Stock Market Data Analysis

Saachi Talwai
Department of Computer Science
01FB14ECS198
PES University
Bangalore, India
saachi1208@gmail.com

Sai Manasa NC
Department of Computer Science
01FB14ECS202
PES University
Bangalore, India
manasa.nallabotula@gmail.com

Spoorti Nidagundi
Department of Computer Science
01FB14ECS246
PES University
Bangalore, India
spoorti22596@gmail.com

V. Udayani
Department of Computer Science
01FB14ECS271
PES University
Bangalore, India
akhila.vaka@gmail.com

Abstract— The Stock market analysis is interesting technique that helps in determining stock price evolution. The aim of our project is to analyze stock data for a few companies listed in the National Stock exchange of India. A stock exchange market depicts savings and investments that are advantageous to increase the effectiveness of the national economy. The dataset of a single company contains its share price data and turnover data over a period of time.

Keywords—Stock markets ,National Stock Exchange , technical analysis , fundamental analysis,ARIMA Model

I. INTRODUCTION

Stock market analysis is mandatory prior to any financial investment. The most efficient way to forecast the future is to understand the present. Stock market analysis and forecasting is the process of studying and investigating data on existing stocks and trying to predict how they will do in the market. It helps the investors to make decision on whether the company is stable, growing and has an improving future based on stock reports.
Identifying trends or patterns is very important. But the market never moves in a straight line. A stock will never fall continuously on a given day and rise on another.

Analysis of stocks is useful for new investors to invest in stock market based on the various factors.

Various factors influence the stock values, some of them are -
- *Demand and Supply*
- *Corporate results:* profits or progress of the company
- *Popularity:* Main Strength in hands of share buyer.

The Indian stock market plays an important role in financing Indian corporate sector.

The principal functions of stock markets are:
- Enabling mobilizing resources for investment
- Monitoring and providing liquidity for investors
- Disciplining company management

The two major stock exchanges in India are National Stock Exchange and Bombay Stock Exchange .We are using National Stock Exchange.

Two techniques used by investors to predict stock behavior:

1. Fundamental analysis
2. Technical analysis

Fundamental analysis –
Is mainly concerned with determination of intrinsic value of stocks by analyzing the fundamental factors of economy, industry and company.

Technical analysis –
Is concerned with predicting the future trend movements of stocks by using historic data and then take decision if trend movement is low or high.[4.]

First we discuss about already existing system of stock analysis, research papers related to it and various assumptions made. In the next part, we present the approach used and limitations. We conclude by providing future enhancements and various references have been included.

Different approaches used [1]:

- ARIMA Model:
  This method uses an *Autoregressive integrated moving average* model for stock price forecasting.

Summary:

Attempts have been made to develop a prediction model for forecasting the stock market trends based on various models like ARIMA and some others using Neural network system. The results obtained by the ARIMA model helped us to predict the stock market values and in turn guided the investors whether they have to buy/hold/sell a share.

II. PROBLEM STATEMENT

Developing an efficient predictive model that will assist investors take appropriate decisions regarding which company they should invest using efficient tools in R. The aim of our project is to analyze stock data for a few companies listed in the National Stock exchange of India.
Comparative analysis for intra-industry
Stock data.

Using the data we intend to:

- Study the stock trend of the company: ranging from highest share price and a given time to the lowest as well as the opening and closing prices.
- Create a predictive model to estimate future stock prices – finding the best prediction technique
- Do a comparative analysis of stock market data between

- Two competing companies within the same industry
- Between Industries
- Estimate the right time for an investor to buy, hold or sell shares.

III. APPROACH

We aim to do an analysis of how different companies perform in the same industry (with respect to share data).We also would like analyze the progress of industries as a whole predicting which industry is leading amongst the chosen domains.

