level, interest rates, tax rates etc.	Historical data
Learning methods	Extraction of trading rules from charts	Simple trading rules extraction	Regression analysis on attributes is used
Type of Tools	Charts are used	Trading rules	RNN, ANN, Linear Regression, etc.
Implementation	Daily basis prediction	Long –term basis prediction	Long –term basis prediction

(Comparative Study of Prediction Techniques)

1.2 Purpose

The purpose of this project are as follows:

- To identify factors affecting share market.
- To generate the pattern from large datasets of NSE stock market for prediction.
- To predict an approximate value of share price.
- Perform and Experiment different algorithms if the share price is not correctly predicted by the current model.
- Perform traditional methods of Technical Analysis and Time Series Analysis regularly.
- Regularly check the accuracy of the prediction.
- Predict and inform the user about the status of the company as accurately as possible through Textual Analysis.

Company Selection :

This program will predict the stocks of following companies registered in NSE.

This companies are more or less popular among stock market dealers community.

- HDFC Bank
- Reliance
- Tata Consultancy Services Limited(TCS)
- Infosys Ltd
- ICICI Bank
- Maruti Suzuki
- Bank of Baroda
- State Bank of India(SBI)
- Aditya Birla Capital
- Ambuja Cement
- Hyundai Motors
- Tata Motors

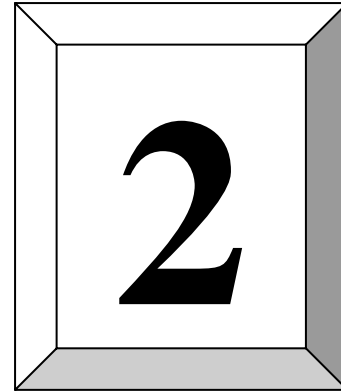
1.3 Scope

The project will be useful for investors to invest in stock market based on the various factors. The project target is to create a program that analyses important parameters which affect share prices and implement these parameters in Machine Learning Algorithm to determine the value that particular share will have in near future as accurately as possible. These predicted and analysed data can be observed by anyone and can know the financial status of companies and their comparisons.

Predicted and analysed data is also useful to Companies themselves. Company and industry can use it to stretch their limitations and enhance their stock value. It can be very useful to even researchers, stock brokers, market makers, government and general people.

The main feature of this project is to generate an approximate forecasting output and create a general idea of future values based on the previous data by generating a pattern. The scope of this project does not exceed than a generalized suggestion tool.

The objective of the system is to give an approximate idea of where the stock market might be headed. It does not give a long-term forecasting of a stock value. There are way too many reasons to acknowledge for the long-term output of a current stock.



Chapter # 2: System Requirement Analysis

- 2.1 Current System Study**
- 2.2 Weakness of Current System**
- 2.3 Problem Identification / Definition**
- 2.4 Requirement of New System**
- 2.5 Feasibility Study**
 - 2.5.1 Technical**
 - 2.5.2 Economical**
 - 2.5.3 Operational**
 - 2.5.4 Schedule (Time Line Chart)**
- 2.6 Development model (Software Process Model)**
- 2.7 Requirement Validation**
- 2.8 Tools and Technology / Minimum Hardware and Software Requirements**
- 2.9 System Architecture**
- 2.10 Data Flow Diagram**
- 2.11 Use –case Diagram**
- 2.12 Activity Diagram**

2. System Requirement Analysis

2.1 Current System Study :

There are various kinds of systems available for stock market forecasting. But each system differs in their approaches.

The approaches used to forecast future directions of share market prices are historically splitted into two main categories: those that rely on technical analysis, and those that rely on fundamental analysis.

The machine learning approach to the latter problem has been declined in several forms, especially during recent years. As an example, good results have been obtained using linear classifiers as the logistic regression one, which has been used to predict the Indian Stock market. More complicated techniques such as Support Vector Machine (SVM), was the best choice for prediction before the rise of neural networks.

Many other systems currently use different set of algorithms including Hidden Markov Models, Bayesian Networks, Multi-Layered Perceptrons (MLP), etc. But in general, most of the time efficient prediction is observed to be done by the use of SVM and several ANNs, rather than other models. It is Important to note the fact that for every company, it may be possible that any different model may be more suitable. It's not necessary that one model used for one company is suitable for other company. (Check References: In every research paper cited here, one may easily observe this issue)

Currently the systems use RNN, CNN, SVM to predict stock prices, with the advent in computational power of the computers. Computers with mediocre computational powers cannot perform heavy algorithms which involve complex neural networks.

The main disadvantage of all current systems is that once you define the model of prediction, you cannot change it. It may be result into heavy loss in the accuracy of the system. So, every current system tends not to follow EMH.

2.2 Weakness of Current System :

Current systems for stock price prediction are generally proprietary software and generally it costs very high to its consumers. At the same time there are very handful amount of systems available which are cost effective and efficient to perform stock market analysis. Moreover, many current systems do not involve textual analysis. Sometimes stock market is more about the emotions of business rather than just doing analysis of historical data. In such scenarios, textual analysis is helpful.

We also introduce some indices which give information of the weight of time series analysis, fundamental analysis and technical analysis on the final future value of the share.

2.3 Project Definition / Problem Identification :

Forecasting of stock market is gaining more attention as the profitability of investors in the stock market mainly depends on the predictability. If the direction of the market is successfully predicted the investors can yield enough profits out of market using prediction.

Stock price prediction is rather a hazardous operation. A good analyst is therefore not the one who is always right, but someone who is at average, someone who has higher efficiency than his colleagues.

Sometimes overall impression of the company is more helpful in predicting its stock price, therefore textual analysis of relevant newspaper articles, financial news and impressions have also been into consideration.

In the last few years, it has become clear that ANN have become part of this class of analysts. ANNs are programs that are based on the geometry of the human brain. We analysed some research papers and performed some experiments with Data Mining tool WEKA and found that Machine learning model which use Combination of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) are most efficient and they are also fast and it can also be able to run in mediocre hardware specifications.

Investing in a good stock but at a bad time can have disastrous result, while investing in a stock at the right time can bear profits. Financial investors of today are facing this problem of trading as they do not properly understand as to which stocks to buy or which stocks to sell in order to get optimal result. So, the purposed project will reduce the problem with suitable accuracy faced in such real time scenarios.

