Stock Market Prediction for Algorithmic Trading using Machine Learning Techniques & Predictive Analytics: An Excel based automated application integrated with R and D3.JS

Prepared By:

Ms. Chandrima Chowdhury

DOEACC 'C' Level

As a Partial Fulfilment of 'C' Level Programme of NIELIT

National Institute of Electronics & Information Technology

New Delhi

TABLE OF CONTENTS

ABSTRACT

Stock price forecasting is a popular and important topic in financial and academic studies. Share Market is an untidy place for predicting since there are no significant rules to estimate or predict the price of share in the share market. Many methods like technical analysis, fundamental analysis, time series analysis and statistical analysis, etc. are all used to attempt to predict the price in the share market.

In this project we attempt to implement a Predictive Modeling and Technical Indicators Analysis approach to predict stock market prices by developing an automated stock data collection and predictive analysis tool. Predictive Modeling is very effectively implemented in forecasting stock prices, returns, and stock modeling and the most frequent methodologies are the Decision Tree algorithm and the Regression Algorithm. This project is for Indian users as the prediction is done on the listed companies of National Stock Exchange of India's NIFTY index. We outline the design of the Predictive models with its salient features and customizable parameters, and design visually interactive trend charts for stock technical indicators analysis. We select a certain group of parameters with relatively significant impact on the share price of a company. With the help of statistical analysis, the relation between the selected factors and share price is formulated which can help in forecasting accurate results. Although, share market can never be predicted, due to its vague domain, this project aims at applying Predictive Modeling Machine Learning techniques and stock indicator concepts in forecasting the stock prices.

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 the growth of the industry and commerce of the country that eventually affects the economy 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 stock market is the primary source for any company to raise funds for business expansions. It is based on the concept of demand and supply. If the demand for a company's stock is higher, then the company share price increases and if the demand for company's stock is low then the company share price decrease.

The **National Stock Exchange of India Limited** (*NSE*) is the leading stock exchange of India, located in Mumbai. The NSE was established in 1992 as the first demutualized electronic exchange in the country. NSE was the first exchange in the country to provide a modern, fully automated screen-based electronic trading system which offered easy trading facility to the investors spread across the length and breadth of the country.

The **NIFTY 50** index is National Stock Exchange of India's benchmark broad based stock market index for the Indian equity market. It represents the weighted average of 50 Indian company stocks in 12 sectors and is one of the two main stock indices used in India, the other being the BSE Sensex.

Due to involvement of many number of industries and companies, it contain very large sets of data from which it is difficult to extract information and analyze their trend of work manually. The application developed in this project, not only helps in prediction the future movement if the stock in the market, but also automate the data retrieval, trend analysis, predictive analysis and insights generation of a stock, just at the click of a button. Stock market analysis and prediction will reveal the market patterns and predict the time to purchase stock. The successful prediction of a stock's future price could yield significant profit. This is done using large historic market data of 12 months in this project, to represent varying conditions and confirming that the time series patterns have statistically significant predictive power for high probability of profitable trades and high profitable returns for the competitive business investment.

Statement about the problem

The Stock Market is a complex and dynamical system, & is influenced by many factors that are subject to uncertainty. So, it is a difficult task to forecast stock price movements. Due to technology and globalization of business & financial markets it is important to predict the stock prices more quickly & accurately. Automated User friendly Trading application can be developed based on financial predictive indicator algorithms & machine learning techniques to predict the performance of stocks in NSE's Nifty 50 Index.

My role in the Project

I will be involved in the end to end activities for the creation of this project under the guidance of my supervisor and serve as an "Individual Contributor" for this project fulfilling the following roles:

- Idea Generation – *Research on Indian Stock Market and Prediction Analysis of the NIFTY stocks*
- Planning – *Design blueprint of the methodology for Stock Data Retrieval & Buy/Sell Prediction*
- Design and Development – *Developing prediction models based on the planned methodology and designing the prediction application tool for stock data retrieval and stock movements analysis*
- Testing & Documentation – *Validating the accuracy of prediction models, reducing the error rate and documenting the results*

What contribution would the Project make?

- Making predictions is an interesting exercise, but the real fun is looking at how well these forecasts would play out in the actual market. Using the Stock Movement Predictor application, developed in this project, an investor can "play" the stock market using our in-built **prediction models (Decision Tree & Regression Analysis)** over an evaluation period. The investors will use a strategy informed by our model which they can then compare to the simple strategy of buying and holding the stock over an entire period.
- Movement of Stock Market can also be predicted by **Technical Analysis**. The main two component of Technical Analysis are price and volume and on these two data whole stock market can predicted. Stock Market movement is nothing but a mix and match of Mathematics and Human psychology and Technical Analysis is all about these two attributes.
    - When market moves positive people invest expecting a further positive movement but when small downtrend is seen people book profits expecting, small generated profits would be lost.
    - When market goes down, people hold thinking that they would exit if the market falls further. They average it out and thus increase their losses.
    - The correct way is to cut the losses by placing correct stop loss and add up more stocks with the profits with maintaining trailing stop losses but unfortunately 97% traders and investors do it exactly opposite. Only 3% people do it in correct way and hence they are called as **smart investors**

This project aims to contribute to the **smart investors** by easing out their market movement analysis. The tool developed in this project generates trend charts and prediction models using latest data for a NIFTY stock at the click of a button, thus speeding up their investing analysis and decision making process.

Limitations:

Following are some of the limitations of this project:

- A piece of data which is **missing in this project is the intraday prices**, i.e. the prices minute by minute. However, intraday prices are not as freely available as interday prices and are considered a commodity in themselves. To get hold of such a dataset would incur a large cost, one that is not within the budget of a project such as this.
- Another important piece of **missing data is the order book**. The order book is a record of live buy and sell orders for a particular stock. Successful orders are matched off against the order book by the exchange. It is easy to imagine that the order book contains useful data. For instance, the weighted average of orders might be predictive of the price. However access to this data is extremely costly and far beyond what most casual investors can afford, let alone the budget for this project.
- We are trying to quantify the true value of a company when almost every company has in some way or another some purely qualitative value. Fundamental Analysis methods do not attempt to capture these qualitative values. For example, it is not possible to quantify the value of a brand, the size of its customer base, or a competitive advantage. Until these values are quantified, it leaves a large gap in what an algorithmic style approach can achieve. For instance, what algorithm, would have valued WhatsApp at $22 billion while still making an year-on-year loss? In fundamental analysis we are **limited to purely quantitative company metrics**.

Objective and Scope of the Project:

Objective:

- **To add to the academic understanding of stock market prediction:**
  - With a greater understanding of how the market moves, investors will be better equipped to prevent another financial crisis.
  - Evaluate some existing strategies from a rigorous scientific perspective and provide a quantitative evaluation of new strategies.
- **Provide an automated Stock Prediction Tool to Traders to:**
  - Make Buy/Sell Decisions
  - Distinguish between conservative and risky stocks

Scope:

- No attempt is made in this project at portfolio management. Portfolio management is largely an extra step done after an investor has made a prediction on which direction any particular stock will move. The investor may choose to allocate funds across a range of stocks in such a way to minimize his or her risk. For instance, the investor may choose not to invest all of their funds into a single company lest that company takes unexpected turn. A more common approach would be for an investor to invest across a broad range of stocks based on some criteria he has decided on before. **This project will focus exclusively on predicting the daily trend (price movement) of individual stocks. The project will make no attempt to deciding how much money to allocate to each prediction. More so, the project will analyze the accuracies of these predictions.**

- A distinction must be made between the trading algorithms studied in this project and high frequency trading (HFT) algorithms. HFT algorithms make little use of intelligent prediction and instead rely on being the fastest algorithm in the market. These algorithms operate in fractions of a second. **The algorithms presented in this report will operate on the order of days and will attempt to be truly predictive of the market.**

Theoretical background

In the last few decades forecasting of stock returns has become an important field of research. In most of the cases the researchers had attempted to establish a linear relationship between the input macroeconomic variables and the stock returns. After the discovery of nonlinearity in the stock market index returns, many literatures have come up in nonlinear statistical modeling of the stock returns, most of them required that the nonlinear model be specified before the estimation is done. But since stock market return is noisy, uncertain, chaotic and nonlinear in nature, Predictive Modeling & Machine Learning has evolved in capturing the structural relationship between a stock's performance and its determinant factors more accurately than many other statistical techniques.

In literature, different sets of input variables are used to predict stock returns. In fact, different input variables are used to predict the same set of stock return data. Some researchers used input data from a single time series where others considered the inclusion of heterogeneous market information and macro-economic variables. Some researchers even preprocessed these input data sets before feeding it to the Predictive Model for forecasting.

Min and Lee were doing prediction of bankruptcy using machine learning. They evaluated methods based on Support Vector Machine, multiple discriminant analysis, logistic regression analysis, and three-layer fully connected back-propagation neural networks. Their results indicated that support vector machines outperformed other approaches.

A Decision Tree is a useful and popular classification technique that inductively learns a model from a given set of data. One reason for its popularity stems from the availability of existing algorithms that can be used to build decision trees, such as CART (Breiman et al., 1984), ID3 (Quinlan, 1986), and C4.5 (Quinlan, 1993). These algorithms all learn decision trees from a supplied set of training data, but do so in slightly different ways. As discussed in the introduction, a classifier is built by analyzing training data. That is to say, a classifier is built by analyzing a collection of instances where each instance is composed of a set of attribute values and a known class value. These decision tree algorithms build top down structures that partition instances into separate classes, and it is hoped that these structures generalize well to instances with unknown class values. This would mean that the decision trees have fulfilled their objectives and have indeed discovered some underlying property of the data (Quinlan, 1986).

Tsai and Wang did a research where they tried to predict stock prices by using ensemble learning, composed of decision trees and artificial neural networks. They created dataset from Taiwanese stock market data, taking into account fundamental indexes, technical indexes, and macroeconomic indexes. The performance of Decision Tree + Artificial Neural Network trained on Taiwan stock exchange data showed Fscore performance of 77%. Single algorithms showed F-score performance up to 67%.

Josip Arneric, Elza Jurun, Snjezana pivac describes that technical analysis is done to find out the price movements whereas fundamental analysis is done to predict values by looking at the fundamentals of a particular company. They focuse on technical analysis and define that trend can be of two types on the basis of either time structure or general direction.

Professor Veroljub says that the way of investing is to sell when prices are at top and to buy when prices are at lower whatever the patterns are. In his articles he has discussed the market efficiency theory, Classical theory, confidence theory and Dow Theory. He also differentiates between the Classical and Confidence theory.

Wing-Keung Wong, Meher Manzur and Boon-Kiat Chew (2002) article discuss that the helpful principle of technical analysis is to identify trends and then go with the trend whether it is occurring randomly or due to

fundamental factors. He also discussed the techniques of moving averages and relative strength index (RSI) by applying it on Singapore stock exchanges.

There are many tools and software available out there that provide forecasting of stock market entities, share quantity and share value for a given financial organization. Most of them claim to predict the stock market with near to 100% accuracy but the opinions from the users vary. Some of the popular tools and software with their methodologies are mentioned as follows.

- **inteliCharts Predictive Stock Market Analytics:**
  It is a quantitative modeling tool used for financial time series forecasting. The system is adaptive in its core as it learns the patterns and geometrical relationships defined by historical time series data points, which are unique for each individual stock, index, or another financial instrument.

- **Markettrak**
  Its stock market forecast system consists of two major parts: an extensive database and a forecast model. The forecast model reads the database and then makes a prediction of where the market is headed. From this prediction, it determines a trading position for the Dow Diamonds or the SP500 Spiders. The database and forecast are updated daily at the close of trading.

- **Stock-Forecasting.com**
  www.stock- forecasting.com (Center of Mathematics & Science, Inc., Chicago, United States of America) provides innovative price-prediction technology for active Day Traders, Short- and Long-term Investors. They develop web-based software for stock market forecasting and analysis.

  The Stock-Forecasting software predicts stock prices, generates trading "Buy-Hold-Sell" signals, computes the most profitable company to invest in and analyzes the accuracy of predictions.

Definition of problem

The Stock Market is a complex and dynamical system, & is influenced by many factors that are subject to uncertainty. So, it is a difficult task to forecast stock price movements. Due to technology and globalization of business & financial markets it is important to predict the stock prices more quickly & accurately. Automated User friendly Trading application can be developed based on financial predictive indicator algorithms & machine learning techniques to predict the performance of stocks in NSE's Nifty 50 Index.

Investors prefer stock market investments as they have the opportunity of highest return over other schemes. Nifty (benchmark of NSE India) is a well diversified index consisting of 50 major stocks from 21 sectors of the Indian economy. However, trading through Stock buy/sell prediction computer algorithms is still in its nascent stage in the Indian stock market. The need of an automated user friendly trading predictor system, which predicts stock price upward/downward movements, is necessary in the Indian stock market, given the explosion of algorithmic trading, being one of the most prominent trends in the global financial industry over recent decade. A Stock Prediction Application will be developed in this project using Nifty data, keeping in mind the following three steps:

1. The system has to have some models generating Stock Market predictions using **Financial Stock Predictor** Functions *(**E.g.**: **Williams %R**)* and **Machine Learning** Techniques *(**E.g.**: **Decision Trees**)*

2. **Back testing** of the **Prediction Models** is essential to evaluate the trading system's performance on historical market data and thus determine the viability of the system

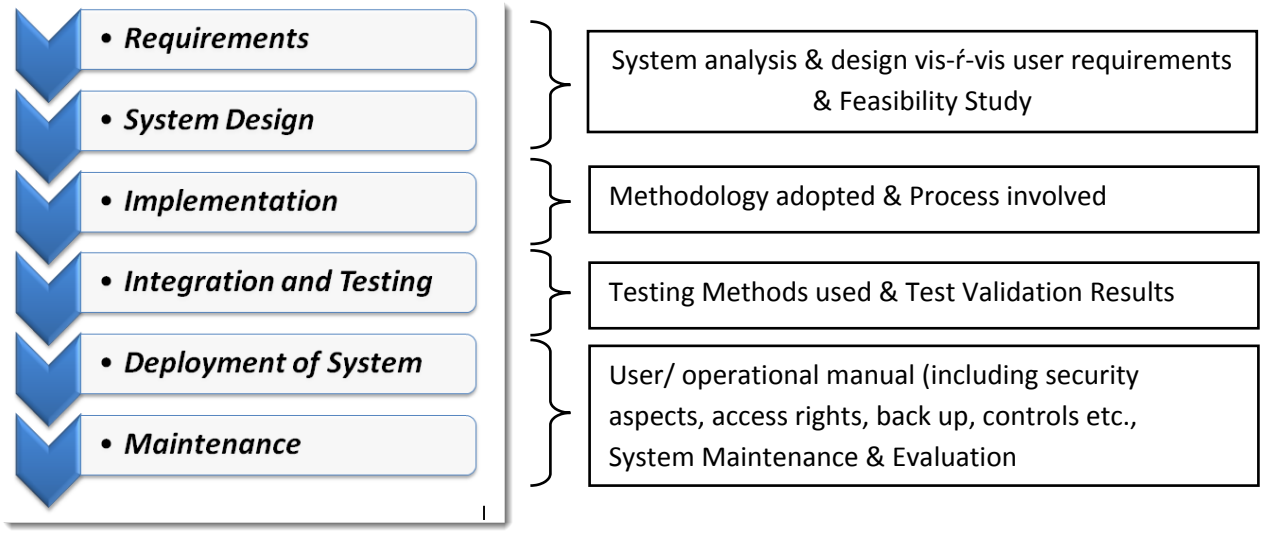
3. An analytical insight has to be provided, of whether the stock is: ***"Bullish"*** or ***"Bearish"***

Project Lifecycle - Waterfall Model

The Project Lifecycle Model used for this project is **Waterfall Model**. The Waterfall Model was the first Process Model to be introduced. It is very simple to understand and use. In a Waterfall model, each phase must be completed before the next phase can begin and there is no overlapping in the phases. Waterfall model is the earliest SDLC approach that was used for software development.

In "The Waterfall" approach, the whole process of project application development is divided into separate phases. The outcome of one phase acts as the input for the next phase sequentially. This means that any phase in the development process begins only if the previous phase is complete. The waterfall model is a sequential design process in which progress is seen as flowing steadily downwards (like a waterfall) through the phases of Conception, Initiation, Analysis, Design, Construction, Testing, Production/Implementation and Maintenance.

As the Waterfall Model illustrates the application development process in a linear sequential flow; hence it is also referred to as a **Linear-Sequential Life Cycle Model**.



a) System Analysis:

i. User Requirement Analysis:

- **System Overview:**

  This system named "**Stock Buy/Sell Predictive Analytics For Trading Of Nifty Stocks Using Predictive Algorithms & Machine Learning Techniques**" is a web application that aims to predict stock market value using Technical stock indicators and Prediction models: Decision Tree & Multiple Linear Regression. This project is intended to solve the economic dilemma created in individuals who want to invest in Stock Market.

- **System Features:**

1. **Stock market prediction:**

   Stock price movements are in somewhat repetitive in nature in the time series of stock values. The prediction feature of this system tries to predict the stock return in the time series value by training the Decision Tree/Regression model or analyzing the trend charts of technical indicators, which involves producing an output and correcting the error.

2. **Automated Prediction & Analysis Application:**

   The system tries to automate the stock analysis for the user by downloading latest data, analyzing technical indicator trends, creating prediction models, validating the prediction models and giving the end results to the users as to whether the stock should be bought/sold or whether the stock is stable/risky, just at the click of a button.

   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 and non-functional requirements. These requirements are listed below:

a. **Functional Requirements:**

   Functional requirement 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 that the system must require are as follows:

i. The system should be able to predict the approximate share price movement.
ii. The system should collect accurate data from the Yahoo Finance website in consistent manner.

b. **Non-Functional Requirements:**

   Non-functional requirement is a description of features, characteristics and attribute 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. Based on these the non-functional requirements are as follows:

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

ii. Feasibility Analysis:

Stock market cannot be accurately predicted. The future, like any complex problem, has far too many variables to be predicted. The stock market is a place where buyers and sellers converge. When there are more buyers than sellers, the price increases. When there are more sellers than buyers, the price decreases. So, there is a factor which causes people to buy and sell. 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.

The proposed system will not always produce accurate results since it does not account for the human behaviors. Factors like change in company's leadership, internal matters, strikes, protests, natural disasters,

and change in the authority cannot be taken into account for relating it to the change in Stock market by the machine.

The objective of the system is to give a 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.

Feasibility studies undergo four major analyses to predict the system to be success and they are as follows:

- Operational Feasibility
- Technical Feasibility
- Schedule Feasibility
- Economic Feasibility

- **Operational Feasibility:**

Operational feasibility is a measure of how well a proposed system solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis phase of system development. Operational feasibility reviews the willingness of the organization to support the proposed system. This is probably the most difficult of the feasibilities to gauge. In order to determine this feasibility, it is important to understand the management commitment to the proposed project. If the request was initiated by management, it is likely that there is management support and the system will be accepted and used. However, it is also important that the employee base will be accepting of the change. The operational feasibility is the one that will be used effectively after it has been developed. If users have difficulty with a new system, it will not produce the expected benefits. It measures the viability of a system in terms of the **PIECES** framework. The **PIECES** framework can help in identifying operational problems to be solved, and their urgency:

1. **Performance:** *Does current mode of operation provide adequate throughput and response time?*

   As compared to traditional methods of manually retrieving the stock data from the web and forecasting the stock prices with large number of manual calculations, this system plays a very important role in designing an application that automates the process of data retrieval and stock movement/price prediction with the help of a user-friendly dashboard, thus making the process easier and faster.