TABLE I. HISTORIC STOCK DATA FROM THE NATION STOCK EXCHANGE OF INDIA DATABASE AVAILABLE ON QUANDLL .COM

Domain	COMPANIES		
	Name	From	To
Steel	1. Tata Steel Limited	2005	2016
	3. JSW Steel Limited	2005	2016
	4.Vardhman Special Steels Limited	2012	2016
Consumer-goods	1. Godrej Consumer Products Limited	2001	2016
	2. Britannia Industries Limited	1998	2016
	3. Dabur India Limited	1998	2016
	4. Nestle India Limited	2010	2016
Banking	1. Kotak Mahindra Limited	2003	2016
	2. Yes Bank Limited	2005	2016
	3.Corporation Bank	1998	2016
	4.Edelweiss ETF - Nifty Bank	2015	2016

FIG I. Dataset Screenshot

The dataset for each company consists of opening, closing, highest, lowest share values, Total Trade Quantity, Turnover (in Lakhs) and the date.

These attributes of dataset will help us value each company and in turn help us in stock forecasting based on the values the investors are looking at

Attributes involved in the dataset:

1. Opening Price: The opening price is the price at
2. which a security first trades upon the opening of an exchange on a given trading day.
3. Closing price: The price of a security at the end of the day's business in a financial market.

4. Low Price: Lowest price at which the stock is traded during the course of the day.

5. Total Trade Quantity: Total quantity of shares traded on a single day.

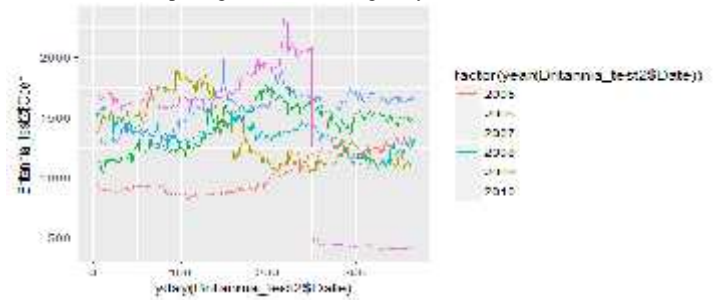6. Turnover: Share turnover is a measure of stock liquidity
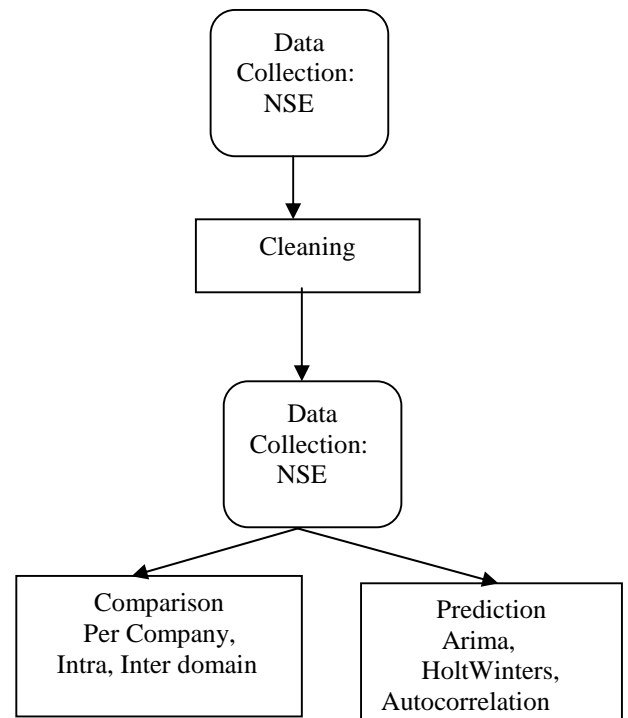
FIG II.
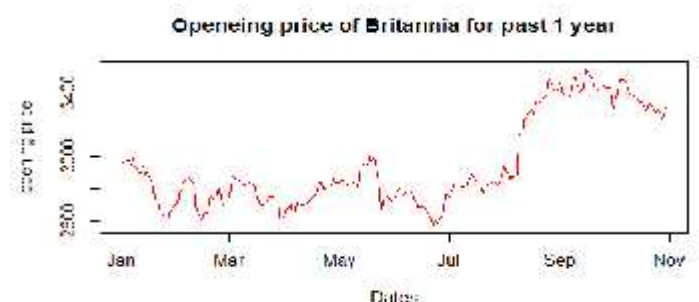
IV. COMPARATIVE ANALYSIS

A. INTRA-INDUSTRY:

We first analyze a company's share trends for a particular year.A simple plot between opening price and months has been used here. This is the opening price trend for Britannia in 2016.It hits its lowest in July.