2.4 Requirement of New System:

It is difficult to find existing systems that are both performance efficient and have high accuracy. Existing systems are also costly. By proposing new system which is open source and give sufficient amount of accuracy to its users will likely to be helpful to take them their decisions.

Day by day, on the internet, availability of financial data, use of social networking, Online opinions and comments, E-Newspaper, Financial articles are increasing at a very high rate. Any new system can make full use of it. By using sentimental analysis and textual analysis it is possible to give probable overall status of the company and sometimes stock market ups and downs are more about status of company rather than just mining the historical data.

2.5 Feasibility Study:

2.5.1 Technical Feasibility:

Simply put, stock market cannot be accurately predicted. The future, like any complex problem, has far too many variables to be predicted. When there are more buyers than sellers, the price increases. When there are more sellers than buyers, the price decreases. It has more to do with emotion than logic. Because emotion is unpredictable, stock market movements will be unpredictable. It's futile to try to predict where markets are going. They are designed to be unpredictable. To deal with such case we have introduced textual analysis which involve analysis of business articles, news, etc. to give the user certain idea of the status of the company.

2.5.2 Economic Feasibility:

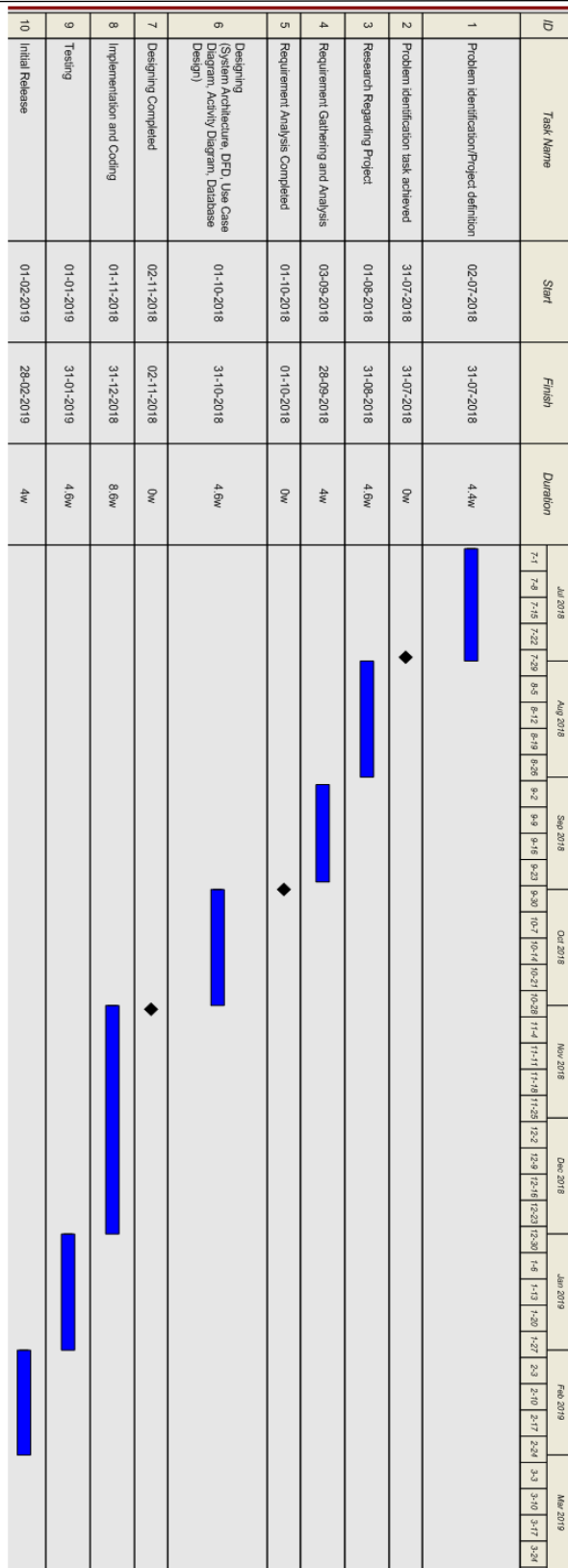
The proposed system is extremely economic feasible. As mentioned above that day by day, availability of financial data is increasing on the internet. Internet is also available with negligible cost today. The system uses internet to fetch the data and process according to certain algorithms which can run on normal hardware specifications. Hardware requirement is also not very fancy.

2.5.3 Operational Feasibility:

The proposed system will not always produce accurate results since it does not account for human behaviours. Factors like change in company's leadership, internal matters, strikes, protests, natural disasters, terrorist attack, change in authority, political affairs, etc. cannot be considered for relating it to change in Stock market by machine at any circumstances.

The objective of the system is to give an approximate idea of where the stock market might be headed. It does not give a long-term forecasting of a stock value. There are way too many reasons to acknowledge for the long-term output of a current stock. Many things and parameters may affect it on the way due to which long term forecasting is just not feasible.

2.5.4 Time-Line Chart:



(Time line chart)

2.6 Software Development model (Software Process Model):

This project will follow the incremental model of software development. The project is decided to follow incremental model due to following reasons:

- Less cost and time will be required to develop the core product.
- This is a smaller scale system.
- Testing each increment is likely to be easier than testing the entire system.
- The feedback providing at each iteration is useful for determining the final requirement of system.

The iteration in the system will be based on the accuracy of the prediction which the system generates. Proper choice of number of hidden layer inputs in ANN can be very important factor in generating the output with higher accuracy. Plus, slight changes in the algorithms will also play an important role with each iteration of the system.