2. **Information:** *Does current mode provide end users and managers with timely, pertinent, accurate and usefully formatted information?*

   System provides end users with timely, pertinent, accurate and usefully formatted information. Since all the stock related information is being pulled from Yahoo Finance against a unique NSE Stock Symbol, it will provide for meaningful and accurate data to the investor. The investing decisions are made by the traditional investors manually. This results in loss of validity of data due to human error. The information handling and the investing decision in the proposed system will be driven by computerized and automatically updated prediction and validation of stock data. The human errors will be minimal. The data will be automatically updated from time to time and will be validated before the data is processed into the system.

3. **Economy:** *Does current mode of operation provide cost-effective information services to the business? Could there be a reduction in costs and/or an increase in benefits?*

Determines whether the system offers adequate service level and capacity to reduce the cost of the business or increase the profit of the business. The deployment of the proposed system, manual work will be reduced and will be replaced by an IT savvy approach. Moreover, it has also been shown in the economic feasibility report that the recommended solution is definitely going to benefit economically in the long run. The system is built on Excel, R and JavaScript. Excel and Javascript do not need any additional installation; they are in-built in every system. R needs installation but it is free software. So, overall the application is very economically feasible.

4. **Control:** *Does current mode of operation offer effective controls to protect against fraud and to guarantee accuracy and security of data and information?*

   As all the data is pulled from Yahoo Finance, which is a public stock data provider, it does not contain any confidential information which can be misused, so on that contrast there should be no use of any security corner for this system.

5. **Efficiency:** *Does current mode of operation makes maximum use of available resources, including people, time, and flow of forms?*

   Efficiency work is to ensure a proper workflow structure to store patient data; we can ensure the proper utilization of all the resources. It determines whether the system makes maximum use of available resources including time, people, flow of forms, minimum processing delay. In the current system a lot of time is wasted as the investing decisions are made by the traditional investors manually. The proposed system will be a lot efficient as it will be driven by computerized and automatically updated prediction and validation of stock data. The data will be automatically updated from time to time and will be validated before the data is processed into the system.

6. **Services:** *Does current mode of operation provide reliable service? Is it flexible and expandable?*

   The system is desirable and reliable services to those who need it and also whether the system is flexible and expandable or not. The proposed system is very much flexible for better efficiency and performance of the organization. The scalability of the proposed system will be inexhaustible as the storage capacity of the system can be increased as per requirement. This will provide a strong base for expansion. The new system will provide a high level of flexibility.

- ## Technical Feasibility:

  A large part of determining resources has to do with assessing technical feasibility. It considers the technical requirements of the proposed project. The technical requirements are then compared to the technical capability of the organization. The systems project is considered technically feasible if the internal technical capability is sufficient to support the project requirements.

  The analyst must find out whether current technical resources can be upgraded or added to in a manner that fulfils the request under consideration. This is where the expertise of system analysts is beneficial, since using their own experience and their contact with vendors they will be able to answer the question of technical feasibility. The essential questions that help in testing the technical feasibility of a system include the following:

1. Is the project feasible within the limits of current technology?
2. Does the technology exist at all?
3. Is it available within given resource constraints?
4. Is it a practical proposition?

5. Manpower- programmers, testers & debuggers
6. Software and hardware
7. Are the current technical resources sufficient for the new system?
8. Can they be upgraded to provide to provide the level of technology necessary for the new system?
9. Do we possess the necessary technical expertise, and is the schedule reasonable?
10. Can the technology be easily applied to current problems?
11. Does the technology have the capacity to handle the solution?
12. Do we currently possess the necessary technology?

Automated Stock Prediction system deals with the modern technology system that needs the well efficient technical system to run this project. All the resource constrains must be in the favor of the better influence of the system. Keeping all this facts in mind we had selected the favorable hardware and software utilities to make it more feasible.

1. **Details of Hardware and Software Used:**

- **Hardware:**
  - **Processor:** Intel ® Core ™ i5-6440HQ CPU @ 2.60GHz 2.59GHz
  - **Installed Memory (RAM):** 8.00 GB (7.64 GB Usable)
  - **System type:** 64-bit Operating System

- **Software**
  - **Back-end Data Service Provider:** Yahoo! Finance
  - **Operating System:** Windows 7
  - **Front-end Tool:** Microsoft Excel 2007 (with VBA and Analysis Toolpak VBA configuration)
  - **Back-end Tools:**
    - IIS Server local host configuration
    - Phantom JS Server
    - R for Windows 3.4.4
  - **Markup Languages:** HTML, XHTML
  - **Scripts:** D3.JS (Javascript Library based on Angular JS), VBA
  - **Style Sheets:** CSS

2. **Why Yahoo! Finance:**

Yahoo! Finance is a media property that is part of Yahoo!'s network. It provides financial news, data and commentary including stock quotes, press releases, financial reports, and original content.

The application needs latest stock OHLVC (Open Price, High Price, Low Price, Volume, Close Price) data for each NIFTY stock. Since most Finance sites tend to be overloaded with advertising and tetchy about scrapers, this can be a little challenging at first.

As the purpose of this project is to develop an automated data fetch application, the stock data should be automatically retrieved from the web, which can be done through Web scraping R. In such a case, Yahoo Finance is a good source for extracting financial data as the format of the data is mostly consistent for this source and is easily accessible through the R packages like "Rvest"

3. **Why Microsoft Excel (VBA):**

Microsoft excel has been used in this project to design the front end dashboard for the user. Few macros have also been coded using VBA for running data models and integrating the front end and backend with R and D3.JS. As Microsoft Excel is a most common platform and easily accessible to all kinds of users, an effort has been made in this project to enhance the features of Excel by adding D3 visuals and R models into it, to fulfill all the needs of a stock investor in a simple and flexible platform.

VBA (Visual Basic Applications) is a programming language which is developed by Microsoft to be used for the Microsoft office package such as Word, Access, Excel and others. It is used to customize the applications to meet the needs of the business. It is a powerful and convenient tool to perform an operation repeatedly and also helps in analyzing the data. VBA is used to access the functions of applications and controls them within some other applications. Marketing Sales reporting and analysis is done in an effective and efficient way using VBA.

VBA in excel is used to generate, format and print reports using graphical representations like charts. The reports are generated with ease and it is simple with the help of VBA. The reports are generated using various options as per the need of the management.

4. **Why R?**

R is a programming language and free software environment for statistical computing and graphics that is supported by the R Foundation for Statistical Computing.

One of the great advantages of using R for data analysis is the amount of data that can be imported over the web. This is practical because a database can be downloaded or updated with a simple command, avoiding all the manual and tedious work of collecting data manually. It is also easy to share code, as anyone can download the exact same dataset with a single line of code.

Importation of stock data from Yahoo Finance can be performed using specific packages in CRAN (Comprehensive R Archive Network) and web scraping techniques.

R Programming also gives a broad variety of statistical (direct and nonlinear modeling), techniques, which can be used for Decision Tree analysis for the purpose of this project.

5. **Why D3.JS?**

D3.**JS** is a JavaScript library for producing dynamic, interactive data visualizations in web browsers. It makes use of the widely implemented SVG, HTML5, and CSS standards. Techan JS is a visual, stock charting (Candlestick, OHLC, indicators) and technical analysis library built on D3.

For the purpose of this project, an attempt has been made to enhance the visuals Technical Trend Charts' visuals in Excel by integrating Excel with D3 and presenting the D3 visuals on Excel dashboard with user friendly tooltips and labels.

- **Economical Feasibility:**

Economic analysis could also be referred to as cost/benefit analysis. It is the most frequently used method for evaluating the effectiveness of a new system. In economic analysis the procedure is to determine the

benefits and savings that are expected from a candidate system and compare them with costs. If benefits outweigh costs, then the decision is made to design and implement the system. An entrepreneur must accurately weigh the cost versus benefits before taking an action.

Possible questions raised in economic analysis are:

1. Is the system cost effective?
2. Do benefits outweigh costs?
3. The cost of doing full system study
4. The cost of business employee time
5. Estimated cost of hardware
6. Estimated cost of software/software development
7. Is the project possible, given the resource constraints?
8. What are the savings that will result from the system?
9. Cost of employees' time for study
10. Cost of packaged software/software development
11. Selection among alternative financing arrangements (rent/lease/purchase)

The concerned business must be able to see the value of the investment it is pondering before committing to an entire system study. If short-term costs are not overshadowed by long-term gains or produce no immediate reduction in operating costs, then the system is not economically feasible, and the project should not proceed any further. If the expected benefits equal or exceed costs, the system can be judged to be economically feasible. Economic analysis is used for evaluating the effectiveness of the Proposed System. The economical feasibility will review the expected costs to see if they are in-line with the projected budget or if the project has an acceptable return on investment. At this point, the projected costs will only be a rough estimate. The exact costs are not required to determine economic feasibility. It is only required to determine if it is feasible that the project costs will fall within the target budget or return on investment. A rough estimate of the project schedule is required to determine if it would be feasible to complete the systems project within a required timeframe. The required timeframe would need to be set by the organization.

Costs & Benefit Analysis:

It is the process of analyzing the financial facts associated with the system development projects performed when conducting a preliminary investigation. The purpose of a cost/benefit analysis is to answer questions such as:

- Is the project justified (because benefits outweigh costs)?
- Can the project be done, within given cost constraints?
- What is the minimal cost to attain a certain system?
- What is the preferred alternative, among solutions?

Following is the figure showing the approx. amount of cost and benefit to the system:

Tangible Cost:-

DEVELOPMENT COSTS	
Windows 7	INR 1500

MS Office	INR -
R Studio	INR -
Phantom JS	INR -
Monitor	INR 4000
Key Board	INR 450
Mouse	INR 200
Hard Drive	INR 2000
RAM	INR 600
Graphics	INR 1500
Processor	INR 2000
TOTAL DEVELOPMENT COST	**INR 15000**

Intangible Cost:-

OPERATIONAL COSTS	
Software Upgrades	INR 1000
Hardware Upgrades (1 PCs)	INR 1000
User Training	INR 1000
Network Technician + Computer Operator	INR 50000
TOTAL OPERATIONAL COSTS	**INR 53000**

- **Schedule Feasibility:**

We may have the technology, but we may not have the skills required to properly apply that technology. All information systems professionals can learn new technologies, but that learning curve will impact the schedule of the project.

System Planning (PERT & Gantt Chart):

PERT (program evaluation and review technique) and **CPM** (critical path method), are available to assist the the management of a large-scale project that requires coordinating numerous activities. A myriad of details must be considered in planning how to coordinate all these activities, in developing a realistic schedule, and then in monitoring the progress of the project. These techniques make heavy use of *networks* to help plan and display the coordination of all the activities.

The original versions of PERT and CPM had some important differences. However, they also had a great deal in common, and the two techniques have gradually merged further over the years.

A **Gantt Chart** is a timeline that is used as a project management tool to illustrate how the project will run. You can view individual tasks, their durations and the sequencing of these tasks. View the overall timeline of the project and the expected completion date.

The **Gantt chart** and the **PERT chart** are probably the two best known charts in project management. Each of these can be used for scheduling, but because **Gantt charts** don't illustrate task dependencies and **PERT charts** can be confusing, PMs often use both.

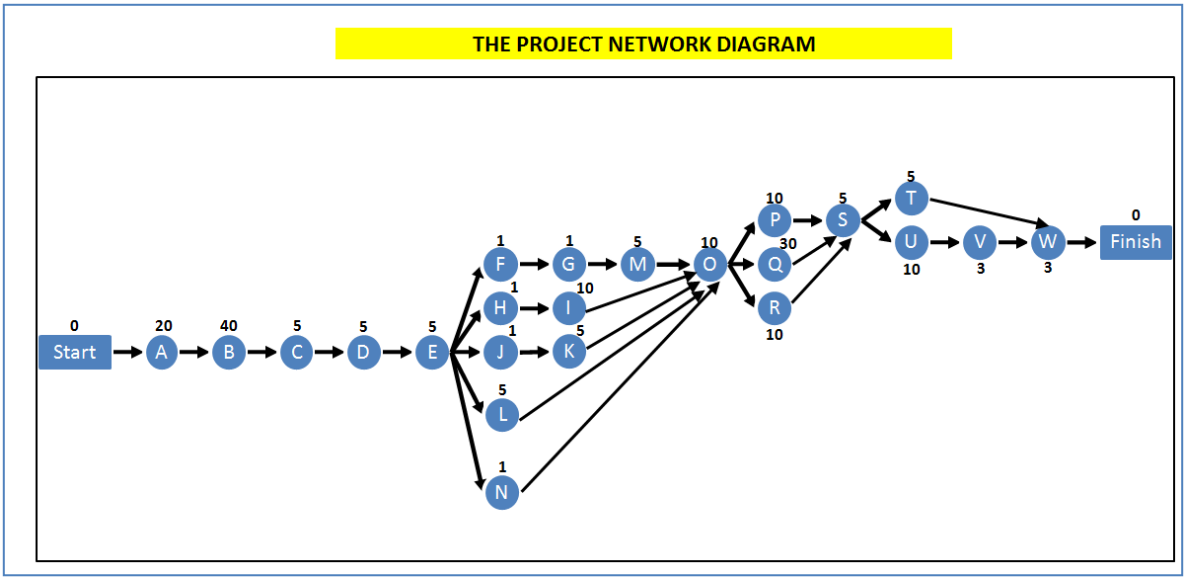
1. **Activity List for the project:**

Following is the Activity List for this project:

Activity	Activity Description	Immediate Predecessors	Estimated Duration
A	Research on Application of Stock market prediction and automated trading in the international stock market	-	20 days
B	Plan and design an excel application product blueprint to perform automated stock market prediction of NSE India - NIFTY stocks	A	40 days
C	Requirements Analysis and Feasibility study of the application	B	5 days
D	Design implementation methodology of creating the application	C	5 days
E	Decide the Hardware & Software to be used	D	5 days
F	Retrieve latest Nifty stocks' list from NSE	E	1 day
G	Retrieve latest 12 months' data for each Nifty stock from Yahoo Finance	F	1 day
H	Set up IIS localhost server on system to develop D3.JS web pages	E	1 day
I	Develop Dummy Stock Technical Indicators Trend Charts using retrieved data using D3.Js	H	10 days
J	R for Windows installation	E	1 day
K	Develop Decision Tree Stock Price Movement Prediction Model in R	J	5 days
L	Develop Multiple Linear Regression - Stock Close Price Prediction Model in Excel	E	5 days
M	Automate data retrieval from NSE and Yahoo Finance using R Web Scraping	G	5 days
N	Phantom JS Installation	E	1 day
O	Develop code to Integrate R, Excel and D3.JS using Phantom JS and VBA shell code	I,K,L,M,N	10 days
P	Develop Code to Render Decision tree Model and Related Analysis on Excel Dahsboard using integration code	O	10 days
Q	Develop Code to Render Technical Indicators Trend Charts and Related Analysis on Excel Dashboard using integration code	O	30 days
R	Develop Code to Automate Regression Model Generation and Related Analysis on Excel Dashboard using Analysis Toolpak – VBA	O	10 days
S	Design interactive Excel dashboard integrating all the codes	P,Q,R	5 days
T	Develop VBA automation code to download latest stock data, develop prediction models, develop technical indicator charts, analyze the results and conclude findings at the click of a button	S	5 days
U	Validate the results of the prediction models and calculate accuracy percentage in Excel	S	10 days
V	Modify the variables used in the models as per the validation to increase accuracy percentage	U	3 days
W	Include the results validation automation in the VBA automation code and reframe the concluded findings as per the accuracy results	T,V	3 days

2. Using A Network To Visually Display A Project:

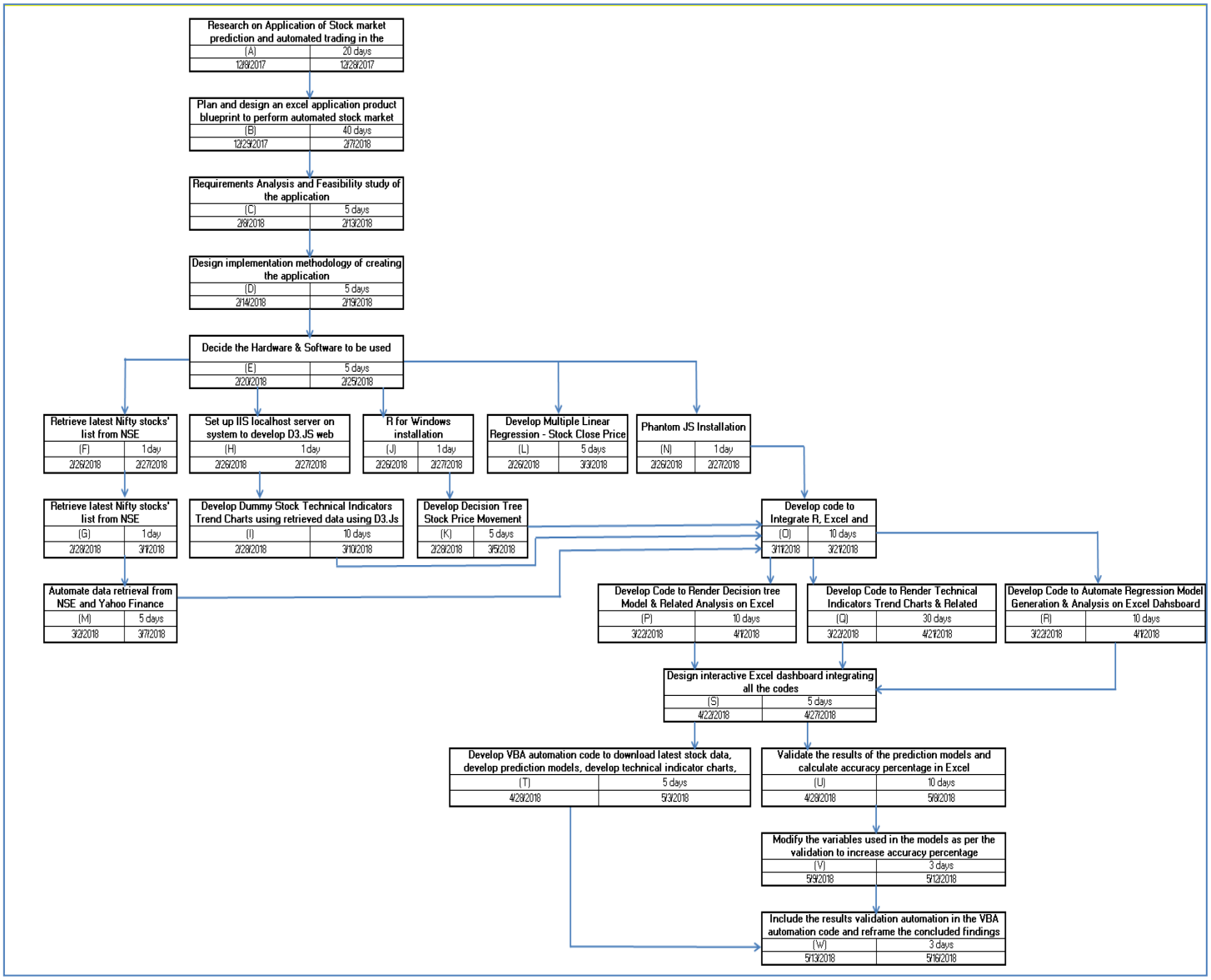
a. The Basic Network Diagram:



Activity Code:

A.	Research on Application of Stock market prediction and automated trading in the international stock market
B.	Plan and design an excel application product blueprint to perform automated stock market prediction of NSE India - NIFTY stocks
C.	Requirements Analysis and Feasibility study of the application
D.	Design implementation methodology of creating the application
E.	Decide the Hardware & Software to be used
F.	Retrieve latest Nifty stocks' list from NSE
G.	Retrieve latest 12 months' data for each Nifty stock from Yahoo Finance
H.	Set up IIS localhost server on system to develop D3.JS web pages
I.	Develop Dummy Stock Technical Indicators Trend Charts using retrieved data using D3.Js
J.	R for Windows installation
K.	Develop Decision Tree Stock Price Movement Prediction Model in R
L.	Develop Multiple Linear Regression - Stock Close Price Prediction Model in Excel
M.	Automate data retrieval from NSE and Yahoo Finance using R Web Scraping
N.	Phantom JS Installation
O.	Develop code to Integrate R, Excel and D3.JS using Phantom JS and VBA shell code
P.	Develop Code to Render Decision tree Model and Related Analysis on Excel Dahsboard using integration code
Q.	Develop Code to Render Technical Indicators Trend Charts and Related Analysis on Excel Dashboard using integration code
R.	Develop Code to Automate Regression Model Generation and Related Analysis on Excel Dahsboard using Analysis Toolpak - VBA
S.	Design interactive Excel dashboard integrating all the codes
T.	Develop VBA automation code to download latest stock data, develop prediction models, develop technical indicator charts, analyze the results and conclude findings at the click of a button
U.	Validate the results of the prediction models and calculate accuracy percentage in Excel
V.	Modify the variables used in the models as per the validation to increase accuracy percentage
W.	Include the results validation automation in the VBA automation code and reframe the concluded findings as per the accuracy results

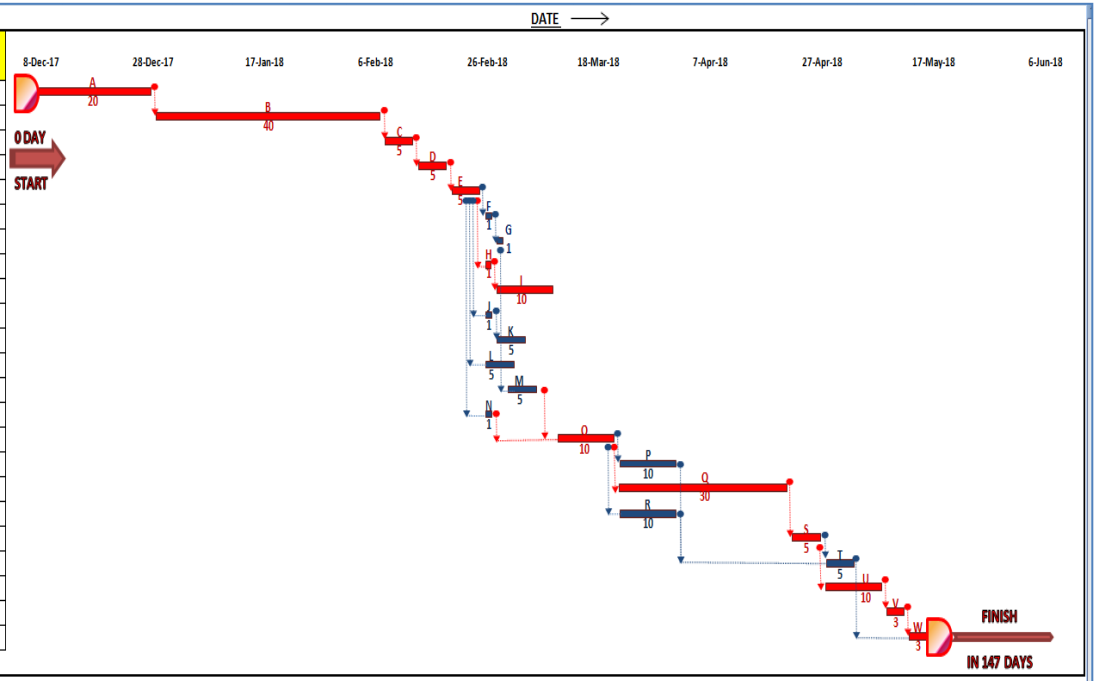
b. PERT Chart:



c. GANTT Chart:

ACTIVITY LIST

TASK	TASK DESCRIPTION	START DATE	DURATION	END DATE	PRE-DECESSORS
A	Research on Application of Stock market prediction	12/8/2017	20	12/28/2017	-
B	Plan and design an excel application product blue	12/29/2017	40	2/7/2018	A
C	Requirements Analysis and Feasibility study of the	2/8/2018	5	2/13/2018	B
D	Design implementation methodology of creating th	2/14/2018	5	2/19/2018	C
E	Decide the Hardware & Software to be used	2/20/2018	5	2/25/2018	D
F	Retrieve latest Nifty stocks' list from NSE	2/26/2018	1	2/27/2018	E
G	Retrieve latest 12 months' data for each Nifty stock	2/28/2018	1	3/1/2018	F
H	Set up IIS localhost server on system to develop D3	2/26/2018	1	2/27/2018	E
I	Develop Dummy Stock Technical Indicators Trend	2/28/2018	10	3/10/2018	H
J	R for Windows installation	2/26/2018	1	2/27/2018	E
K	Develop Decision Tree Stock Price Movement Pred	2/28/2018	5	3/5/2018	J
L	Develop Multiple Linear Regression - Stock Close I	2/26/2018	5	3/3/2018	E
M	Automate data retrieval from NSE and Yahoo Finar	3/2/2018	5	3/7/2018	G
N	Phantom JS Installation	2/26/2018	1	2/27/2018	E
O	Develop code to Integrate R, Excel and D3.JS using	3/11/2018	10	3/21/2018	I,K,L,M,N
P	Develop Code to Render Decision tree Model and F	3/22/2018	10	4/1/2018	O
Q	Develop Code to Render Technical Indicators Tren	3/22/2018	30	4/21/2018	O
R	Develop Code to Automate Regression Model Gene	3/22/2018	10	4/1/2018	O
S	Design interactive Excel dashboard integrating all	4/22/2018	5	4/27/2018	P,Q,R
T	Develop VBA automation code to download latest s	4/28/2018	5	5/3/2018	S
U	Validate the results of the prediction models and c	4/28/2018	10	5/8/2018	S
V	Modify the variables used in the models as per the	5/9/2018	3	5/12/2018	U
W	Include the results validation automation in the VB	5/13/2018	3	5/16/2018	T,V



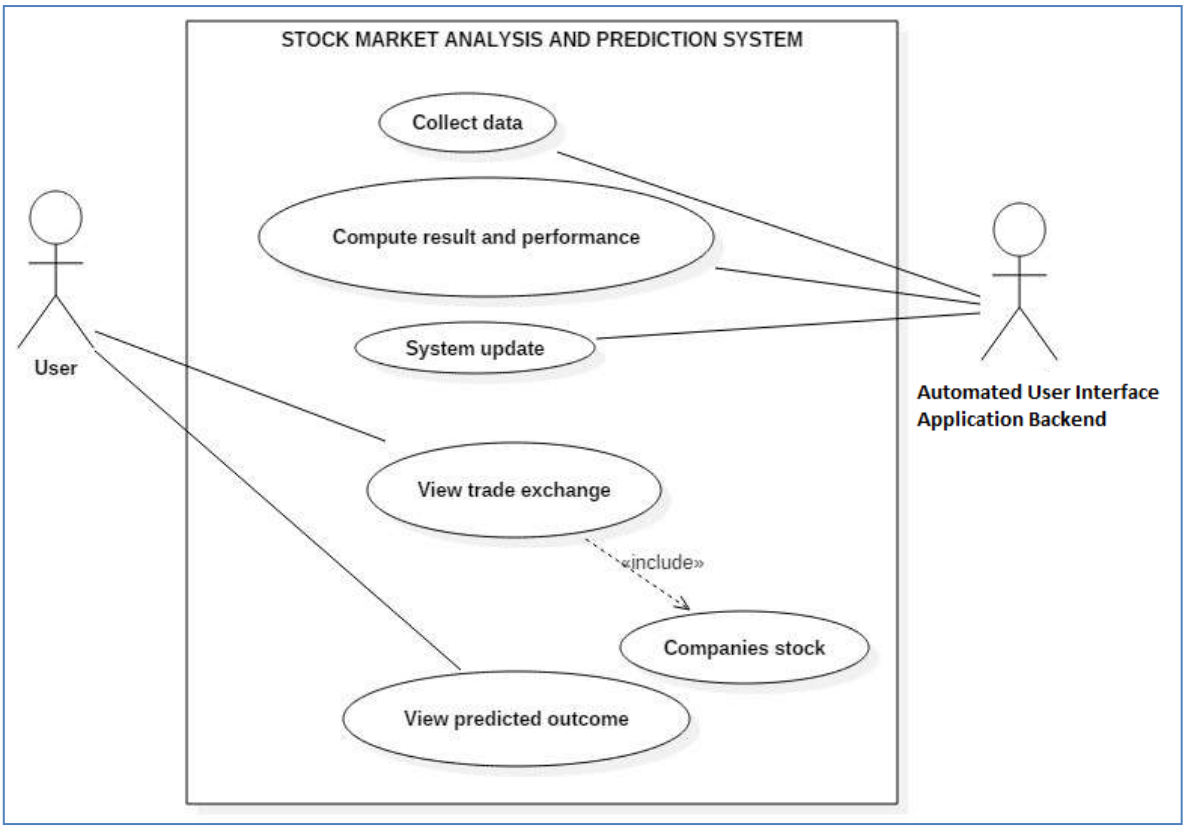
CRITICAL PATH: A + B + C + D + E + H + I + O + Q + S + U + V + W

TOTAL NO. OF DAYS: 20 + 40 + 5 + 5 + 5 + 1 + 10 + 10 + 30 + 5 + 10 + 3 + 3

= (147) DAYS

b) System Design & Architecture:

i. Use Case Diagram for the system:



Use case index:

Use case ID	Use case name	Primary actor	Scope	Complexity	Priority
1	Collect data	Automated UI Appln Backend	In	High	1
2	Compute result & prepare	Automated UI Appln Backend	In	High	1
3	System update	Automated UI Appln Backend	In	High	1
4	View trade exchange	User	In	Medium	2
5	Company stock	User	In	Medium	1
6	View predicted outcome	User	In	High	1

Use case description:

Use case ID: 1
Use case name: Collect data
Description: Every required stock data of NSE will be available in Yahoo Finance. Automated User Interface Application Backend will be able to collect the data for system.

Use case ID: 2
Use case name: Compute result and performance
Description: Prediction result will be handled and generated by Automated User Interface Application Backend. The system will be built, through which the result of prediction and system performance will be analyzed.

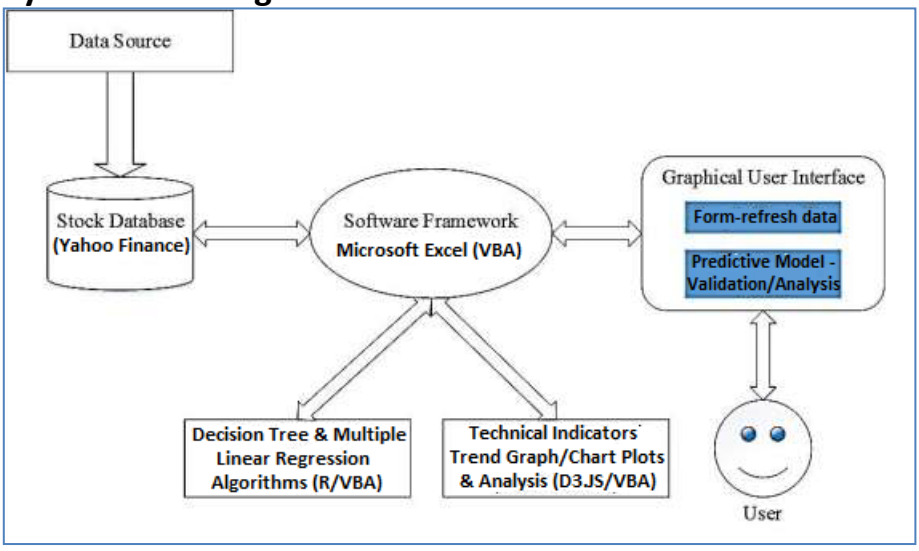
Use case ID: 3
Use case name: System update
Description: With the change of market and technology regular update of system is required. The predicted result of stock exchange and their actual price will be auto-updated by the Automated User Interface Application Backend on regular basis.

Use case ID: 4
Use case name: View traded exchange
Description: Company trading which is held at NSE India can be viewed by user.

Use Case ID: 5
Use Case Name: Company Stock
Description: It is extended feature of view traded exchange. This includes the stock value of particular company.

Use Case ID: 6
Use Case Name: View predicted outcome
Description: This use case is the most important in whole project. The key feature of this project is to predict the value of Nifty stocks. Thus, this will be available in user interface and viewer can observe them.

ii. System Flow diagram:



iii. Data Flow Diagram:

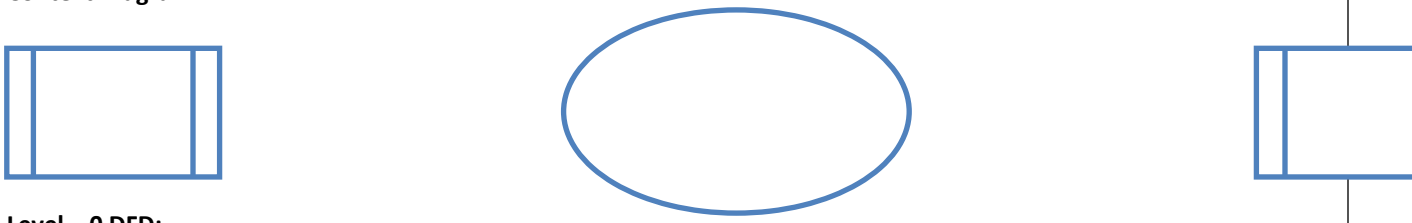
A data flow diagram shows the way information flows through a process or system. It includes data inputs and outputs, data stores, and the various sub processes the data moves through. DFDs are built using standardized symbols and notation to describe various entities and their relationships.

Data flow diagram levels:

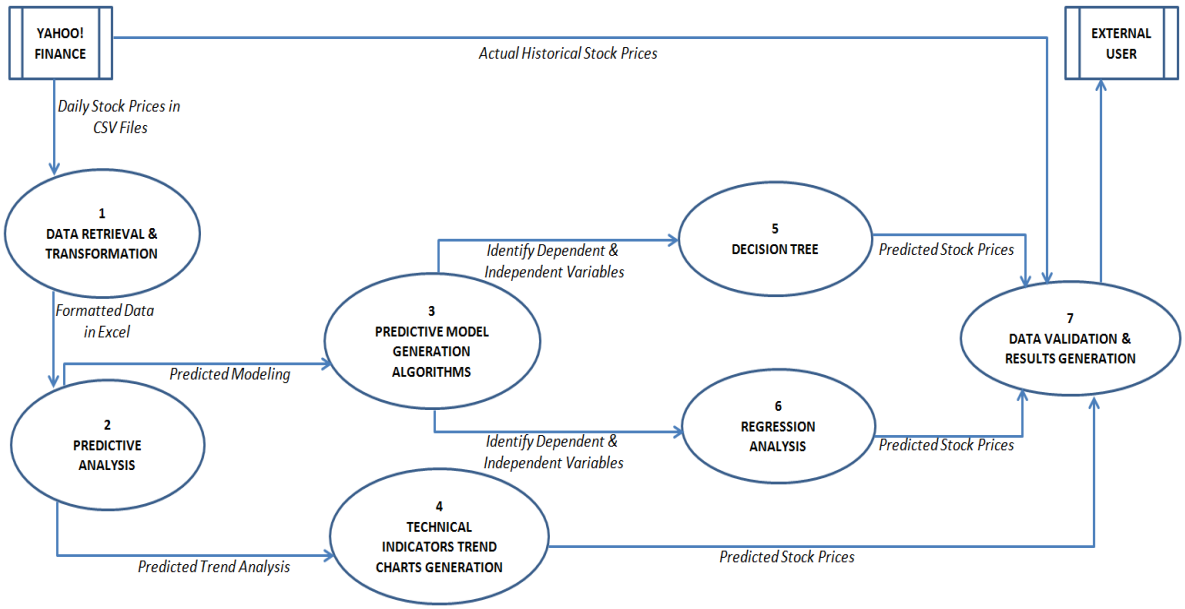
Data flow diagrams are also categorized by level. Starting with the most basic, level 0, DFDs get increasingly complex as the level increases.

Context Diagrams/Level 0 DFDs are the most basic data flow diagrams. They provide a broad view that is easily digestible but offers little detail.

Context Diagram:

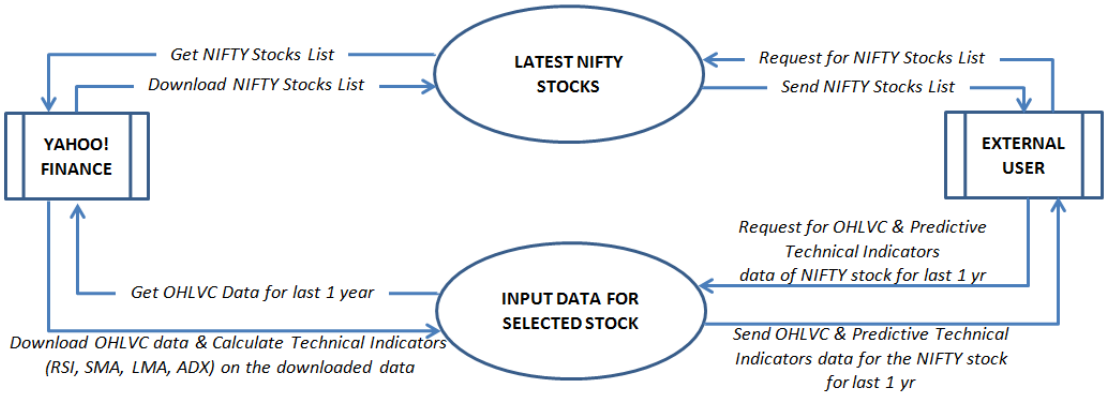


Level – 0 DFD:

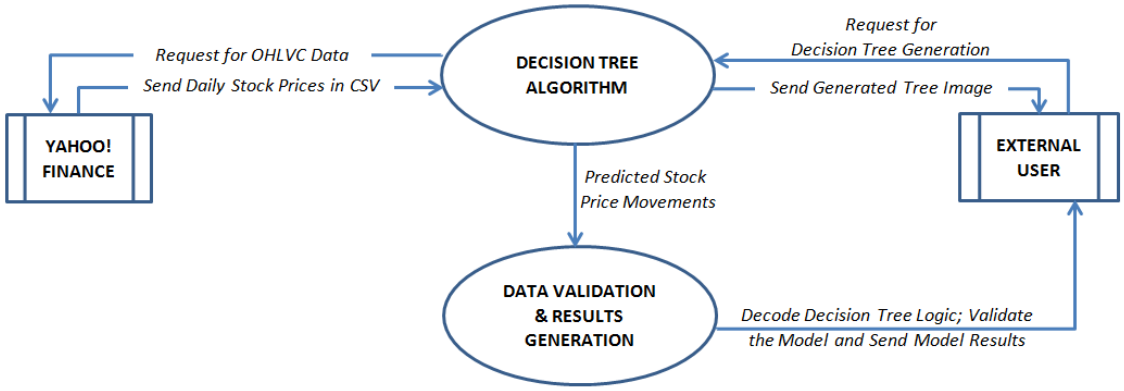


Level 1 DFDs go into more detail than a Level 0 DFD. In a level 1 data flow diagram, the single process node from the context diagram is broken down into sub processes. As these processes are added, the diagram will need additional data flows and data stores to link them together.