3

FIG III. Opening Price over a range of years.

We then used ggplots to compare a company's share performance over years. Though, 2010 first shows the best opening price performance has a sudden fall around day 250 in September.

FIG IV.

Closing Prices of all Steel Industries

Everyone wants to know how ones competitor is performing, and so it is necessary to draw comparisons between competing companies of an Industry. An R tool "dygraph" has been used for this analysis. JSW steel seems to be a front runner in 2016 with respect to Closing price. Many such comparisons have been made and analyzed which helps an amateur trader decide which is that industry/company that is leading in the market,

A software tool called *Tableau* has been used to see the average

Total Trade Quantity of different companies within a single domain. This gives a clear picture of which company in the consumer goods domain is currently leading.
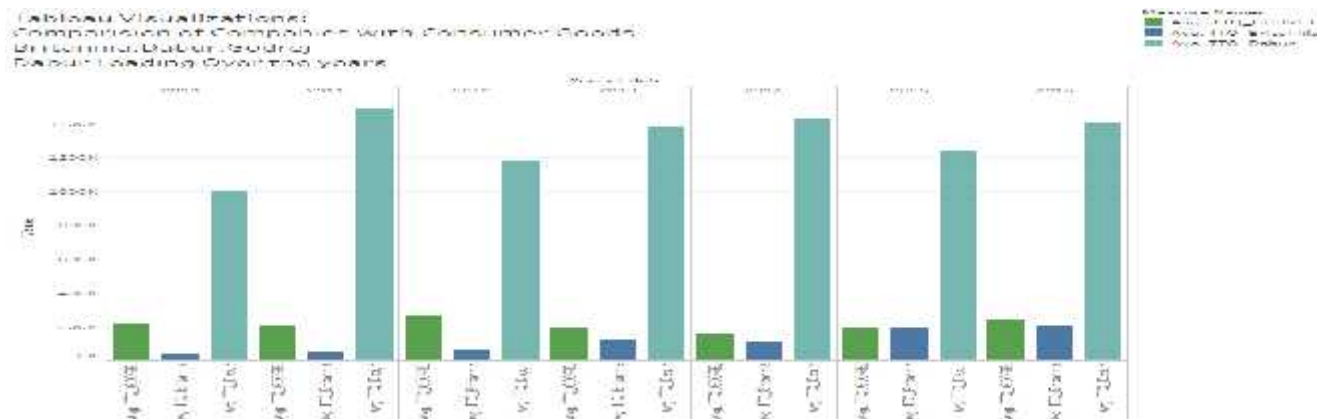

FIG V.