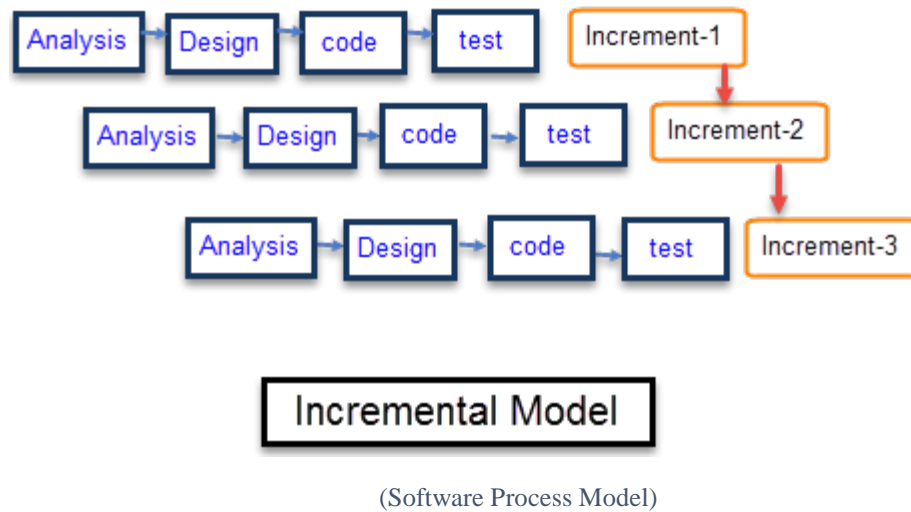
The incremental model combines elements of linear sequential model with the iterative philosophy of prototyping. Each linear sequence produces a deliverable “increment” of the software. We call first increment as the core product. In core product, basic requirements are added but some unknown supplementary features will remain undelivered. This core product will be used by customer to evolve the system and next increment is planned to develop.

During first requirement analysis phase, we will communicate with customers to specify as many requirements as possible. The first version of product with minimal and essential feature will be launched. Based on the feedback and experience of that version, list of additional features will be added. This process is repeated following the delivery of each increment, until the complete product is produced.

It may also possible that there is some change in the requirements of the customers over the time and therefore new design and implementation for this program would have to be prepared. Incremental model is quite flexible in this scenario.

There is not huge risk involved in this system. Moreover, the system architecture is also not that much complex. So, there is no need for spiral model of software development.

After some crucial increments, the final product will be deployed.



2.7 Tools and Technology / Minimum Hardware and Software Requirements:

- Operating System: Windows 7 or later versions.
- RAM: 2 GB or above.
- Processor: Intel or AMD.
- Tools and Technology used:
 - Historical data from finance.yahoo.in
 - Python nsepy library
 - Python pandas library
 - Python pyplot library
 - Python sklearn library
 - Various Financial News and Business articles.

2.8 Requirement Validation:

It's a process of ensuring the specified requirements meet the customer needs. It's concerned with finding problems with the requirements.

These problems can lead to extensive rework costs when these they are discovered in the later stages, or after the system is in service.

The cost of fixing a requirements problem by making a system change is usually much greater than repairing design or code errors. Because a change to the requirements usually means the design and implementation must also be changed, and re-tested.

So, this phase will validate the requirements which are gathered and an assurance with the customers that their requirements are the same as prescribed earlier.

After the extensive analysis of the problems in the system, we are familiarized with the requirement that the current system needs. The requirement that the system needs is categorized into the functional requirements and non-functional requirements. These requirements are listed below:

Functional Requirements

Functional requirements are the functions or features that must be included in any system to satisfy the business needs and be acceptable to the users. Based on this, the functional requirements of the system must require are as follows:

- The system should be able to generate an approximate share price.
- The system collects the accurate data from internet using various library and tools as prescribed in section 2.7 in consistent manner and regularly.
- The system will also count technical indicators and judge according to those indicators.
- The system will perform Textual analysis using online business articles and news.

Non-Functional Requirements

Non-functional requirement is a description of features, characteristics and attributes of the system as well as any constraints that may limit the boundaries of the proposed system. The non-functional requirements are essentially based on the performance, information, economy, control and security efficiency and services. These non-functional requirements are as follows:

- The system should provide better accuracy.
- The system should have simple interface for users to use.
- To perform efficiently in short amount of time.

The system assessment on the stocks from India's Bombay Stock Exchange is carried out. For given day's open index, day's high, day's low, volume and adjacent values along with the stock news textual data, our forecaster will forecast the final index price for given trading date.

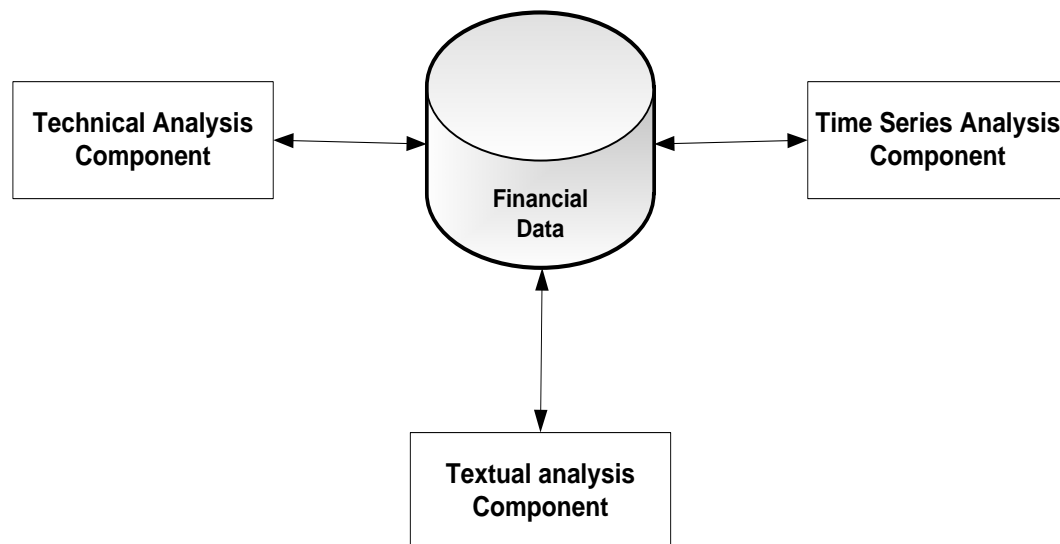
2.9 System Architecture:

The proposed system will be of Repository Architecture.

The majority of systems that use large amounts of data are organized around a shared database or repository. All data in a system is managed in a central repository that is accessible to all system components. Components do not interact directly, only through the repository.

This model is therefore suited to applications in which data is generated by one component, stored in central repository and used by another. Stock Price Prediction Program is such a system.