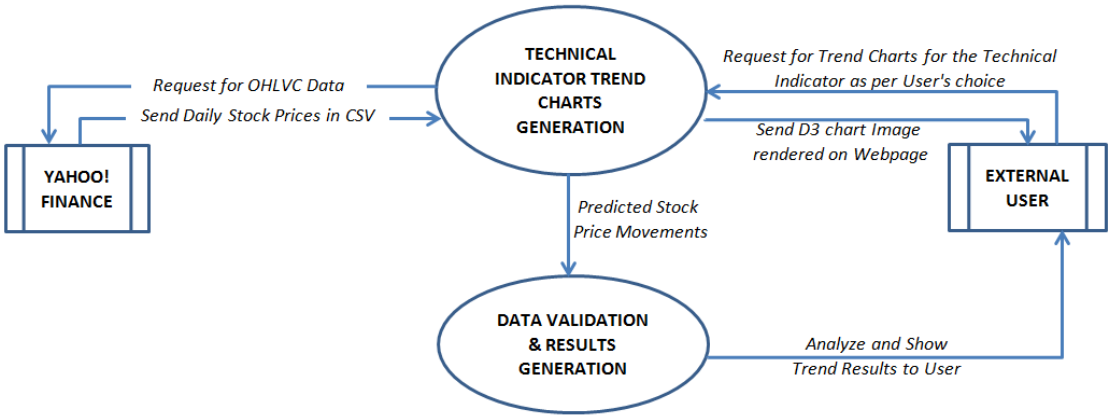
Level – 1 DFD (Data Retrieval & Transformation):



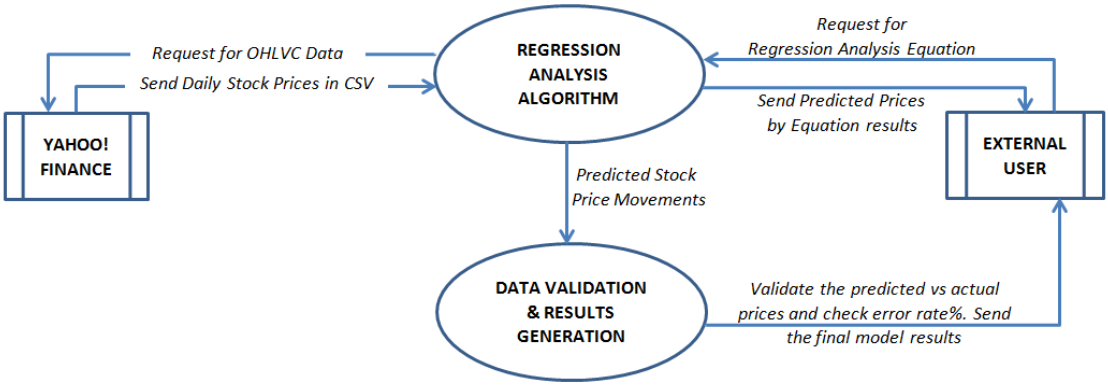
Level – 1 DFD (Decision Tree):



Level – 1 DFD (Technical Indicators Trend Charts Generation):

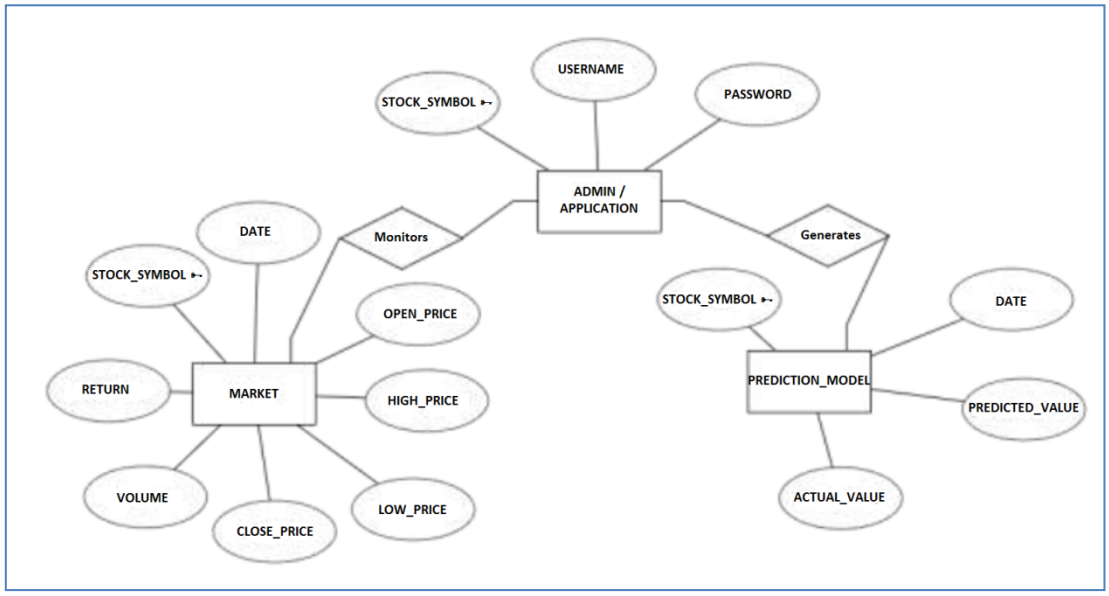


Level – 1 DFD (Regression Analysis):



Entity Relationship Diagram:

An entity–relationship model describes interrelated things of interest in a specific domain of knowledge. A basic ER model is composed of entity types and specifies relationships that can exist between instances of those entity types.



c) System Implementation:

i. Methodology adopted:

I. Automated Extraction of Latest Nifty stocks' Data using R:

Following are the Data extraction steps that would be followed to ensure that relevant data is available for performing Predictive Analysis. Data extraction can be done using Data scraping techniques in R:

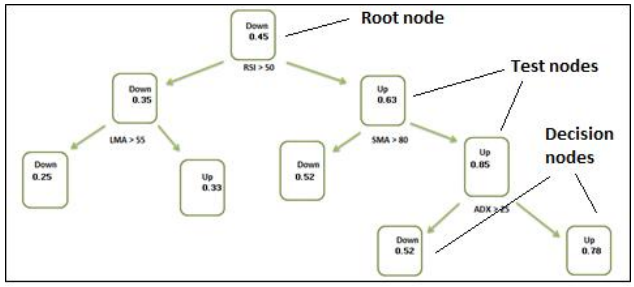
a. Download Latest Nifty 50 stocks List from Yahoo Finance

b. Download Last One Year of data effective till Today's date from Yahoo Finance for each Nifty 50 stock. Downloaded data will include the following information for each stock for last 1 year:

i. Date of Trading
ii. Open, High, Low, Close Price & Volume on the given date

II. Decision Tree Analysis of Nifty Stocks using R:

For the purpose of this project, we want to predict which way stocks will go using decision trees. We will download stock data from Yahoo Finance. We will add some technical indicators such as RSI (Relative Strength Index), SMA (Simple Moving Average), LMA (Linearly Weighted Moving Average) and ADX (Average Directional Index) to this dataset. Technical indicators will be calculated using basic stock values (OHLCV) in our case which will help us predict stock movements. Our machine learning algorithm will make use of the values from the technical indicators to make more accurate predictions. We will predict the daily change (Up/Down) using these technical indicators, by applying a machine learning decision tree algorithm on this dataset using R.

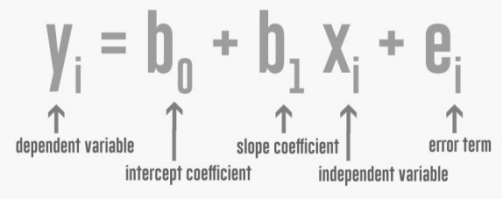


The root node is the main node in a decision tree. After the root node, each test node splits the data into further parts according to some set criteria. The data is split based on the indicators which are the test nodes.

The final node is the leaf node. In the decision tree diagram, the nodes containing the Up/Down class along with the probability value indicating the probability of the target attribute are the leaf nodes. The left of any node is considered as 'yes', meaning the question asked at the node (e.g. RSI >50 if answered yes, the algorithm navigates to left side of the tree and if not then right). So for example, this stock on a given day had its RSI at 39 then, so travel to the right of the tree check if the SMA is > 80, if yes then travel to the leaf which says, 'the stock has a 0.52 probability of going down'.

III. Linear Regression Analysis using Advanced Excel Tools and VBA:

Linear Regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X. Calculation is done based on the equation below, also called as regression expression:



For the purpose of this project, we would be predicting close price of stock on a daily basis, which would be our Dependent Variable and following would be the Dependent Variables:

a) Open, High, Low Price & Stock Volume
b) ADX (Average Directional Index)
c) SMA (Simple Moving Average)
d) LMA (Linearly Weighted Moving Average)
e) RSI (Relative Strength Index)

Independent Variables (X)

Dependent Variable (Y) => Close Price

IV. Technical Indicators Trend Analysis for Stock Prediction using D3.JS:

Technical indicators are focused on historical trading data, such as price, volume, and open interest, rather than the fundamentals of a business, like earnings, revenue, or profit margins.

There are two basic types of technical indicators:

- **Overlays** - Technical indicators that use the same scale as prices. Example: Moving averages.
- **Oscillators** - Technical indicators that oscillate between a local minimum and maximum are plotted above or below a price chart. Examples include the MACD or RSI.

For the purpose of this project, we will be viewing and analyzing the trend of the following technical indicators, to predict whether a stock is “Bullish” or “Bearish”:

a) **Average Directional Index (ADX):**

The Average Directional Index (**ADX**) represents a group of directional movement indicators that form a trading system. The **Plus Directional Indicator (+DI)** and **Minus Directional Indicator (-DI)** are derived from smoothed averages of these differences, and measure trend ***direction*** over time. The **Average Directional Index (ADX)** is in turn derived from the smoothed averages of the difference between +DI and -DI, and measures the ***strength*** of the trend (regardless of direction) over time. *In general, the bulls have the edge when +DI is greater than -DI, while the bears have the edge when -DI is greater.*

b) **Aroon Indicator:**

The **Aroon** indicator is a technical indicator that is used to identify when trends change direction. In essence, the indicator measures the time that it takes for the price to reach the highest and lowest points over a given timeframe as a percentage of total time.

c) **Average True Range (ATR):**

The **average true range indicator** is an oscillator, meaning the ATR will oscillate between peaks and valleys. The key ways through which the traders use the indicator are “Gauging a stock's volatility” and “Stop Loss/Exiting a Trade”.

d) **Heikin Ashi:**

Heikin-Ashi Candlesticks use the open-close data from the prior period and the open-high-low-close data from the current period to create a combo candlestick. Heikin-Ashi Candlesticks can be used to identify trending periods, potential reversal points and classic technical analysis patterns.

e) **Ichimoku Cloud:**

The **Ichimoku Cloud** defines support and resistance, identifies trend direction, gauges momentum and provides trading signals. The Ichimoku Cloud consists of five plots:

a. **Tenkan-sen (Conversion Line):** (9-period high + 9-period low)/2))
On a daily chart, this line is the midpoint of the 9-day high-low range.

b. **Kijun-sen (Base Line):** (26-period high + 26-period low)/2))
On a daily chart, this line is the midpoint of the 26-day high-low range.

c. **Senkou Span A (Leading Span A):** (Conversion Line + Base Line)/2))
This is the midpoint between the Conversion Line and the Base Line.

d. **Senkou Span B (Leading Span B):** (52-period high + 52-period low)/2))

On the daily chart, this line is the midpoint of the 52-day high-low range.

e. **Chikou Span (Lagging Span):** Close plotted 26 days in the past.

f) **Moving Average Convergence/Divergence oscillator (MACD):**

The **Moving Average Convergence/Divergence oscillator (MACD)** turns two trend-following indicators, moving averages, into a momentum oscillator by subtracting the longer moving average from the shorter moving average. **MACD** Line is the 12-day Exponential Moving Average (EMA) less the 26-day EMA. The MACD **histogram** is positive when the MACD Line is above its Signal line & negative when the MACD Line is below its Signal line. Convergence occurs when the moving averages move towards each other. Divergence occurs when the moving averages move away from each other.

g) **Relative Strength Index (RSI):**

The **Relative Strength Index (RSI)** is a momentum oscillator that measures the speed and change of price movements. The RSI oscillates between zero and 100.

h) **Stochastic Oscillator:**

The **stochastic oscillator** is a momentum indicator comparing the closing price of a security to the range of its prices over a certain period of time. For this indicator, in a market trending upward, prices close near the high, & in a market trending downward, prices close near the low.

i) **Williams %R:**

Williams %R is a momentum indicator that is the inverse of the Fast Stochastic Oscillator. Williams %R reflects the level of the close relative to the highest high for the look-back period. %R corrects for the inversion by multiplying the raw value by -100. As a result, the Fast Stochastic Oscillator and Williams %R produce the exact same lines, only the scaling is different.

V. Develop Front-end Dashboard for Visualization & Validation of Predictive Analysis in Excel:

a. Provide **User Friendly options and tools for the front end user** to select any of the Nifty stocks from a given list

b. Allow the Users to **Download latest data** for these stocks and View their overall analysis **on the click of a button**

c. Integrate Excel with R and D3.JS using VBA to project the output of Decision Tree and Technical Indicator charts on excel

d. Perform **Data Validation** of the Prediction models generated by Decision Tree and Linear Regression and validate the results to test the accuracy of the data

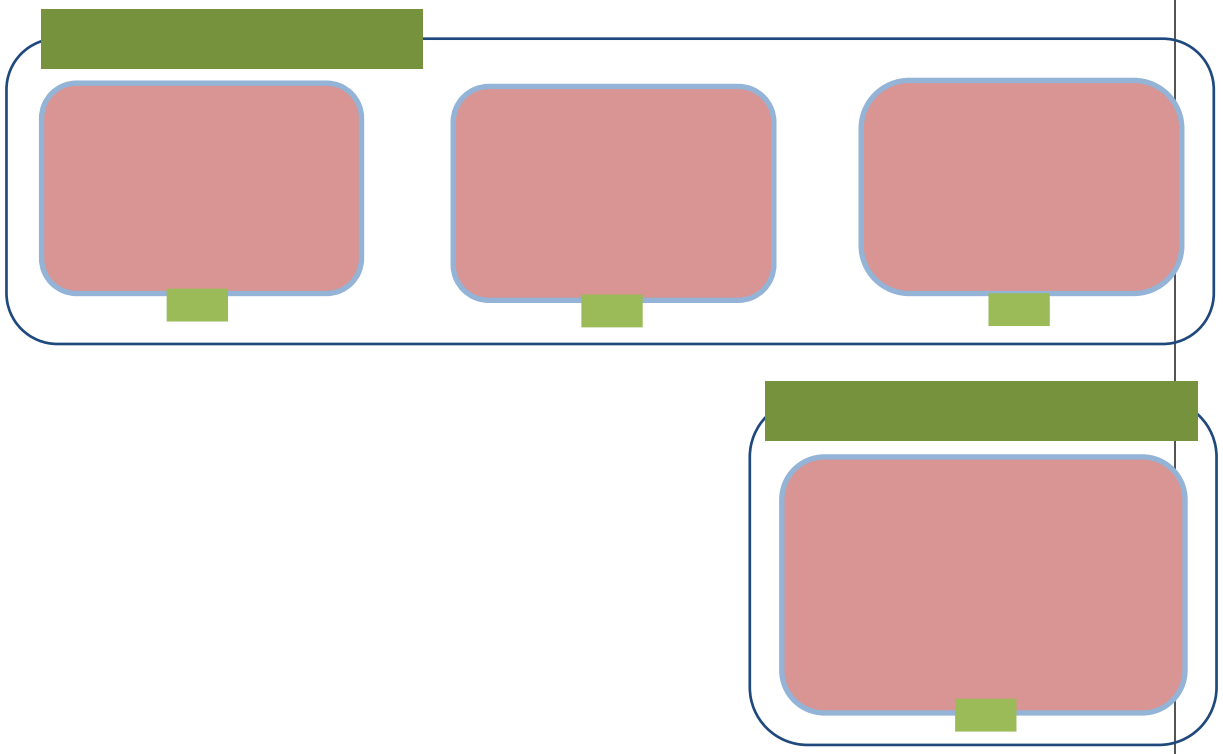
e. Perform a "**Bullish/Bearish**" **analysis** on each Prediction model & Technical Indicator Trends

f. Provide a Summary View to show the results of each indicator/model's results and give an overall picture of a particular stock, answering the following questions for the investor for a stock:

- Is the stock **Bullish or Bearish**? Or, in other words, is the stock ideal to be bought or to be sold on the next day?
- Is the stock a **stable or a risky stock**? Is the frequency of change in the upward and downward movement of the stocks high or low? This will help the conservative users to be aware of the risky stocks, and would also give an idea to the investors who are ready to take risks and make higher profits, as to which stock is better for them.

ii. Process Involved:

Process Implementation Flow Diagram:



*10**Data Validation will be performed simultaneously with each predictive model generation*

A. Stock Data Retrieval:

1. Web scraping code is written in R using the Rvest package to extract the NIFTY list and the OHLVC data for latest 12 months for a stock to be chosen and selected by the User in the Excel Dashboard.
2. This R code is integrated with VBA using Shell scripts in order to download the data on User's request.
3. In the downloaded stock data, the independent variables and the dependent variables are identified in order to develop the "Decision Tree" & "Regression" prediction models:

Variable	Decision Tree Model	Regression Model
Open Price	Independent	Independent
High Price	Independent	Independent
Low Price	Independent	Independent
Volume	Independent	Independent
RSI (Relative Strength Index)	Independent	Independent
SMA (Simple Moving Average)	Independent	Independent
LMA (Linearly Weighted Moving Average)	Independent	Independent
ADX (Average Directional Index)	Independent	Independent
Close Price	Independent	Dependent

Close Price Movement (Upward/Downward)	Dependent	-

B. Decision Tree Model:

4. Generate a Decision Tree in R with the help of the "Quantmod" and "Rpart" package, using the Independent and Dependent variables of the stock selected by the User.

C. Regression Analysis:

5. Generate a Regression Equation with the help of "Analysis Toolpak" in VBA using the Independent and Dependent variables of the stock selected by the User.

D. Technical Indicator Trend Charts:

6. Set up a Local IIS website for the project directory. Using IIS server, generate Technical indicator trend charts using D3.JS TechanJS library on an HTML local webpage.

E. Front End Excel Dashboard:

7. Use Phantom JS to connect R with VBA. VBA will send User's requested stock data to R and R will send the Decision Tree image to VBA

8. Use R to connect D3 with VBA. VBA will send User's requested stock data to D3.JS and D3.JS will send the Trend Chart images to VBA

9. Prepare an interactive Front-end Excel Dashboard in VBA to show the results of Decision Tree image/validation, Regression Analysis results, and Technical indicator trend chart images with interactive data hover tooltips. Provide interactive User-friendly drop down options and buttons in the dashboard so that the User can get the desired results.

F. Data Validation & Final Results: *Data Validation will be performed simultaneously with each predictive model generation.*

10. Validate the Prediction Models in VBA and Trend Chart observations and check the accuracy level by predicting historical data:

    a. **Decision Tree:** Validate by checking if the predicted movement is equal to the actual movement for last 12 months. Get the accuracy % as Number of days of accurate movement divided by Number of days of Total Movement. Ideally the accuracy % benchmark should be 65 to 70%.

    b. **Regression Analysis:** Check the error rate % of predicted close price vs actual close price for last 12 months and get the average error rate. The ideal average error rate should be less than 0.1%.

    c. **Technical Indicator Trend Charts:** Observe and Compare the Trend of Price/Price movement over 6 months and derive conclusions based on similar results from most of the charts. Also view the trend of stable/risky movement of price.