B. INTER-INDUSTRY:

Monthly cum Yearly Analysis of Turnover for Inter Industry data for the three domains. Banks in black, steel in blue, Consumer goods in pink. Bank performance peaks around the end of 2016.
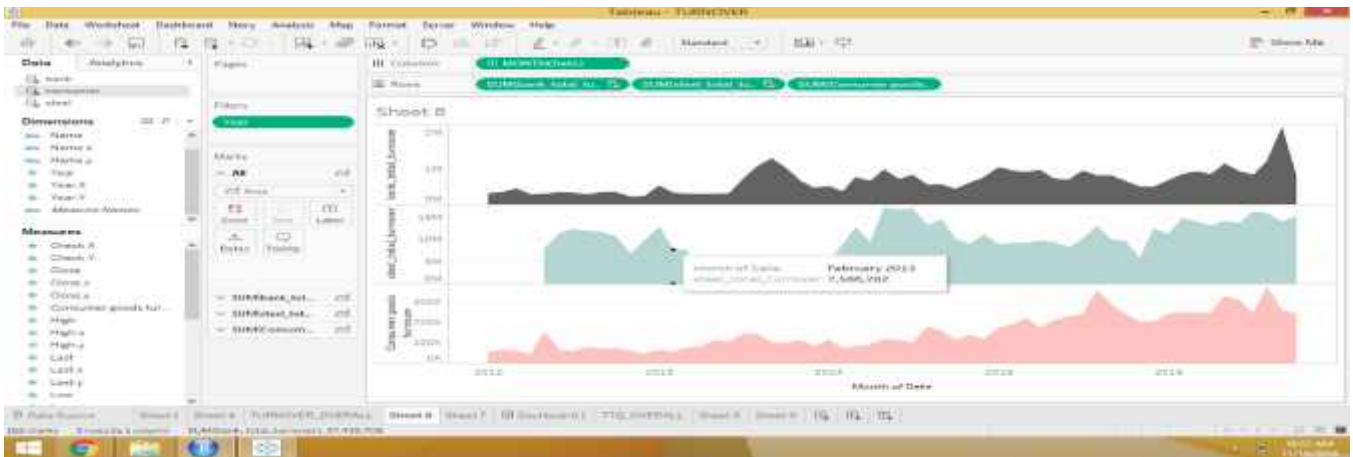

FIG VI. TURNOVER INTER-INDUSTRY

V.   PREDICTIVE ANALYSIS:

- Using Autocorrelation and Holt-winters.
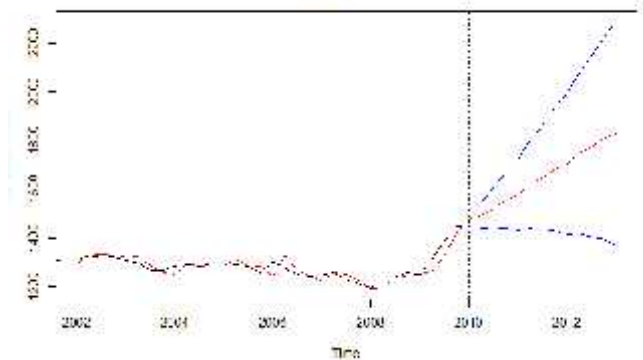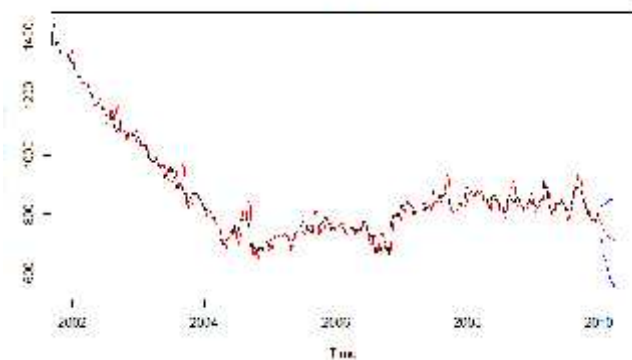- Using the ARIMA model.

Holt-Winters : Holt-Winters' method involves three smoothing parameters to smooth the data, the trend, and the seasonal index.Apply holt-winters function on the data and plot the exponential smoothing plot. We also set the interval type as "predict", and use the default 0.95 confidence level.

FIG VII.

FIG VIII.

A. AUTO-CORRELATION AND HOLT-WINTERS FILTERING:

Autocorrelation, it is the similarity between observations as a function of the time lag between them. Autocorrelation plots are a commonly-used tool for checking randomness in a data set. If random, such autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero.

Lag : a distributed lag model is a model for time series data in which a regression equation is used to predict current values of a dependent variable based on both the current values of an explanatory variable and the lagged (past period) values of this explanatory variable.