It is an efficient way to share large amounts of data. There is no need to transmit data explicitly from one component to another. However, components must operate around an agreed repository data model. It may be difficult to distribute the repository over a number of machines. Although it is possible to distribute a logically centralized repository, there may be problems with data redundancy and inconsistency. Such things can happen while collecting financial data from internet regularly, and sometimes used redundantly or inconsistently by time series analysis and technical analysis.



(System Architecture)

The various reasons why Repository Architecture is suitable for Stock Price Prediction System are listed below:

- We have a system in which large volumes of information are generated that has to be stored for a long time and later used by another component.
- It is a data-driven system, where the insertion of data in the repository triggers an action.

Other features of this type of architecture are:

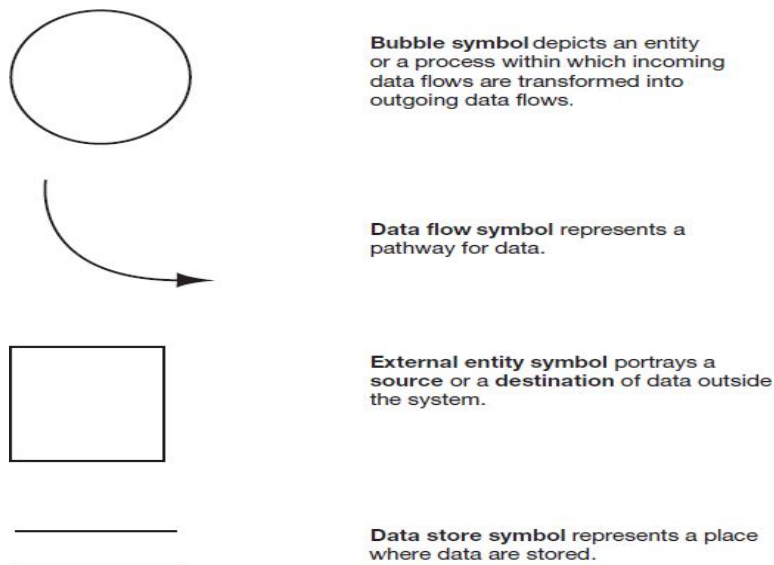
- Components are independent; they do not need to know of the existence of other components. (Time Series Analysis, Technical Analysis, and Textual Analysis)
- The repository is a single point of failure so problems in the repository affect the whole system. If the financial data not collected properly, the problem can be created.

2.10 Data-Flow Diagram:

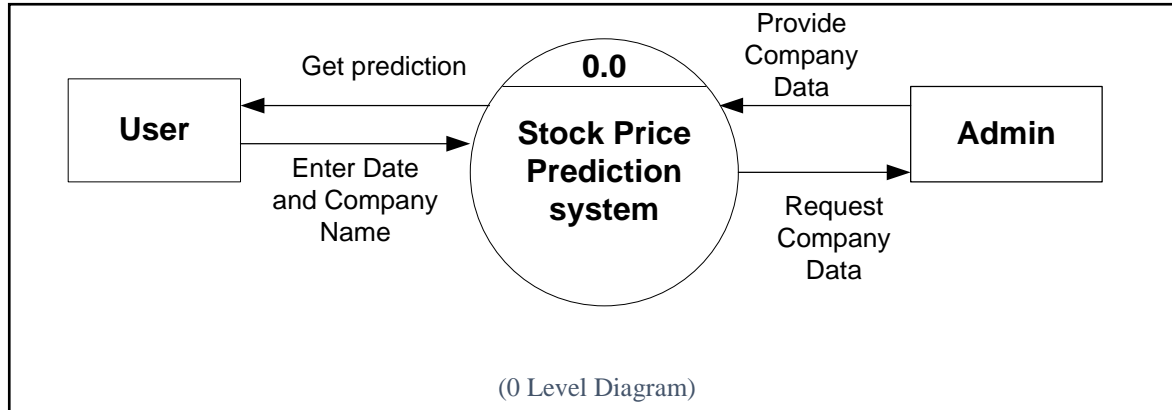
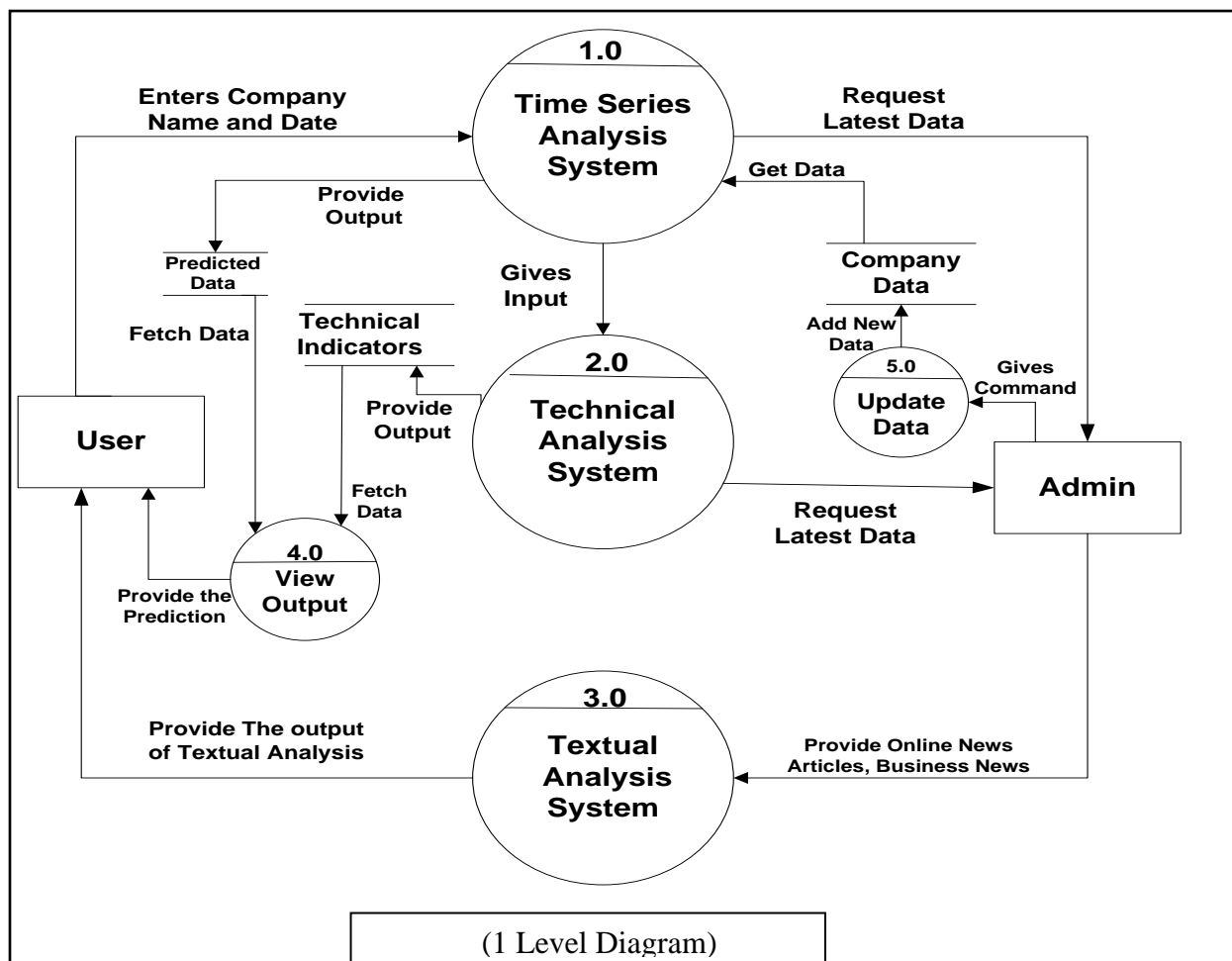
Data Flow Diagrams illustrate how data is processed by current system in terms of input and outputs. It shows the flow of data from external entities into the system, shows how data moves from one process to another, as well as logical storage. The logical storage is derived from the entities of ER Diagram.