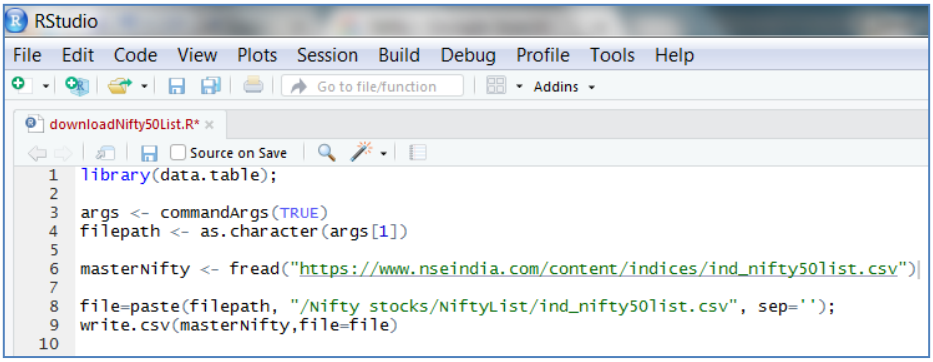
    Summarize the final results of the prediction.

- **Back-end System Implementation:**

1. **Data Sources:**

a. **NSE (National Stock Exchange):** The National Stock Exchange of India Limited is the leading stock exchange of India, located in Mumbai. The NIFTY 50 index is National Stock Exchange of India's benchmark broad based stock market index for the Indian equity market. The latest NIFTY stocks list is retrieved on User's request from NSE's website using R Web scraping. Following is the link to get the latest NIFTY list:
https://www.nseindia.com/content/indices/ind_nifty50list.csv

Below is a snapshot of the **R Code** to scrape and download NIFTY list from NSE's website:



Following is the Data Table generated from this data:

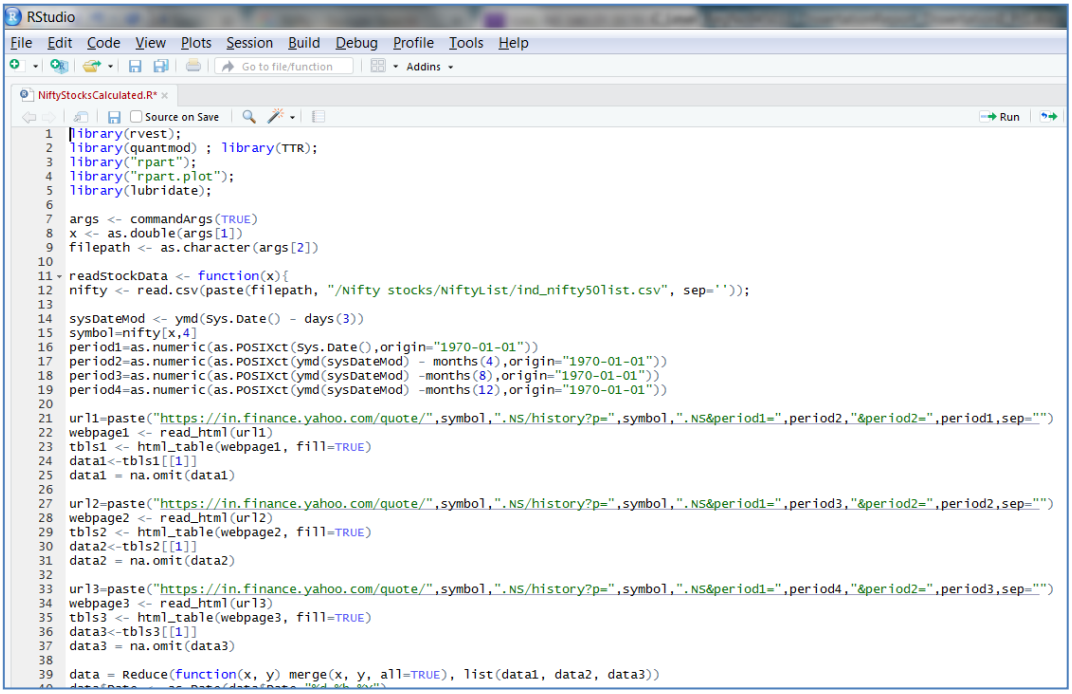
NIFTY_LIST
STOCK_SYMBOL
SERIAL_NUM
COMPANY_NAME
INDUSTRY
SERIES
ISIN_CODE

Following is a snapshot of the downloaded list:

	Company Name	Industry	Symbol	Series	ISIN Code
1	Adani Ports and Special Economic Zone Ltd.	SERVICES	ADANIPORTS	EQ	INE742F01042
2	Asian Paints Ltd.	CONSUMER GOODS	ASIANPAINT	EQ	INE021A01026
3	Axis Bank Ltd.	FINANCIAL SERVICES	AXISBANK	EQ	INE238A01034
4	Bajaj Auto Ltd.	AUTOMOBILE	BAJAJ-AUTO	EQ	INE917I01010
5	Bajaj Finance Ltd.	FINANCIAL SERVICES	BAJFINANCE	EQ	INE296A01024
6	Bajaj Finserv Ltd.	FINANCIAL SERVICES	BAJAJFINSV	EQ	INE918I01018
7	Bharat Petroleum Corporation Ltd.	ENERGY	BPCL	EQ	INE029A01011
8	Bharti Airtel Ltd.	TELECOM	BHARTIARTL	EQ	INE397D01024
9	Bharti Infratel Ltd.	TELECOM	INFRATEL	EQ	INE121J01017
10	Cipla Ltd.	PHARMA	CIPLA	EQ	INE059A01026
11	Coal India Ltd.	METALS	COALINDIA	EQ	INE522F01014
12	Dr. Reddy's Laboratories Ltd.	PHARMA	DRREDDY	EQ	INE089A01023
13	Eicher Motors Ltd.	AUTOMOBILE	EICHERMOT	EQ	INE066A01013
14	GAIL (India) Ltd.	ENERGY	GAIL	EQ	INE129A01019
15	Grasim Industries Ltd.	CEMENT & CEMENT PRODUCTS	GRASIM	EQ	INE047A01021
16	HCL Technologies Ltd.	IT	HCLTECH	EQ	INE860A01027
17	HDFC Bank Ltd.	FINANCIAL SERVICES	HDFCBANK	EQ	INE040A01026
18	Hero MotoCorp Ltd.	AUTOMOBILE	HEROMOTOCO	EQ	INE158A01026
19	Hindalco Industries Ltd.	METALS	HINDALCO	EQ	INE038A01020
20	Hindustan Petroleum Corporation Ltd.	ENERGY	HINDPETRO	EQ	INE094A01015
21	Hindustan Unilever Ltd.	CONSUMER GOODS	HINDUNILVR	EQ	INE030A01027
22	Housing Development Finance Corporation Ltd.	FINANCIAL SERVICES	HDFC	EQ	INE001A01036
23	I T C Ltd.	CONSUMER GOODS	ITC	EQ	INE154A01025
24	ICICI Bank Ltd.	FINANCIAL SERVICES	ICICIBANK	EQ	INE090A01021
25	Indiabulls Housing Finance Ltd.	FINANCIAL SERVICES	IBULHSGFIN	EQ	INE148I01020
26	Indian Oil Corporation Ltd.	ENERGY	IOC	EQ	INE242A01010
27	IndusInd Bank Ltd.	FINANCIAL SERVICES	INDUSINDBK	EQ	INE095A01012
28	Infosys Ltd.	IT	INFY	EQ	INE009A01021
29	Kotak Mahindra Bank Ltd.	FINANCIAL SERVICES	KOTAKBANK	EQ	INE237A01028
30	Larsen & Toubro Ltd.	CONSTRUCTION	LT	EQ	INE018A01030
31	Lupin Ltd.	PHARMA	LUPIN	EQ	INE326A01037
32	Mahindra & Mahindra Ltd.	AUTOMOBILE	M&M	EQ	INE101A01026
33	Maruti Suzuki India Ltd.	AUTOMOBILE	MARUTI	EQ	INE585B01010
34	NTPC Ltd.	ENERGY	NTPC	EQ	INE733E01010
35	Oil & Natural Gas Corporation Ltd.	ENERGY	ONGC	EQ	INE213A01029
36	Power Grid Corporation of India Ltd.	ENERGY	POWERGRID	EQ	INE752E01010
37	Reliance Industries Ltd.	ENERGY	RELIANCE	EQ	INE002A01018

b. **Yahoo! Finance:** Yahoo! Finance is a media property that is part of Yahoo!'s network. It provides financial news, data and commentary including stock quotes, press releases, financial reports, and original content. The NIFTY stocks list downloaded from NSE is provided to the end User. The end User chooses a stock from this list and requests for latest 12 months' OHLVC (Open, High, Low, Volume, Close Price) data for the chosen stock. This data is downloaded from Yahoo! Finance using R Web scraping. Below is an example of a link to download the stock: GAIL's data from Yahoo! Finance: https://in.finance.yahoo.com/quote/GAIL.NS/history?p=GAIL.NS

Below is a snapshot of the **R Code** to scrape and download stock OHLVC data from Yahoo! Finance:



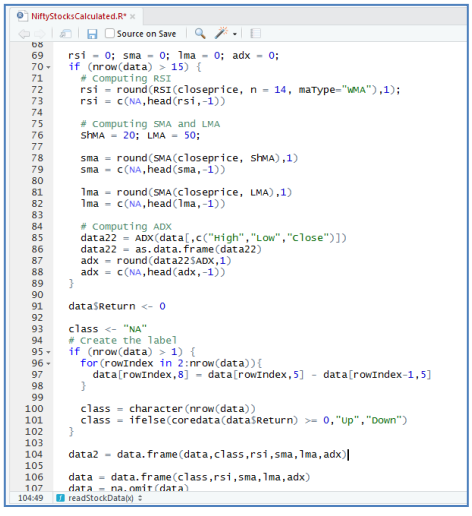
2. **Data Transformation:**

a. **R:**

R is a programming language and free software environment for statistical computing and graphics that is supported by the R Foundation for Statistical Computing.

i. **Stock Data Transformation:**

In the same code, calculate Return, Class (Price Movement – Up/Down), RSI, SMA, LMA and ADX, for each date and append to the existing dataset. Below is the code snapshot:



Following is the Data Table generated after download and transformation of the stock data:

OHLVC_DATA
STOCK_SYMBOL
DATE 🗝
OPEN_PRICE
HIGH_PRICE
LOW_PRICE
CLOSE_PRICE
ADJUSTED_CLOSE_PRICE
VOLUME
RETURN
CLASS
RSI
SMA
LMA
ADX

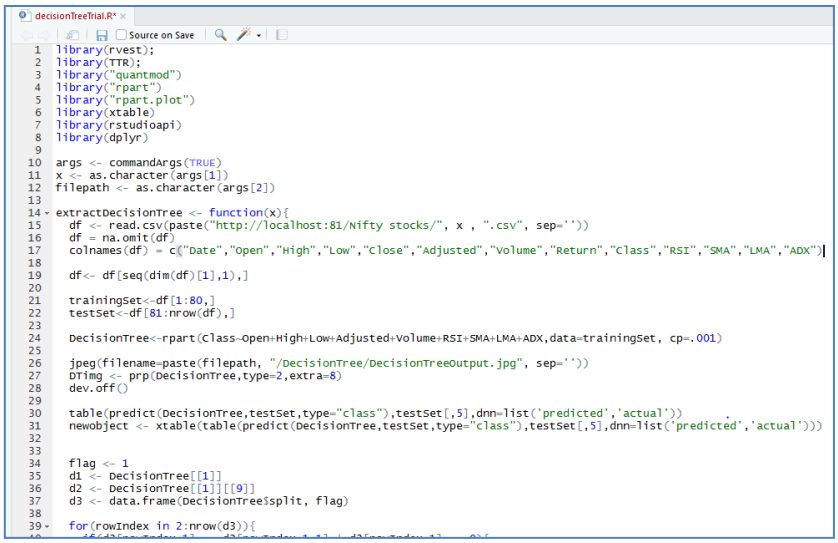
Following is a snapshot of the final stock dataset:

Date	Open	High	Low	Close	Adjusted	Volume	Return	class	rsi	sma	lma	adx
12-Jun-18	384	390.75	382.35	388.9	388.9	1920763	6.6	Up	55.1	383.5	388.1	22
11-Jun-18	384	387.7	380.1	382.3	382.3	2030226	1.8	Up	51.9	384.9	387.6	23.5
8-Jun-18	381	382.75	375.1	380.5	380.5	1232441	0.45	Up	50.3	386.5	387.3	24.4
7-Jun-18	378	384.8	376.55	380.05	380.05	1770108	4.25	Up	43.5	388	387	25.6
6-Jun-18	374	377.5	373.35	375.8	375.8	3927114	2.2	Up	39.4	389.9	386.7	25.8
5-Jun-18	378.75	379.35	371.5	373.6	373.6	2767220	-4.95	Down	42.5	391.9	386.4	26
4-Jun-18	388	388.95	376.05	378.55	378.55	2481566	-9.05	Down	50.4	393.6	386.1	26.7
1-Jun-18	394.1	395.4	382.3	387.6	387.6	3219572	-4.55	Down	53.8	394.6	385.7	28.1
31-May-18	375.5	397.6	369.25	392.15	392.15	11133305	17.75	Up	22.5	394.8	385.2	29.6
30-May-18	374	378.95	370	374.4	374.4	3110433	-3.45	Down	24.5	396.2	385.1	28.2
29-May-18	379.8	383.85	373.2	377.85	377.85	3394813	-1.05	Down	24.7	397.7	385.2	26.9
28-May-18	382.5	383.45	376.2	378.9	378.9	1807896	-1	Down	24.9	398.8	385.3	25.8
25-May-18	371.5	382	366.5	379.9	379.9	2372618	8.65	Up	3.6	399.7	385.4	24.3
24-May-18	374.95	378.6	364.7	371.25	371.25	3299522	-3	Down	7.3	400.7	385.7	22.1
23-May-18	379.8	383.1	373.5	374.25	374.25	1560605	-5.5	Down	11.7	401.6	385.7	20.3
22-May-18	382.7	389.05	377.75	379.75	379.75	1861819	-3.95	Down	15.6	402	385.9	18.6
21-May-18	389.35	392.55	381.5	383.7	383.7	1544488	-5.8	Down	21.7	401.8	385.8	17.2
18-May-18	396.15	397.1	384.05	389.5	389.5	3895361	-8.4	Down	32.9	401.7	386.1	15.8
17-May-18	404	405.5	395.6	397.9	397.9	1614257	-5.2	Down	44	401	386.2	15.6
16-May-18	407.75	409.75	400.25	403.1	403.1	1897907	-6.05	Down	59.2	400	386.3	16
15-May-18	399	412	397.4	409.15	409.15	3089340	-0.8	Down	62.8	399	386.3	16.4
14-May-18	413.95	414	407.4	409.95	409.95	5517825	-2.75	Down	71.3	397.6	386.3	16.9
11-May-18	412.35	414.5	410.85	412.7	412.7	1442596	1.9	Up	69.2	395.6	386.3	16.8
10-May-18	413.85	414.4	409.1	410.8	410.8	1552936	-2.55	Down	75.4	394	386.2	16.8
9-May-18	412.1	415.3	409.6	413.35	413.35	1408466	0.55	Up	74.8	392.7	386	16.6
8-May-18	413.1	415.65	410.1	412.8	412.8	1945992	0.4	Up	73.9	390.9	385.7	16.3
7-May-18	410	414.3	405.95	412.4	412.4	2582072	4.15	Up	70.7	389.2	385.3	16.2
4-May-18	401.95	409.6	391.3	408.25	408.25	4782981	11.25	Up	60.6	387.6	385.1	16.6
3-May-18	399.7	406.3	393.55	397	397	3161687	-4.1	Down	68.8	386.2	385.2	17.4
2-May-18	408	411.2	397.3	401.1	401.1	3047308	-6.2	Down	81	384.4	385.4	17.7
30-Apr-18	405.8	411.3	402.4	407.3	407.3	2264144	4.85	Up	76	382.5	385.6	17.3
27-Apr-18	398.3	405.9	398.3	402.45	402.45	8210865	5.35	Up	70.1	380.1	385.6	17.5
26-Apr-18	388.65	398.5	388.1	397.1	397.1	3663748	5.2	Up	64.2	378.5	385.6	18.5
25-Apr-18	394	397.45	390.2	391.9	391.9	3205946	-0.35	Down	65	377.2	385.9	19.3
24-Apr-18	387.25	394.5	387.1	392.25	392.25	3093161	5.9	Up	59.3	375.7	386.2	20.5
23-Apr-18	381.45	389.55	375.55	386.35	386.35	3614484	5.2	Up	53.5	374.3	386.5	21.7
20-Apr-18	386	387.25	378.95	381.15	381.15	2178241	-5.4	Down	63.7	373.4	387.1	23.4

ii. **Decision Tree Model:**

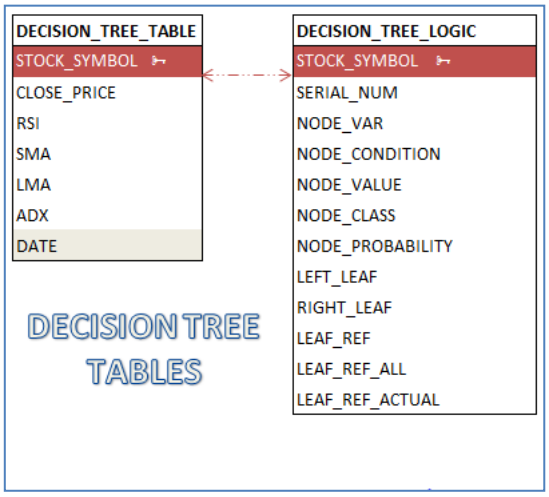
The R code developed for the Decision tree model uses packages like “Quantmod”, “RPart” , “RVest”, “DPlyr”, etc. and generates a decision tree based on the identified Independent and Dependent variables.

Following is a snapshot of the R Code:



```
library(rvest);
library(TTR);
library("quantmod")
library("rpart")
library("rpart.plot")
library(xtable)
library(rstudioapi)
library(dplyr)

args <- commandArgs(TRUE)
x <- as.character(args[1])
filepath <- as.character(args[2])

extractDecisionTree <- function(x){
  df <- read.csv(paste("http://localhost:81/Nifty stocks/", x , ".csv", sep=''))
  df = na.omit(df)
  colnames(df) = c("Date","Open","High","Low","Close","Adjusted","Volume","Return","Class","RSI","SMA","LMA","ADX")

  df<- df[seq(dim(df)[1],1),]

  trainingSet<-df[1:80,]
  testSet<-df[81:nrow(df),]

  DecisionTree<-rpart(Class~Open+High+Low+Adjusted+Volume+RSI+SMA+LMA+ADX,data=trainingSet, cp=.001)

  jpeg(filename=paste(filepath, "/DecisionTree/DecisionTreeOutput.jpg", sep=''))
  DTimg <- prp(DecisionTree,type=2,extra=8)
  dev.off()

  table(predict(DecisionTree,testSet,type="class"),testSet[,5],dnn=list('predicted','actual'))
  newobject <- xtable(table(predict(DecisionTree,testSet,type="class"),testSet[,5],dnn=list('predicted','actual')))


  flag <- 1
  d1 <- DecisionTree[[1]]
  d2 <- DecisionTree[[1]][[9]]
  d3 <- data.frame(DecisionTree$split, flag)

  for(rowIndex in 2:nrow(d3)){
```

Below are the data tables generated at this step:



Below is a snapshot of the related dataset generated in R:

Related Dataset:

	var	CD	index	UD	V
1	Volume	<	10627965	Down	0.5375
2	RSI	>=	52.3	Down	0.589041
3	Adjusted	<	341.765	Down	0.692308
4	<leaf>	NA	NA	Down	0.825
5	<leaf>	NA	NA	Up	0.75
6	Open	>=	339.58	Up	0.666667
7	<leaf>	NA	NA	Down	0.625
8	<leaf>	NA	NA	Up	0.846154
9	<leaf>	NA	NA	Up	1

iii. Integration of R and D3.JS using Phantom JS:

The technical indicator trend charts are generated on a local HTML webpage using D3.JS. In order to produce these charts on the dashboard, the D3.JS webpage is rendered in R using PhantomJs. Then, using the "Webshot" package in R, the chart images are exported and downloaded from the webpage.