The function acf computes (and by default plots) estimates of the auto covariance or autocorrelation function. We compute it for closing prices.
The blue dotted line is the 95% confidence interval. We can see that there is the 4th and the 12th lag significant in the ACF Plot.
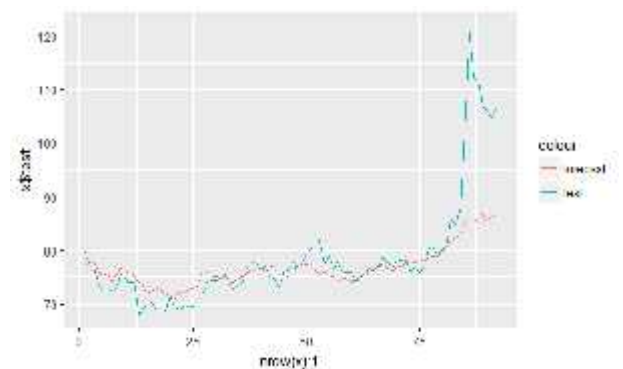
FIG IX.

FIG X.
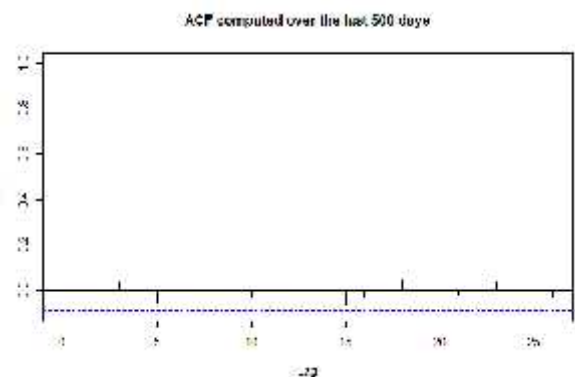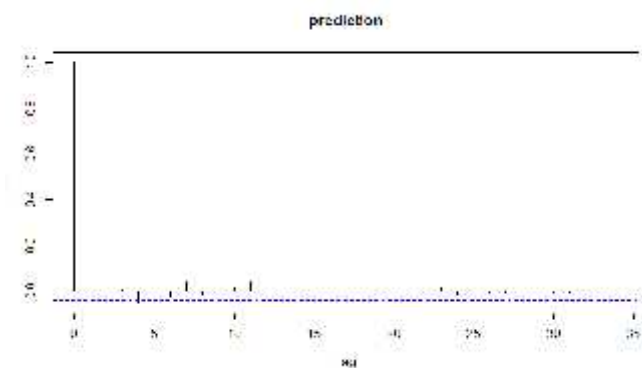
B. ARIMA MODEL:

This model was used for forecasting the future points. We have forecasted closing price values for some of our companies and also calculated a error percentage for the same.

FIG X. FORECASTED VALUES FOR VSSL

VI. CONCLUSION

Attempts have been made to develop a prediction model for forecasting the stock market trends based on various models like ARIMA, Auto-Correlation model and concepts of Holt-Winter's Filtering methods have been implemented.

The results obtained by the ARIMA model helped us to predict the stock market values and in turn guide the investors whether they have to buy/hold/sell a share during a particular period of time.

Also, various comparisons between many attributes have been shown in order to help a trader rank each domain/industry/company.

VII. CONTRIBUTION :

Saachi Talwai : Comparative analysis
Spoorti Nidagundi : Comparative analysis
Manasa N : Predictive analysis
Udayani . V : Predictive analysis

VIII. ACKNOWLEDGMENT

IX. REFERENCES:

[1] "Time_Series_Data_Analysis_for_Stock_ Market_Prediction_using_Data_Mining_ Techniques_with_R",2015.

[2] A study of Technical –Analysis-in-different-sector-Stocks",2010.

[3] http://www.slideshare.net/hemanthcrpatn a/a-study-of-technical-analysis-in-different-sectors-stocks,

[4] http://link.springer.com/chapter/10.1007/ 3-540-45718-6_38.

[5] https://www.researchgate.net/publication/ 286890497_Time_Series_Data_Analysis _for_Stock_Market_Prediction_using_Da ta_Mining_Techniques_with_R,