Starting from a set of high-level functions that a system performs, and DFD model hierarchically represents various sub-functions.

There are certain symbols which represents the components of a DFD.

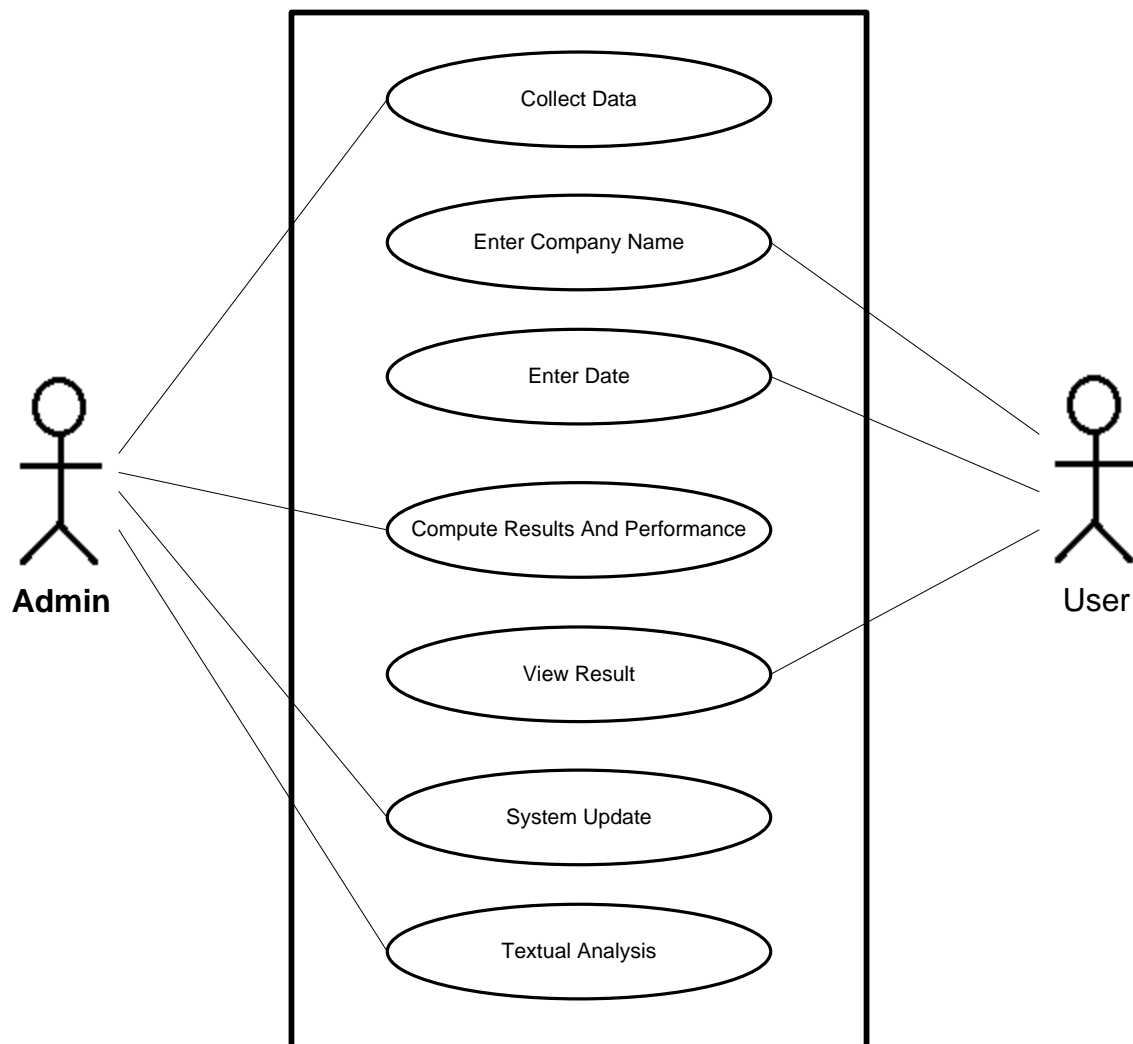


(DFD Symbol)

Context Level Diagram :Level-1 Data Flow Diagram :

2.11 Use-Case Diagram:

The use case diagram generally describes the functional requirements of the system. A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. A use case diagram can identify the different types of users of a system and the different use cases and will often be accompanied by other types of diagrams as well.