Following are the code snapshots:

R code to render D3.JS webpage in R using PhantomJS:

getCsv.R*

Source on Save

```
library(rvest)

args <- commandArgs(TRUE)
x <- as.character(args[1])
filepath <- as.character(args[2])

getCSVForChart <- function(x){

  system(paste("phantomjs ", filepath, "/JS/getCsv.js ", x, " ", filepath, sep=''))

  scraping_wiki <- read_html("http://localhost:81/D3Codes/1.html")

  adx <- scraping_wiki %>%
    html_nodes("#csvAdx") %>% html_table()

  adxCsv <- adx[[1]]

  aroon <- scraping_wiki %>%
    html_nodes("#csvAroon") %>% html_table()

  aroonCsv <- aroon[[1]]

  atr <- scraping_wiki %>%
    html_nodes("#csvATR") %>% html_table()

  atrCsv <- atr[[1]]

  heiken <- scraping_wiki %>%
    html_nodes("#csvHeiken") %>% html_table()

  heikenCsv <- heiken[[1]]

  ichimoku <- scraping_wiki %>%
    html_nodes("#csvIchimoku") %>% html_table()

  ichimokuXAxis <- scraping_wiki %>%
    html_nodes(".load-ichimoku") %>% html_nodes("g") %>% html_nodes("g")

  ichimokuTick <- ichimokuXAxis[5] %>%
```

R Code to export D3 chart images using Webshot package:

phantom.R*

```r
library(webshot)

args <- commandArgs(TRUE)
x <- as.character(args[1])
filepath <- as.character(args[2])

extractD3Chart <- function(x){
  URL <- paste("http://localhost:81/D3Codes/index.html?",x,sep='')
  webshot(URL, paste(filepath, "/images/adx.jpg", sep=''), cliprect = c(0, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/aroon.jpg", sep=''), cliprect = c(500, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/atr.jpg", sep=''), cliprect = c(1000, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/heiken.jpg", sep=''), cliprect = c(1500, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/ichimoku.jpg", sep=''), cliprect = c(2005, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/macd.jpg", sep=''), cliprect = c(2510, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/rsi.jpg", sep=''), cliprect = c(3010, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/stochastic.jpg", sep=''), cliprect = c(3515, 5, 960, 510))
  webshot(URL, paste(filepath, "/images/williams.jpg", sep=''), cliprect = c(4015, 5, 960, 510))
  #webshot(URL, delay = 0.5)
}

extractD3Chart(x)
```

b. D3.JS:

D3.js is an **AngularJS** based JavaScript library for producing dynamic, interactive data visualizations in web browsers. It makes use of the widely implemented SVG, HTML5, and CSS standards. It is the successor to the earlier Protovis framework. **TechanJS** is a visual, stock charting (Candlestick, OHLC, indicators) and technical analysis library built on D3.

Using the OHLVC stock data, trend charts and related back-end datasets for last 6 months are developed on a local IIS server based HTML webpage.

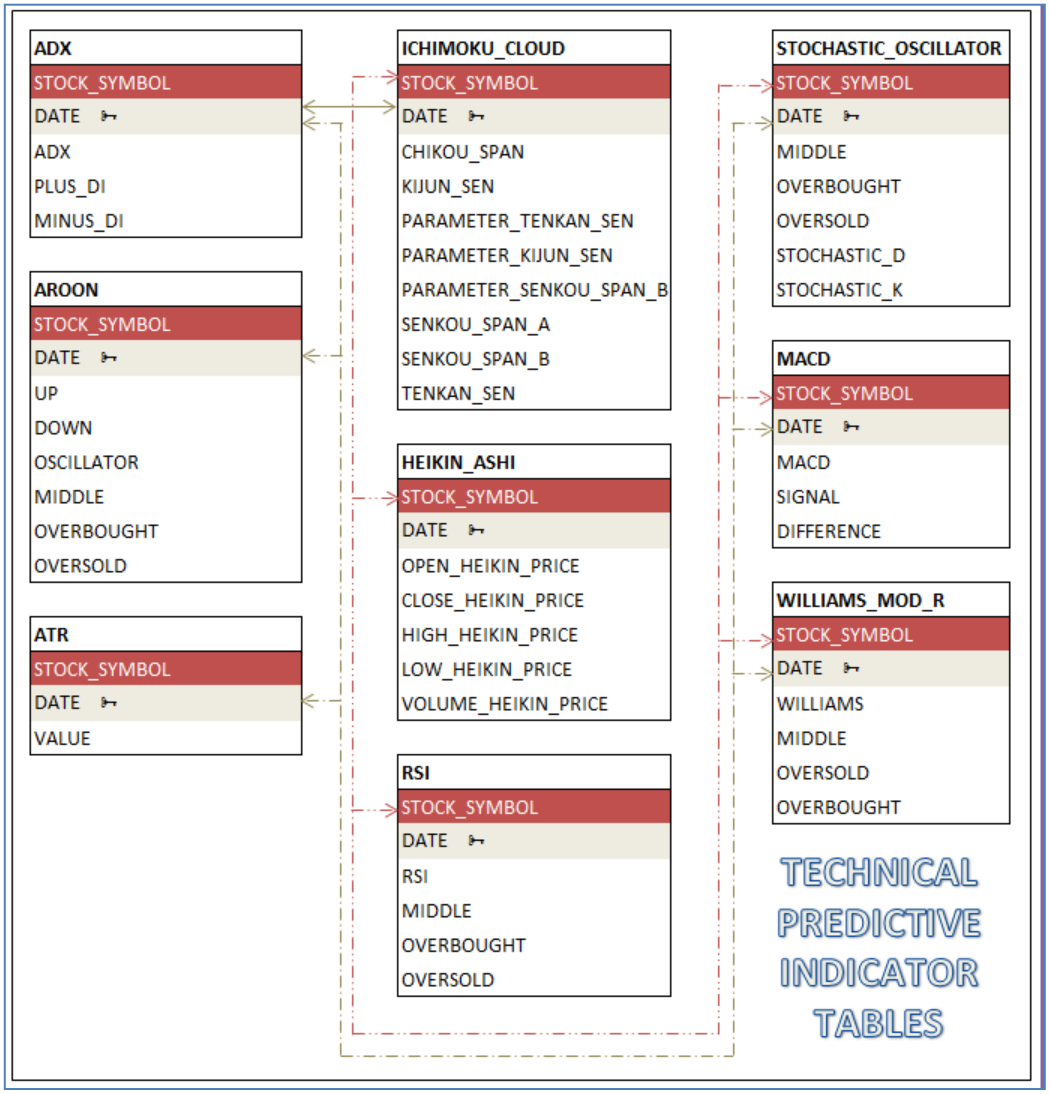
Following is a snapshot of the D3 Techan JS Code:

index.html - Notepad

File Edit Format View Help

```
<script src="/js/d3.v4.min.js"></script>
<script src="/js/techan.min.js"></script>
<script>
    var margin = {top: 20, right: 20, bottom: 30, left: 50},
            width = 960 - margin.left - margin.right,
            height = 500 - margin.top - margin.bottom;

    var parseDate = d3.timeParse("%d-%b-%y");

    var x = techan.scale.financetime()
            .range([0, width]);

    var y = d3.scaleLinear()
            .range([height, 0]);

    var adx = techan.plot.adx()
            .xScale(x)
            .yScale(y);

    var aroon = techan.plot.aroon()
            .xScale(x)
            .yScale(y);
    var accessor = aroon.accessor();

    var candlestick = techan.plot.candlestick()
            .xScale(x)
            .yScale(y);

    var atrtrailingstop = techan.plot.atrtrailingstop()
            .xScale(x)
            .yScale(y);

    var ichimoku = techan.plot.ichimoku()
            .xScale(x)
            .yScale(y);

    var macd = techan.plot.macd()
            .xScale(x)
            .yScale(y);

    var rsi = techan.plot.rsi()
            .xScale(x)
            .yScale(y);

    var stochastic = techan.plot.stochastic()
            .xScale(x)
            .yScale(y);


    var heikinashi = techan.plot.heikinashi()
            .xScale(x)
            .yScale(y);

    var heikinashiIndicator = techan.indicator.heikinashi();

    var williams = techan.plot.williams()
            .xScale(x)
```

Following are the data tables generated related to the charts:



c. VBA (Excel):

VBA stands for Visual Basic for Applications. It is a combination of the Microsoft's event-driven programming language Visual Basic with Microsoft Office Applications such as Microsoft Excel.

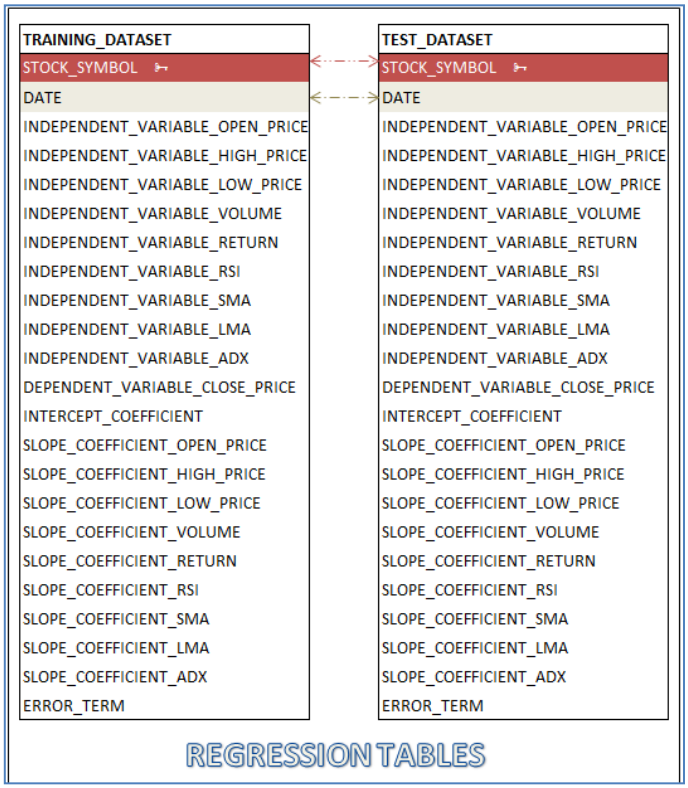
i. **Regression Analysis:**

Regression Analysis has been done using the "Analysis Toolpak – VBA" add-in in Excel as per the identified independent and dependent variables.

Following is the VBA code snapshot for generating Regression analysis model and performing model validation:

```
Sub regressionAnalysis()
    Application.ScreenUpdating = False
    Application.DisplayAlerts = False

    ThisWorkbook.Worksheets("Regression Analysis").Select
    ThisWorkbook.Worksheets("Regression Analysis").Range("B6:L6").Select
    ThisWorkbook.Worksheets("Regression Analysis").Range(Selection, Selection.End(xlDown)).ClearContents
    ThisWorkbook.Worksheets("Regression Analysis").Range("N6:X6").Select
    ThisWorkbook.Worksheets("Regression Analysis").Range(Selection, Selection.End(xlDown)).ClearContents

    Set wb = Workbooks.Open(ThisWorkbook.path & "\Nifty stocks\" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisWorkbook.Worksheets("Nifty Symb
    Range("A2").AutoFilter Field:=10, Criteria1:="<>NA"
    Range("A2").AutoFilter Field:=11, Criteria1:="<>NA"
    Range("A2").AutoFilter Field:=12, Criteria1:="<>NA"
    Range("A2").AutoFilter Field:=13, Criteria1:="<>NA"
    Range("A2:D2").Select
    Range(Selection, Selection.End(xlDown)).Copy
    ThisWorkbook.Activate
    ThisWorkbook.Worksheets("Regression Analysis").Range("B6").PasteSpecial xlPasteValues
    wb.Activate
    Range("G2:H2").Select
    Range(Selection, Selection.End(xlDown)).Copy
    ThisWorkbook.Activate
    ThisWorkbook.Worksheets("Regression Analysis").Range("F6").PasteSpecial xlPasteValues
    wb.Activate
    Range("J2:M2").Select
    Range(Selection, Selection.End(xlDown)).Copy
    ThisWorkbook.Activate
    ThisWorkbook.Worksheets("Regression Analysis").Range("H6").PasteSpecial xlPasteValues
    wb.Activate
    Range("F2").Select
    Range(Selection, Selection.End(xlDown)).Copy
    ThisWorkbook.Activate
    ThisWorkbook.Worksheets("Regression Analysis").Range("L6").PasteSpecial xlPasteValues
    wb.Close

    ThisWorkbook.Worksheets("Regression Analysis").Range("B5:L" & ThisWorkbook.Worksheets("Regression Analysis").Range("C5").End(xlDown).Row).Sort Key1:=Th
    ThisWorkbook.Worksheets("Regression Analysis").Range("B124:L" & ThisWorkbook.Worksheets("Regression Analysis").Range("C5").End(xlDown).Row).Copy
    ThisWorkbook.Worksheets("Regression Analysis").Range("N6").PasteSpecial xlPasteValues
    ThisWorkbook.Worksheets("Regression Analysis").Range("B124:L" & ThisWorkbook.Worksheets("Regression Analysis").Range("C5").End(xlDown).Row).ClearConten

    ThisWorkbook.Worksheets("Regression Analysis").Range("AB8:AJ33").ClearContents

    Application.Run "ATPVBAEN.XLAM!Regress", ThisWorkbook.Worksheets("Regression Analysis").Range("$L$5:$L$123"), _
```

Following are the Data Tables generated for Regression Analysis:



ii. **Decision Tree Model Validation:**

Below is a snapshot of the VBA code which is used in the Dashboard to validate the Decision Tree model:

```
Sub getDecisionTreeLogic()
    Application.ScreenUpdating = False
    Application.DisplayAlerts = False

    'DT sheet
    ThisWorkbook.Activate
    Sheets("DT").Visible = True
    ThisWorkbook.Worksheets("DT").Cells.ClearContents

    Set wb = Workbooks.Open(ThisWorkbook.path & "\DecisionTree\DecisionTreeOutput_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisWorkbook.Work
    Range("A:F").Copy

    ThisWorkbook.Activate
    Sheets("DT").Select
    ThisWorkbook.Worksheets("DT").Range("A:F").PasteSpecial xlPasteValues

    wb.Close

    ThisWorkbook.Worksheets("DT").Range("G1").Value = "Left Leaf"
    ThisWorkbook.Worksheets("DT").Range("H1").Value = "Right Leaf"
    ThisWorkbook.Worksheets("DT").Range("I1").Value = "Leaf Ref"
    ThisWorkbook.Worksheets("DT").Range("J1").Value = "Leaf Ref All"
    ThisWorkbook.Worksheets("DT").Range("K1").Value = "Leaf Ref Actual"

    ThisWorkbook.Worksheets("DT").Range("G2").Formula = "=IF(C2=""NA"","""",A3)"
    ThisWorkbook.Worksheets("DT").Range("H2").Formula = "=IF(C2=""NA"","""",IF(B3=""<leaf>"",A4,MAX(H3:INDIRECT(ADDRESS(K2,8)))+1))"
    ThisWorkbook.Worksheets("DT").Range("I2").Formula = "=IF(C2=""NA"",IF(AND(C2=""NA"",C3<>""NA""),ROW(),""""),"""")"
    ThisWorkbook.Worksheets("DT").Range("J2").Formula = "=MIN(I2:$I$" & Range("F2").End(xlDown).Row & ")"

    ThisWorkbook.Worksheets("DT").Range("G2:J2").Copy
    ThisWorkbook.Worksheets("DT").Range("G3:J" & Range("F2").End(xlDown).Row).PasteSpecial xlPasteAll

    formulaLeaf
    ThisWorkbook.Worksheets("DT").Range("K2").Copy
    ThisWorkbook.Worksheets("DT").Range("K3:K" & Range("F2").End(xlDown).Row).PasteSpecial xlPasteAll

    ThisWorkbook.Activate
    Sheets("DT").Visible = xlVeryHidden

    'DTValues sheet
    ThisWorkbook.Activate
    Sheets("DTValues").Visible = True

    Set wb1 = Workbooks.Open(ThisWorkbook.path & "\Nifty stocks\" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisWorkbook.Worksheets("Nifty Symb
```

iii. **Integrating R codes with VBA:**

The Shell scripts are used to call R scripts from VBA (including the scripts that use Phantom JS to integrate D3.JS with R). The Shell script send Stock Symbol chosen by the user to R and R accordingly retrieves and transforms the data for that stock. Following is a sample of Shell scripts VBA snapshot:

```
Sub runRScript()
    Application.ScreenUpdating = False

    Dim shell As Object
    Set shell = VBA.CreateObject("WScript.Shell")
    Dim waitTillComplete As Boolean: waitTillComplete = True
    Dim style As Integer: style = 0
    Dim errorCode As Integer
    Dim path As String

    path = "RScript " & ThisWorkbook.path & "\RCodes\NiftyStocksCalculated.R " & ThisWorkbook.Worksheets("Nifty Symbols").Range("D2").Value & " " & ThisWor

    errorCode = shell.Run(path, style, waitTillComplete)

    Application.ScreenUpdating = True
End Sub

Sub runPhantomScript()
    Application.ScreenUpdating = False

    Dim shell As Object
    Set shell = VBA.CreateObject("WScript.Shell")
    Dim waitTillComplete As Boolean: waitTillComplete = True
    Dim style As Integer: style = 0
    Dim errorCode As Integer
    Dim path As String

    path = "RScript " & ThisWorkbook.path & "\RCodes\phantom.R " & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisWorkbook.Worksheets("Nifty Symb
    errorCode = shell.Run(path, style, waitTillComplete)

    Application.ScreenUpdating = True
End Sub

Sub runCSVScript()
    Application.ScreenUpdating = False

    Dim shell As Object
    Set shell = VBA.CreateObject("WScript.Shell")
    Dim waitTillComplete As Boolean: waitTillComplete = True
    Dim style As Integer: style = 0
    Dim errorCode As Integer
    Dim path As String

    path = "RScript " & ThisWorkbook.path & "\RCodes\getCsv.R " & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisWorkbook.Worksheets("Nifty Symbo
```

iv. Copying Trend Charts and Decision Tree images on the Excel Dashboard:

Below is the code to copy the exported D3 Trend Charts and R Decision Tree images on the Excel Dashboard:

```
Sub refreshOnClick()
    Application.ScreenUpdating = False

    Dim sheetName As String
    sheetName = ActiveSheet.Name

    Dim FSO
    Set FSO = CreateObject("Scripting.FileSystemObject")

    ThisWorkbook.Worksheets("Predictive Indicator").Range("T1").Value = 1
    UserForm1.Show (vbModeless)
    UserForm1.Repaint
    UserForm1.Label1.Caption = "Generating latest data from Yahoo Finance. Please hold on for few minutes...."
        runRScript
    UserForm1.Repaint
    UserForm1.Label1.Caption = "Generating Predictive Indicators. Please hold on for few minutes...."
        runPhantomScript
    UserForm1.Repaint
    UserForm1.Label1.Caption = "Generating Analytic Trends. Please hold on for few minutes...."
        runCSVScript

        FSO.copyfile ThisWorkbook.path & "\images\adx.jpg", ThisWorkbook.path & "\images\adx_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisW
        FSO.copyfile ThisWorkbook.path & "\images\aroon.jpg", ThisWorkbook.path & "\images\aroon_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & T
        FSO.copyfile ThisWorkbook.path & "\images\atr.jpg", ThisWorkbook.path & "\images\atr_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisW
        FSO.copyfile ThisWorkbook.path & "\images\heiken.jpg", ThisWorkbook.path & "\images\heiken_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" &
        FSO.copyfile ThisWorkbook.path & "\images\ichimoku.jpg", ThisWorkbook.path & "\images\ichimoku_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("
        FSO.copyfile ThisWorkbook.path & "\images\macd.jpg", ThisWorkbook.path & "\images\macd_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & Thi
        FSO.copyfile ThisWorkbook.path & "\images\rsi.jpg", ThisWorkbook.path & "\images\rsi_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("A" & ThisW
        FSO.copyfile ThisWorkbook.path & "\images\stochastic.jpg", ThisWorkbook.path & "\images\stochastic_" & ThisWorkbook.Worksheets("Nifty Symbols").Ran
        FSO.copyfile ThisWorkbook.path & "\images\williams.jpg", ThisWorkbook.path & "\images\williams_" & ThisWorkbook.Worksheets("Nifty Symbols").Range("

    UserForm1.Repaint
    UserForm1.Label1.Caption = "Generating Decision Tree...."
        runDecisionTreeScript
        FSO.copyfile ThisWorkbook.path & "\DecisionTree\DecisionTreeOutput.jpg", ThisWorkbook.path & "\DecisionTree\DecisionTreeOutput_" & ThisWorkbook.Wor

    copyCSVFiles

    refreshChart

    ActiveWorkbook.Worksheets("Decision Tree").Image1.Picture = LoadPicture(ThisWorkbook.path & "\DecisionTree\DecisionTreeOutput_" & ThisWorkbook.Workshee
    getDecisionTreeLogic
    regressionAnalysis
    DecisionTreeValidation
```

v. D3 Trend Charts Hover Data Tooltips:

The D3 Trend Charts exported and copied on the dashboard are just images. To make these charts interactive, the data labels related to each point in the chart are displayed as a tooltip. Below is the code to generate the hover tooltips:

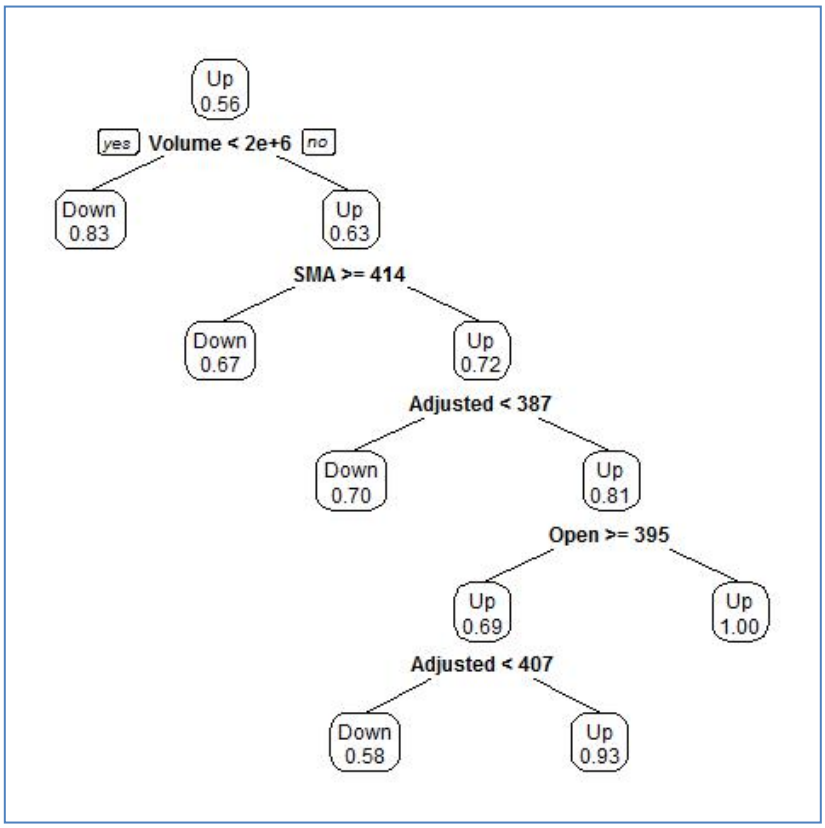
```
Declare Function GetCursorPos Lib "user32" (lpPoint As POINTAPI) As Long
Type POINTAPI
    x As Long
    y As Long
End Type

Function CurosrXY_Pixels(numRow, chartName)
    Dim lngStatus As Long
    Dim typWhere As POINTAPI

    lngStatus = GetCursorPos(typWhere)

    Dim partSlice As Double

    'ThisWorkbook.Worksheets("Predictive Indicator").Range("I2").Value = typWhere.x

    partSlice = ThisWorkbook.Worksheets("Predictive Indicator").Image1.Width / numRow

    ThisWorkbook.Worksheets("Predictive Indicator").TextBox1.Left = typWhere.x / partSlice

    Dim firstHeading, secondHeading, thirdHeading, fourthHeading, fifthHeading, sixthHeading, seventhHeading, eighthHeading, ninthHeading As String
    Dim firstVal, secondVal, thirdVal, fourthVal, fifthVal, sixthVal, seventhVal, eighthVal, ninthVal As String

    firstHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("A1").Value), "_", " ")
    secondHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("B1").Value), "_", " ")
    thirdHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("C1").Value), "_", " ")
    fourthHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("D1").Value), "_", " ")
    fifthHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("E1").Value), "_", " ")
    sixthHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("F1").Value), "_", " ")
    seventhHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("G1").Value), "_", " ")
    eighthHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("H1").Value), "_", " ")
    ninthHeading = Replace(UCase(ThisWorkbook.Worksheets(chartName).Range("I1").Value), "_", " ")

    If (typWhere.x < 95 Or typWhere.x > 1017) Then
        ThisWorkbook.Worksheets("Predictive Indicator").TextBox1.Visible = False
    Else
        firstVal = Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.WorksheetFunction.Match(typWhere.x, T
        secondVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.W
        thirdVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.Wo
        fourthVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.W
        fifthVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.Wo
        sixthVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.Wo
        seventhVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.
        eighthVal = Application.WorksheetFunction.Text(Application.WorksheetFunction.Index(ThisWorkbook.Worksheets(chartName).Range("$A:$I"), Application.W
```

- **Frontend System Implementation:**

1. **Decision Tree Model Design:**

The R code developed for the Decision tree model uses packages like "Quantmod", "RPart" , "RVest", "DPlyr", etc. and generates a decision tree based on the identified Independent and Dependent variables.

Following is a snapshot of a sample Decision Tree generated in R for the stock ***"ADANIPORTS"***:



The logic of the tree goes like this:

IF Volume of stock < 2 million Then
There is an 83% probability that the price of the stock will go Down
Else
If Simple Moving Average (SMA) of stock >= 414 Then
There is a 67% probability that the price of the stock will go Down
Else
If Adjusted Close Price of stock < 387 Then
There is a 70% probability that the price of stock will go Down
Else
If Open Price of stock >= 395 Then
If Adjusted Close Price of stock < 407 Then
58% probability that price will go Down
Else
93% probability that price will go Up
Else
There is a 100% probability that price will go Up

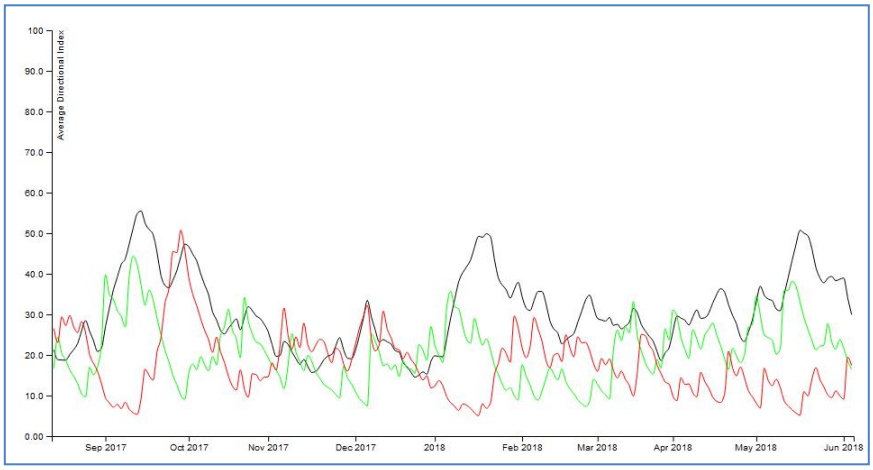
2. **Predictive Technical Indicators Trend Charts Design:**

D3.js is an **AngularJS** based JavaScript library for producing dynamic, interactive data visualizations in web browsers. It makes use of the widely implemented SVG, HTML5, and CSS standards. It is the successor to the earlier Protovis framework. **TechanJS** is a visual, stock charting (Candlestick, OHLC, indicators) and technical analysis library built on D3.

Using the OHLVC stock data, trend charts and related back-end datasets for last 6 months are developed on a local IIS server based HTML webpage.

Following are sample snapshots of the D3 charts generated for different Technical Predictive Indicators for different stocks:

Sample Average Directional Index (ADX) trend chart for the stock "ASIANPAINT":

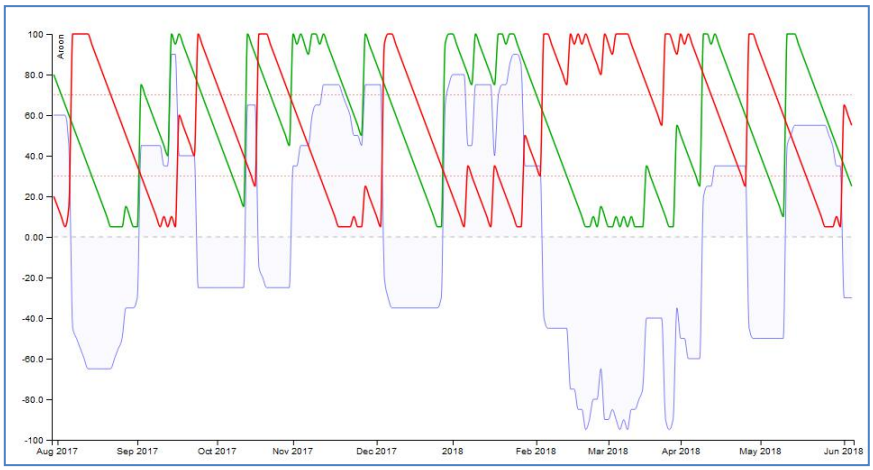


The green represents the Positive Directional Index (DI). The red line represents the Negative Directional Index (DI). The black line is the Average Directional Index (ADX), which is derived from smoothed average of the difference between +DI and –DI.

Average Postive DI for this stock:	20.2963
Average Negative DI for this stock:	17.8993
Difference:	*2.3970*

Bullish Trend: *Ideal to buy the stock now as stock price will increase in future*

Sample Aroon trend chart for the stock "AXISBANK":



The green line measures the strength of the uptrend. The red line measures the strength of the downtrend.

Average Postive Aroon for this stock:	54.7619
Average Negative Aroon for this stock:	56.9264
Difference:	*-2.1645*

The Aroon Up is low, suggesting that the stock has a Bearish Bias:
Ideal to sell the stock now for more profits as price will decrease in future

Sample ATR (Average True Range) trend chart for the stock "BAJAJ-AUTO":

The average true range indicator is a volatility measure of a stock's performance. Below are the key ways traders use the indicator:

1. *Gauging a stock's volatility*
2. *Stop Loss/Exiting a Trade*

Gauging a stock's volatility:

Average ATR:	63.0159
Average Close Price of stock:	3006.5639
Volatility Ratio:	*0.0210*

Stable Stock: *Stock does not have many ups and downs. Can trade if you want to be more conservative.*

Stop Loss/Exiting Trade (Assuming multiplier of 3x for target price & 1x for ATR stop):

Average ATR:	63.0159
Average Close Price of stock:	3006.5639
Target Price:	*$ 3,195.61*
Stop Loss:	*$ 2,943.55*

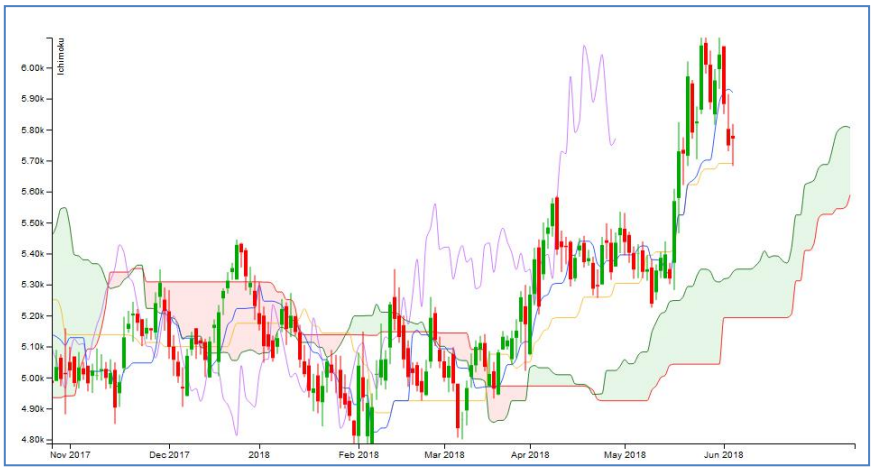
> ***Note:*** *Applying a multiplier to the average true range during a dull trading period will limit you in the potential gains as your profit targets are relative to the most recent trading volatility.*

Sample Heikin Ashi trend chart for the stock "BAJFINANCE":

**** Green sticks indicate strong advance - buying pressure***
**** Red sticks indicate strong decline - selling pressure***

Looking at this chart, Strong Advances are ***greater than*** *Strong Declines*
Ideal to buy stock now

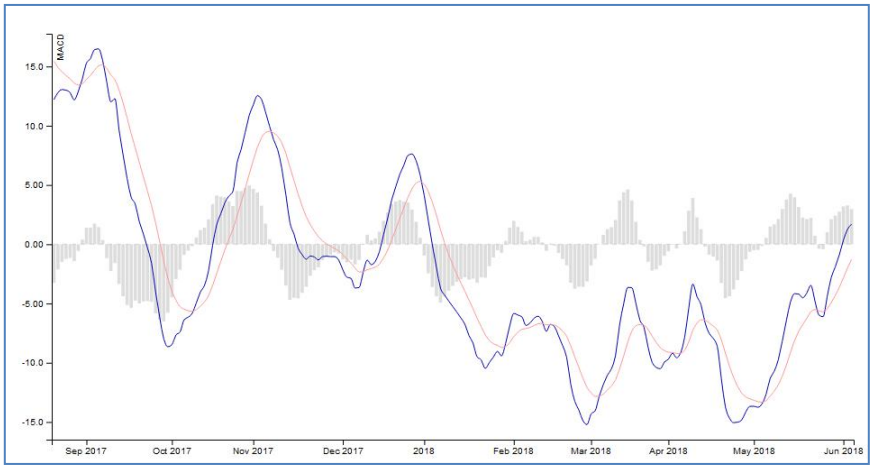
Sample Ichimoku Cloud trend chart for the stock "BAJAJFINSV":



There are two ways to identify the overall trend using the Cloud. **First, the trend is up when prices are above the Cloud, down when prices are below the Cloud and flat when prices are in the Cloud.** Second, the uptrend is strengthened when the Leading Span A (green cloud line) is rising and above the Leading Span B (red cloud line). This situation produces a ***green Cloud***. Conversely, a downtrend is reinforced when the Leading Span A (green cloud line) is falling and below the Leading Span B (red cloud line). This situation produces a ***red Cloud***. Because the Cloud is shifted forward 26 days, it also provides a glimpse of future support or resistance.

Chikou Span:	5302.7245	(Average value less than most of the values)	Bearish Trend
Kijun Sen:	5251.8983	(Average value less than most of the values)	Bearish Trend
Senkou Span A:	5264.9066	(Average value less than most of the values)	Bearish Trend
Senkou Span B:	5218.7250	(Average value less than most of the values)	Bearish Trend
Tenkan Sen:	5277.9150	(Average value less than most of the values)	Bearish Trend
Ideal to sell the stock before further price downfall			

Sample MACD (Moving Average Convergence/Divergence Oscillator) trend chart for the stock "BPCL":



MACD is about the convergence and divergence of the two moving averages. Convergence occurs when the moving averages move towards each other. Divergence occurs when the moving averages move away from each other.

1. Signal Line Crossovers:

The signal line is a 9-day EMA of the MACD Line. As a moving average of the indicator, it trails the MACD and makes it easier to spot MACD turns. A bullish crossover occurs when the MACD turns up and crosses above the signal line. A bearish crossover occurs when the MACD turns down and crosses below the signal line. Crossovers can last a few days or a few weeks, it all depends on the strength of the move.

Number of occurrences with MACD above the signal line:	101
Number of occurrences with MACD below the signal line:	116
More occurrences of Bearish crossover	

2. Centerline Crossovers:

Centerline crossovers are the next most common MACD signals. A bullish centerline crossover occurs when the MACD Line moves above the zero line to turn positive. This happens when the 12-day EMA of the underlying security moves above the 26-day EMA. A bearish centerline crossover occurs when the MACD moves below the zero line to turn negative. This happens when the 12-day EMA moves below the 26-day EMA.

Number of occurrences with MACD above the zero line:	74
Number of occurrences with MACD below the zero line:	143
More occurrences of Bearish crossover	

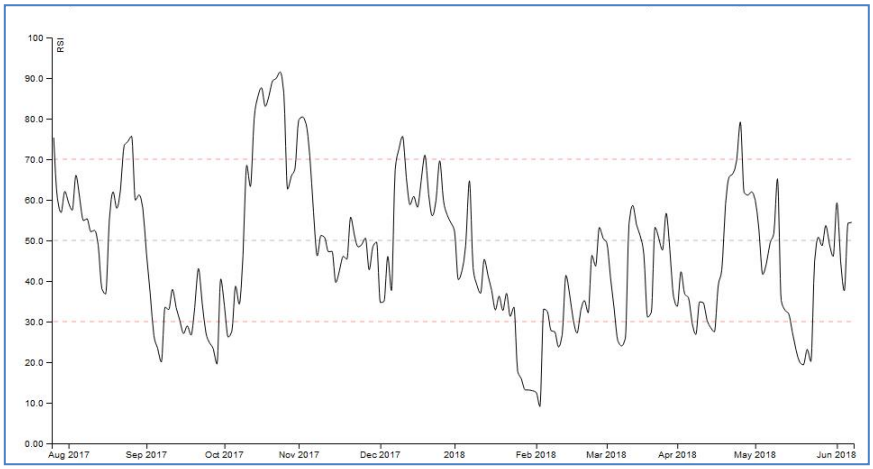
3. Divergences:

A situation that occurs when two lines on a chart move in opposite directions vertically. There are two kinds of divergences: positive and negative. A positive divergence occurs when the indicator moves higher while the stock is declining. A negative divergence occurs when the indicator moves lower while the stock is rising.

Positive Divergence:	54	**Positive Convergence:**	54	More occurrences of Bullish Divergence
Negative Divergence:	54	**Negative Convergence:**	54	

Ideal to sell the stock

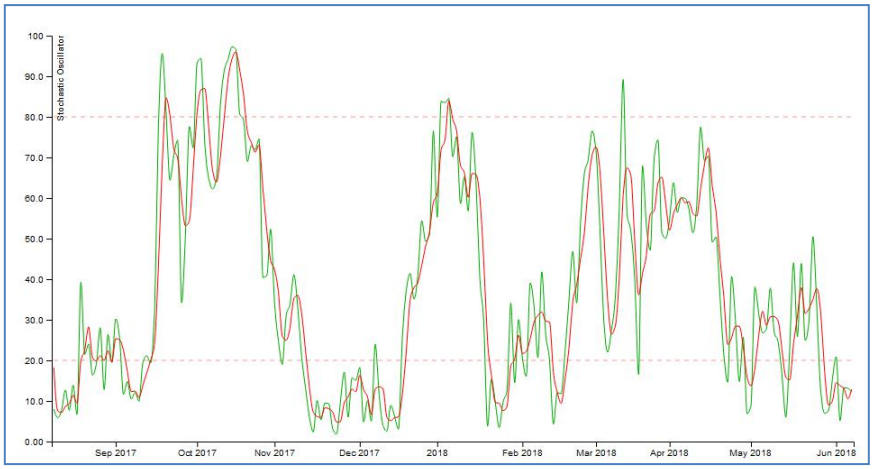
Sample RSI (Relative Strength Index) trend chart for the stock "BHARTIARTL":



RSI - considered overbought when above 70, oversold when below 30 and stable when around 50.