(Use Case Diagram)

Use Case-I:

Actor: Admin

Use-Case: Collect Data

Description:

The admin will collect the financial data from appropriate library using internet connection. The financial Data consist of specified Company's all historical data, starting from date 01-01-2017 to latest possible date. This data may also include the online news articles.

Use Case-II:

Actor: End-User

Use-Case: Enter Company Name

Description:

The user will choose the company name for which he want to know the future stock prices. The admin will in turn perform all operations for this specified company. The available company names which user can chose is listed in section 1.3 of this report.

Use Case-III:

Actor: End-User

Use-Case: Enter Date

Description:

*The user will enter the date for which (s)he wants to know the price of the share, so that he can get idea of where the market is progressing.

Use Case-VI:

Actor: Admin

Use-Case: Compute Results and Performance

Description:

The admin will compute the results according to machine learning algorithms as specified earlier in the project report, which involves using ANNs and SVM algorithms.

Use Case-V:

Actor: End-User

Use-Case: View Result

Description:

The end-user will view the output that is generated by using machine learning techniques as specified above. The result shows the time series analysis and technical analysis performed by the admin. Technical analysis shows certain technical indicators as specified in section 3.1.2.

Use Case-VI:

Actor: Admin

Use-Case: Update

Description:

The financial data is changing constantly. So, the admin will regularly update the financial data and add the latest data to provide more accurate predictions in the future.

Use Case-VII:

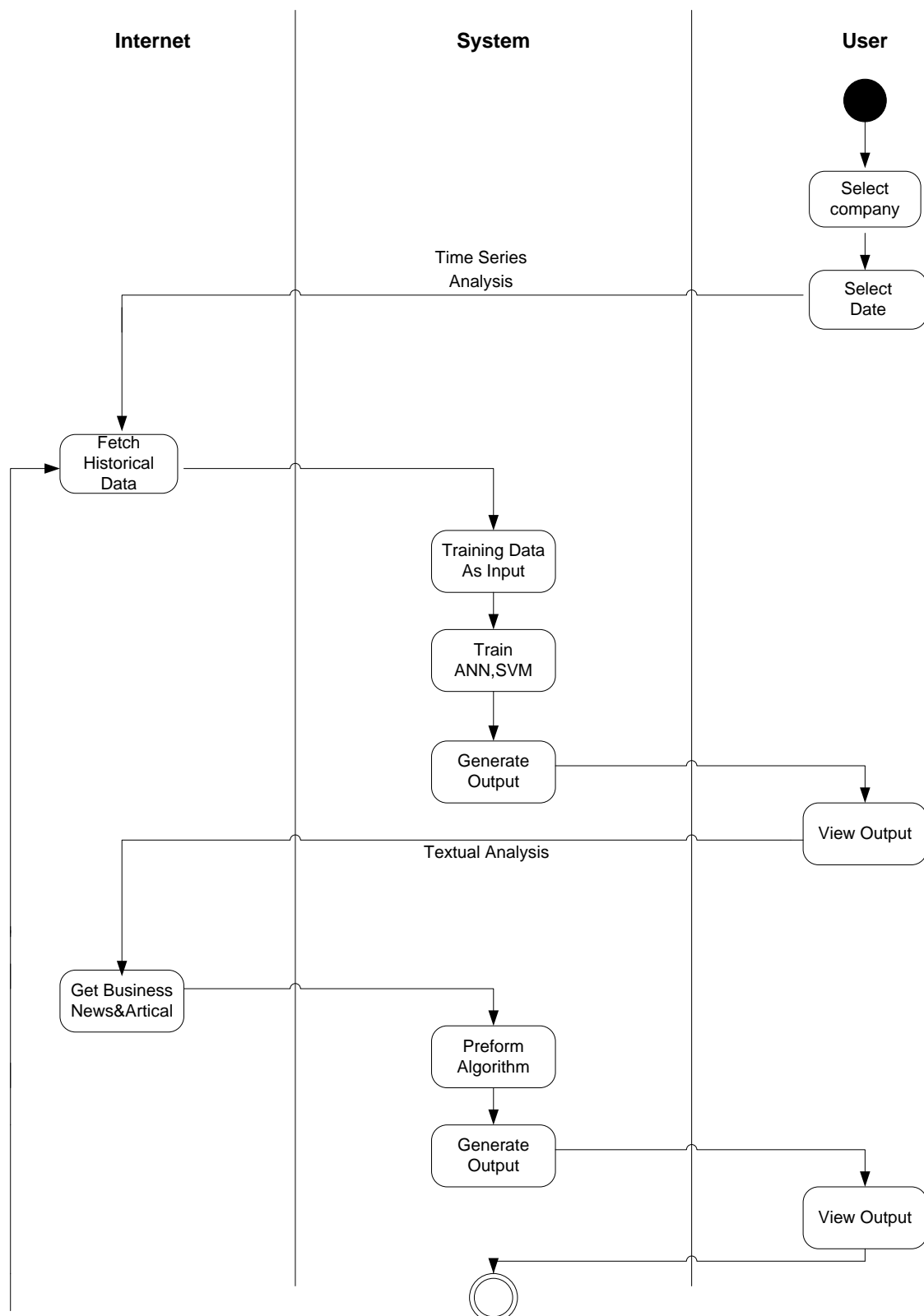
Actor: Admin

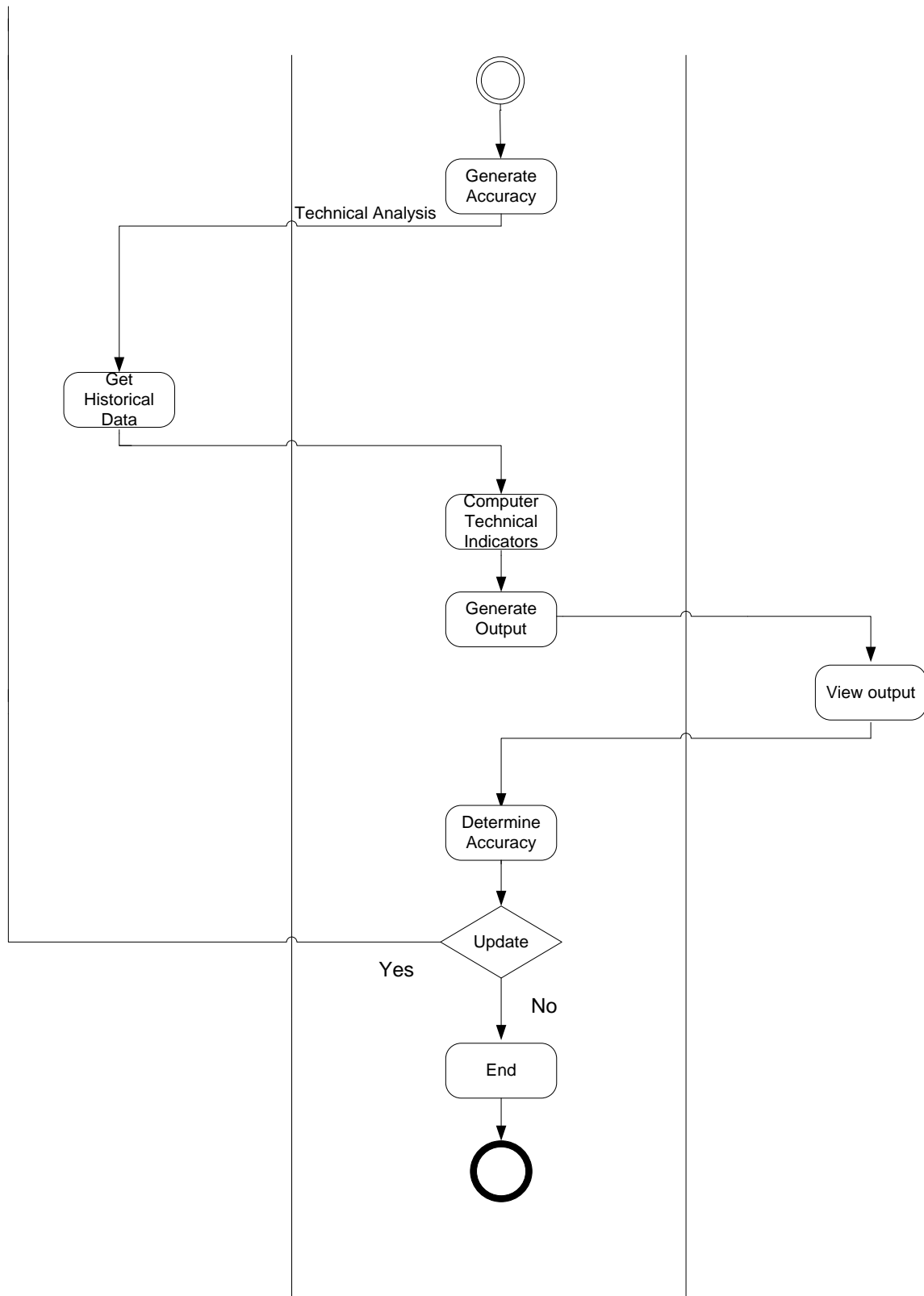
Use-Case: Textual Analysis

Description:

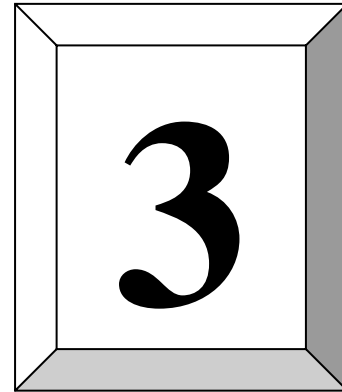
The admin will perform textual analysis. For this, the admin will collect some financial news articles and perform algorithms to give the user overall idea whether the market trend is suitable to buy the share of the specified company or not.

2.12 Activity Diagram:

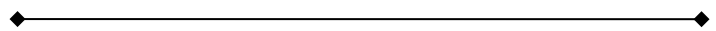




(Activity Diagram)



Chapter # 3: System Design



3.1 Database Design

3.1.1 Entity-Relationship Diagram

3.1.2 Data Dictionary

3.2 System GUI

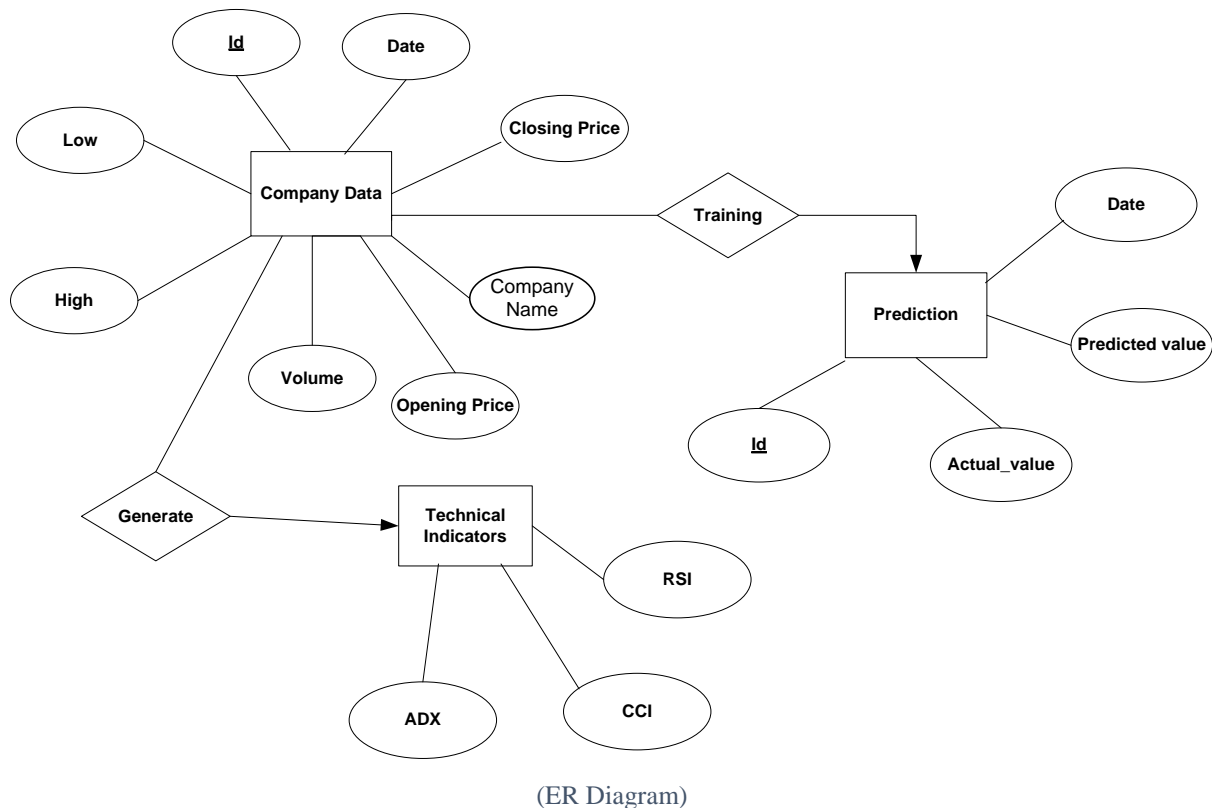
3.1 Database Design:

3.1.1 Entity-Relationship Diagram:

An entity relationship diagram (ERD) shows the relationships of entity sets stored in a database. An entity in this context is an object, a component of data. An entity set is a collection of similar entities. These entities can have attributes that define its properties.

By defining the entities, their attributes, and showing the relationships between them, an ER diagram illustrates the logical structure of databases.

ER diagrams are used to sketch out the design of a database.



Symbols in ERD:**3.1.2 Data Dictionary:**Table Name: Company Data:

Column	Data Type	Size of Field	Primary Key	Constraint	Description
Id	numeric	20	Yes	Not Null	Id of the company
Name	varchar	50	No	Not Null	Name of the Company
Date	date	-	No	Not Null	The date on which the share price is concerned.
High	numeric	(10,2)	No	Not Null	Highest Price of the day
Low	numeric	(10,2)	No	Not Null	Lowest Price of the day
OpeningPrice	numeric	(10,2)	No	Not Null	The Price of the share when the market opens
ClosingPrice	numeric	(10,2)	No	Not Null	The price of the share when the market closes
Volume	numeric	10	No	Not Null	Number of shares owned by shareholders

(Data Dictionary for Company Data)

Table Name: Predicted Data :-

Column	DataType	Size of Field	PrimaryKey	Constraint	Description
Id	numeric	20	Yes	Not Null	Id of the company
Date	date	-	No	Not Null	The date on which the share price is concerned.
PredictedValue	Numeric	(10,2)	No	Not Null	Predicted Opening value of the share on the specified date given by the user.
ActualValue	numeric	(10,2)	No	-	Actual Opening Value of the share on the specified date

(Data Dictionary for Predicted Data)

Table Name: Technical Indicators :-

Column	DataType	Size of Field	PrimaryKey	Constraint	Description
RSI	numeric	(10,2)	No	Not Null	Relative Strength Index ^[1] Suggests the overbought and oversold market signal.
CCI	numeric	(10,2)	No	Not Null	Commodity Channel Index ^[2] Identifies cyclical turns in stock price
ADX	numeric	(10,2)	No	Not Null	Average Directional Index ^[3] Discover if trend is developing

(Data Dictionary for Technical Indicators)

***Notes:**

1. $RSI = 100 - (100 / (1 + RS))$
Where:
Avg Gain = TotalGain/n
Avg Loss = TotalLoss/n
 $RS = \text{Avg Gain} / \text{Avg Loss}$

2. $CCI = 1/0.015((pt - SMA(pt))/MD(pt))$
pt = Typical Price
pt = High+Low+Close/3
MD = Mean Deviation

3. ADX
 $\text{Sum}((+DI - (-DI)) / (+DI + (-DI))) / n$

3.2 System GUI:

The screenshot displays a web-based interface for financial data analysis. At the top, there are two dropdown menus: 'Select Company' with 'HDFC Bank' selected, and 'Select Date' with '25/01/2019' selected. Below the date dropdown is the text 'For which you want to get prediction'. Underneath these are three buttons: 'Get Latest Data', 'Get Prediction', and 'See Historical Data'. The main content area is divided into three sections: 'Time Series Analysis:' showing a 'Predicted Opening Price: 1968.77', 'Technical Indicators:' listing 'RSI : 53.23', 'CCI:', 'Moving Average:', and 'Moving Oscillator:', and 'Textual Analysis:'.

Select Company	HDFC Bank	▼
Select Date	25/01/2019	▼
For which you want to get prediction		
Get Latest Data	Get Prediction	See Historical Data
<u>Time Series Analysis:</u>		
Predicted Opening Price: 1968.77		
<u>Technical Indicators:</u>		
RSI : 53.23		
CCI :		
Moving Average :		
Moving Oscillator :		
<u>Textual Analysis:</u>		

(System GUI)

References

- [1] <https://www.investopedia.com/articles/stocks/09/indian-stock-market.asp>
- [2] Raj Kumar, Anil Balara(2014), “Time Series Forecasting Of Nifty Stock Market Using Weka”, JRPS International Journal for Research Publication & Seminar Vol 05 Issue 02.
- [3] Bashambu Shallu, Sikka Aman, Negi Pallav, “Stock Price Prediction using Neural Networks”, International Journal of Advance Research, Ideas and Innovations in Technology.
- [4] A similar project report: <https://www.slideshare.net/anilsth91/stock-market-analysis-and-prediction>
- [5] <https://www.investopedia.com/terms/t/timeseries.asp>
- [6] <https://image.slidesharecdn.com/efficientmarkethypothesis-111019030158-phpapp01/95/efficient-market-hypothesis-5-728.jpg?cb=1318993357>
- [7] Luca Di Persio, Oleksandr, Honchar(2016), “Artificial Neural Networks architectures for stock price prediction: comparisons and applications”, International Journal of Circuits, Systems and Signal Processing Volume 10.
- [8] “Efficient Machine Learning Techniques for Stock Market Prediction”
- [9] Xinjie Di (2014), “Stock Trend Prediction with Technical Indicators using SVM”, SCPD student from Apple Inc.
- [10] Robert P. Schumaker and Hsinchun Chen, “Textual Analysis of Stock Market Prediction Using Financial News Articles”, Artificial Intelligence Lab, Department of Management Information Systems The University of Arizona, Tucson, Arizona 85721, USA.