Average RSI for last one year as per current time period: 46.8983
This stock is stable, neither overbought nor oversold
Don't buy/sell the stock now! Ideal to wait and observe for further price increase/decrease
Note: Stock does not have many ups and downs. Can trade if you want to be more conservative!

Sample Stochastic Oscillator trend chart for the stock "INFRATEL":



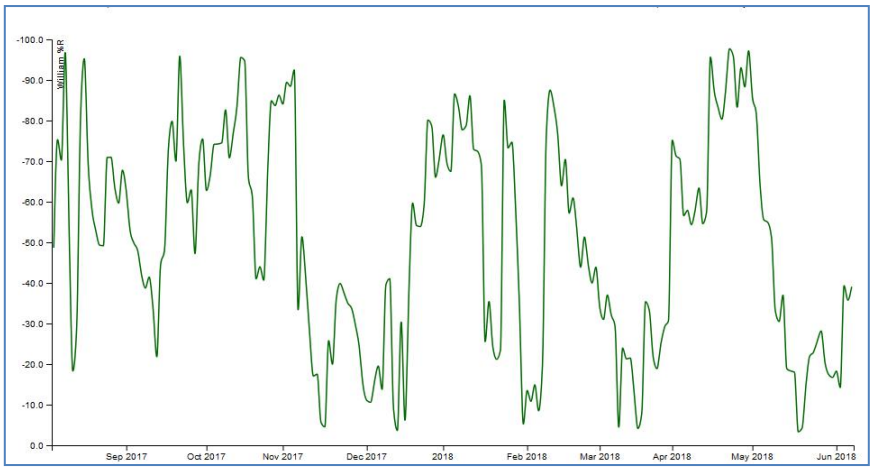
The stochastic oscillator is calculated using the following formula:

%K = 100(C - L14)/(H14 - L14) *(Green line in the chart)*
Where:
C = the most recent closing price
L14 = the low of the 14 previous trading sessions
H14 = the highest price traded during the same 14-day period
%K= the current market rate for the currency pair *(Green line in the chart)*
%D = 3-period moving average of %K *(Red line in the chart)*

Considering the most traditional settings for the oscillator, 20 is typically considered the oversold threshold and 80 is considered the overbought threshold. Readings above 80 indicate a security is trading near the top of its high-low range; readings below 20 indicate the security is trading near the bottom of its high-low range.

Average Stochastic "K" for last one year as per current time period: 37.6980
Average Stochastic "D" for last one year as per current time period: 37.4156
This stock is stable, neither overbought nor oversold
Don't buy/sell the stock now! Ideal to wait and observe for further price increase/decrease
Note: Stock does not have many ups and downs. Can trade if you want to be more conservative!

Sample Williams %R trend chart for the stock "CIPLA":



Williams %R oscillates from 0 to -100. Readings from 0 to -20 are considered overbought. Readings from -80 to -100 are considered oversold.

%R = (Highest High - Close)/(Highest High - Lowest Low) * -100
Lowest Low = lowest low for the look-back period
Highest High = highest high for the look-back period
%R is multiplied by -100 correct the inversion and move the decimal.

Average Williams % R for last one year as per current time period:	-53.2946
This stock is stable, neither overbought nor oversold	
Don't buy/sell the stock now! Ideal to wait and observe for further price increase/decrease	
Note: Stock does not have many ups and downs. Can trade if you want to be more conservative!	

3. **Regression Analysis Model Design:**

Regression Analysis has been done using the "Analysis Toolpak – VBA" add-in in Excel as per the identified independent and dependent variables.

Following is a sample Regression Analysis model Results snapshot for the stock: **"COALINDIA" –**

COALINDIA		INDEPENDENT VARIABLES									DEPENDENT VARIABLE (Daily Adjusted Close)	COALINDIA		INDEPENDENT VARIABLES								DEPENDENT VARIABLE (Daily Adjusted Close)		
Regression Analysis Training	Date	Open	High	Low	Volume	Return	RSI	SMA	LMA	ADX	Actual Price	Regression Analysis Test	Date	Open	High	Low	Volume	Return	RSI	SMA	LMA	ADX	Actual Price	Predicted Price	Error Rate (%)
Training Data	9/18/2017	₹ 261.5	₹ 265.0	₹ 260.1	4114253.00	4.75	75.60	247.70	249.30	23.20	₹ 249.78	Test Data	3/9/2018	₹ 305.0	₹ 307.5	₹ 302.1	2,136,347	-1.4	40.2	306.8	295.5	13.4	₹ 287.1	₹ 288.4	0.5%
	9/19/2017	₹ 262.0	₹ 263.0	₹ 257.1	5364164.00	-6.00	82.10	248.80	249.50	24.10	₹ 244.11		3/12/2018	₹ 305.4	₹ 306.9	₹ 296.0	6,796,993	-6.4	35.8	307.1	296.3	12.6	₹ 281.0	₹ 287.6	2.3%
	9/20/2017	₹ 258.6	₹ 258.7	₹ 253.2	3310089.00	-2.00	62.60	249.80	249.70	24.20	₹ 242.22		3/13/2018	₹ 297.5	₹ 298.7	₹ 290.0	6,306,692	-2.8	23.1	307	297	13.1	₹ 278.4	₹ 282.9	1.6%
	9/21/2017	₹ 255.5	₹ 256.4	₹ 248.5	4473641.00	-2.50	56.50	250.60	249.60	23.40	₹ 239.86		3/14/2018	₹ 291.0	₹ 294.4	₹ 290.7	4,406,927	-1.75	17.9	306.7	297.7	14.3	₹ 276.7	₹ 277.4	0.2%
	9/22/2017	₹ 253.5	₹ 254.5	₹ 248.0	5178307.00	-0.25	48.10	251.20	249.60	21.80	₹ 239.63		3/15/2018	₹ 294.0	₹ 298.9	₹ 292.1	4,487,868	2.4	15.3	306.4	298.3	15.5	₹ 279.0	₹ 275.5	1.3%
	9/25/2017	₹ 254.4	₹ 257.7	₹ 250.3	5684406.00	3.25	44.40	251.90	249.60	20.40	₹ 242.70		3/16/2018	₹ 280.3	₹ 281.4	₹ 276.2	11,016,304	-17.3	24.1	305.8	298.9	15.5	₹ 278.2	₹ 279.8	0.6%
	9/26/2017	₹ 257.5	₹ 261.9	₹ 253.8	4587866.00	3.35	51.30	252.50	249.70	19.40	₹ 245.86		3/19/2018	₹ 279.5	₹ 279.5	₹ 269.1	3,588,096	-5.9	11.7	304.3	299	17.4	₹ 272.3	₹ 256.7	6.1%
	9/27/2017	₹ 261.8	₹ 261.9	₹ 255.3	5990467.00	0.45	58.70	253.70	249.90	19.20	₹ 246.28		3/20/2018	₹ 269.1	₹ 273.3	₹ 268.4	1,804,554	-2.05	9	302.7	299	19.6	₹ 270.3	₹ 257.6	4.9%
	9/28/2017	₹ 260.0	₹ 270.3	₹ 259.0	15893278.00	6.95	60.00	254.60	250.20	19.00	₹ 252.85		3/21/2018	₹ 270.9	₹ 272.6	₹ 269.1	2,739,537	0.65	8	301	298.9	21.8	₹ 270.9	₹ 255.0	6.3%
	9/29/2017	₹ 270.0	₹ 272.4	₹ 267.3	4839652.00	3.05	71.90	256.10	250.50	20.00	₹ 255.73		3/22/2018	₹ 270.0	₹ 273.5	₹ 268.0	3,383,841	-1.5	9.2	299	298.7	23.8	₹ 269.4	₹ 255.6	5.4%
	10/3/2017	₹ 273.0	₹ 274.3	₹ 269.2	2925990.00	1.15	76.20	257.70	250.90	21.20	₹ 256.81		3/23/2018	₹ 266.7	₹ 271.8	₹ 265.0	3,980,861	0.55	7.9	297	298.4	25.7	₹ 270.0	₹ 254.5	6.1%
	10/4/2017	₹ 271.2	₹ 271.9	₹ 268.7	2861280.00	-2.00	78.60	259.00	251.20	22.50	₹ 254.92		3/26/2018	₹ 270.0	₹ 274.5	₹ 268.1	2,277,261	3.5	9.9	294.9	297.7	27.7	₹ 273.5	₹ 252.9	8.1%
	10/5/2017	₹ 270.0	₹ 274.0	₹ 269.1	3173461.00	3.25	72.60	259.80	251.40	23.60	₹ 257.99		3/27/2018	₹ 274.1	₹ 276.8	₹ 272.3	1,888,128	2.75	21.8	293	297	28.9	₹ 276.2	₹ 256.5	7.7%
	10/6/2017	₹ 273.5	₹ 278.5	₹ 273.5	7954834.00	3.45	77.70	260.70	251.60	24.90	₹ 261.25		3/28/2018	₹ 274.7	₹ 285.8	₹ 272.9	6,983,849	7.1	31.1	291.1	296.3	29.5	₹ 283.3	₹ 259.1	9.4%
	10/9/2017	₹ 277.0	₹ 282.0	₹ 276.8	7166260.00	4.65	82.30	261.90	252.00	26.50	₹ 265.64		4/2/2018	₹ 283.4	₹ 283.5	₹ 275.2	2,520,975	-5.8	49.4	289.8	295.8	28.3	₹ 277.5	₹ 264.8	4.8%
	10/10/2017	₹ 281.4	₹ 286.3	₹ 281.4	8451221.00	4.30	87.50	263.20	252.40	28.40	₹ 269.70		4/3/2018	₹ 276.2	₹ 279.6	₹ 275.2	1,641,179	1.45	42.6	288.2	295.3	27.2	₹ 279.0	₹ 260.7	7.0%
	10/11/2017	₹ 286.0	₹ 286.5	₹ 280.8	17103406.00	-2.70	91.90	264.60	253.00	30.50	₹ 267.15		4/4/2018	₹ 278.8	₹ 278.9	₹ 273.2	2,946,453	-4.4	47.8	286.5	295	26.2	₹ 274.6	₹ 261.0	5.2%
	10/12/2017	₹ 281.9	₹ 283.8	₹ 280.2	5889501.00	-0.55	82.60	265.90	253.60	32.30	₹ 266.63		4/5/2018	₹ 276.2	₹ 279.5	₹ 275.5	1,526,718	3.25	42.4	285	294.7	25.5	₹ 277.8	₹ 257.8	7.7%
	10/13/2017	₹ 282.6	₹ 289.0	₹ 282.6	5084089.00	4.85	80.60	267.20	254.30	33.70	₹ 271.21		4/6/2018	₹ 277.7	₹ 278.8	₹ 274.5	1,527,251	-2	50.7	283.5	294.6	24.8	₹ 275.8	₹ 261.6	5.4%
	10/16/2017	₹ 287.3	₹ 291.5	₹ 286.8	5476361.00	3.45	83.90	268.80	255.00	35.50	₹ 274.47		4/9/2018	₹ 276.4	₹ 276.4	₹ 273.5	1,254,956	-1.35	49	282	294.4	24.2	₹ 274.5	₹ 259.0	6.0%
	10/17/2017	₹ 291.0	₹ 291.0	₹ 288.3	3950516.00	-0.65	85.70	270.30	255.80	37.40	₹ 273.85		4/10/2018	₹ 276.0	₹ 281.5	₹ 273.3	2,351,145	6.45	49	280.4	294.2	23.9	₹ 280.9	₹ 257.6	9.1%
	10/18/2017	₹ 290.0	₹ 291.5	₹ 288.3	3633204.00	-0.05	82.80	271.60	256.80	39.10	₹ 273.81		4/11/2018	₹ 283.7	₹ 285.4	₹ 278.1	2,997,141	2.5	62	279.2	293.9	22.5	₹ 283.4	₹ 262.6	7.9%

Regression Equation:

Close Price = *11.75 + 0.06*Open Price + 0.49*High Price + 0.37*Low Price + 0.00*Volume + 0.37*Return + 0.00*RSI + 0.09*SMA + -0.10*LMA + -0.05*ADX*

Average Error Rate: 5.8%	**Predicted price for next trading day:**	**253.28**	***Bearish Trend: Sell the stock***
Training Data:	*9/18/2017*	*To*	*3/8/2018*
Test Data:	*3/9/2018*	*To*	*7/5/2018*

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.997911377
R Square	0.995827117
Adjusted R Sqι	0.995479377
Standard Error	1.038977811
Observations	118

ANOVA

	df	SS	MS	F	Significance F
Regression	9	27821.72317	3091.302575	2863.709569	3.6047E-124
Residual	108	116.5832882	1.079474891		
Total	117	27938.30646			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	11.74849417	3.7930443	3.097378581	0.002488329	4.230022565	19.26696578	4.230022565	19.26696578
Open	0.063148196	0.06777359	0.931752274	0.353542795	-0.071190812	0.197487205	-0.071190812	0.197487205
High	0.489439047	0.058176031	8.413070505	1.81672E-13	0.374124065	0.604754029	0.374124065	0.604754029
Low	0.36511957	0.054801937	6.662530417	1.17022E-09	0.256492628	0.473746513	0.256492628	0.473746513
Volume	-7.58341E-08	3.40775E-08	-2.225343973	0.028136332	-1.43382E-07	-8.28664E-09	-1.43382E-07	-8.28664E-09
Return	0.369955483	0.048549073	7.620237815	1.0346E-11	0.2737228	0.466188165	0.2737228	0.466188165
RSI	0.002160847	0.018730496	0.115365184	0.908369899	-0.034966245	0.039287939	-0.034966245	0.039287939
SMA	0.089700319	0.032381495	2.77011051	0.00659853	0.025514581	0.153886057	0.025514581	0.153886057
LMA	-0.10293296	0.022225169	-4.631369067	1.01856E-05	-0.146987099	-0.05887882	-0.146987099	-0.05887882
ADX	-0.05216682	0.01991888	-2.618963502	0.010087307	-0.091649495	-0.012684145	-0.091649495	-0.012684145

- **Input & Output Screen Design:**

The final steps in the Implementation phase were integrating the front-end and the back-end and prepare a User Interface Excel Dashboard with the following pages:

1. **Index:**

The Index page entails the steps and procedure to be followed before running the application, with step by step explanation of the installations and system setup. The page also provides a Master Button to the user, which gives the user the privilege to download all the back-end data for all the NIFTY stocks at a click. Using this button will take approximately 1 hour 15 min of time to download all the data. However, this will ensure smoother navigation and faster trend views across the various Analysis tabs on this tool. The page also gives an Index with links to other pages in the application. Following is a snapshot of the page:

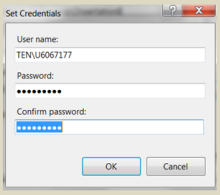
STOCK BUY/SELL PREDICTION APPLICATION FOR AUTOMATED ALGORITHMIC TRADING OF NIFTY STOCKS

Using Predictive Algorithms and Machine Learning Techniques

Pre-requisites before running the application:

1 *Setting up IIS localhost server on your system:*

a) Go To Windows Start and type "Internet Information Services" in the search box
b) Click on "Internet Information Services (IIS) Manager"
c) Click on the arrow besides the system user name under "Connections" to expand the list
d) Right-click on "Sites" and click on "Add Web Site"
e) Put a Site Name of your choice and set the Physical path to the location of the "DissertationII" folder (You would have already extracted, unzipped and saved the DissertationII folder in your local drive)
f) Click on "Pass through Authentication"
g) Click on "Specific User:"
i) Click on "Set"
j) Enter your system user name (along with workgroup name, if any) and your system password



k) Click "OK"
l) Click "OK" again
m) Click on "Test Settings" and ensure the connection is successful before proceeding
n) Assign Port number as "81" and click on OK

One Stop Download of all the Nifty Stocks Data:

Each sheet in this file will let you select a stock and View/Download the latest trends for the selected stock. However, it is highly recommended that you download all the data using the "Download of All Nifty Stocks Data" button below. This will take approximately 1 hour 15 min of time. However, this will ensure smoother navigation and faster trend views across the various Analysis tabs on this tool.

Click on the below button to download all the data. (Approx Download Time: 1 hours 15 minutes):

Download of All Nifty Stocks Data

TABLE OF CONTENTS

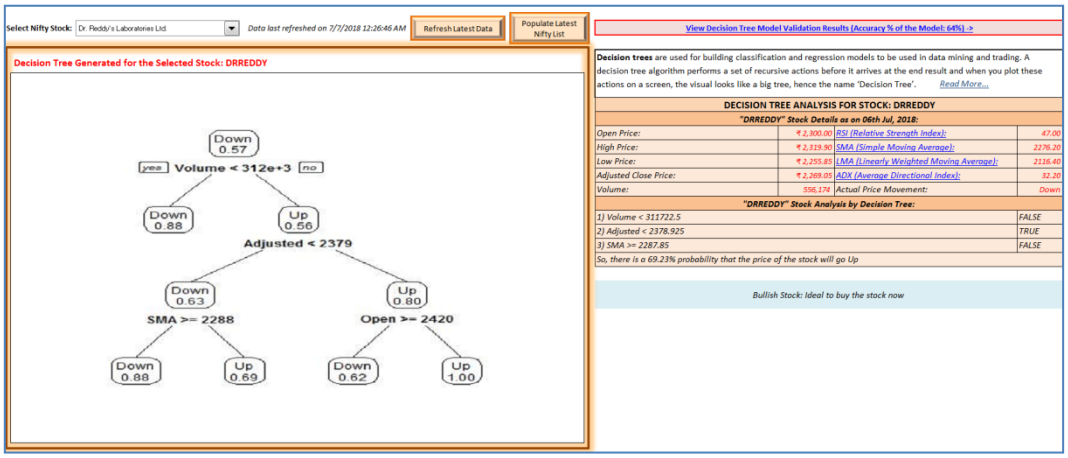
1. DECISION TREE
2. PREDICTIVE INDICATORS
3. LINEAR REGRESSION ANALYSIS
4. SUMMARY & CONCLUSION

2. Decision Tree:

The Decision Tree page provides a dropdown list to the user listing all the NIFTY stock name, from which a user can select one of the stock and view the analysis related to the stock. The page will show when the data was last refreshed and at what time. If the Date of refresh exceeds from the current date, the application fetches current data from Yahoo! Finance server and re-processes the images and validation model. However, if the Date of Refresh does not exceed the current date, the application just refreshes the existing model analysis image and validation on the local drive. The page also provides two buttons. The "Refresh Latest Data" button fetches current data from Yahoo! Finance server, even if the last refreshed date does not exceed the current date. The "Populate Latest Nifty List" button fetches the current NIFTY stocks list from NSE website, as few stocks may move in or out of NSE's Top 50 from time to time. The tree image generated by the Decision Tree model R code is displayed below the dropdown and the buttons. Adjacent to the Image, a short explanation of Decision Tree is mentioned, along with the interpretation of the tree traversal and the final prediction results. On the top of the page, also is a link, which takes the user to the model validation test report and the accuracy percentage. The model validation is done by comparing the Actual Price Movement versus the Predicted Price movement on the historical data.

Following is a snapshot of the page showing Decision Tree Analysis for the stock: **"DRREDDY"** –

Model Results:



As per the model, there is a 69.23% probability that the price of the stock will go Up on the next trading day. The accuracy level of the model is 64%.

Model Validation:

<- Back

Accuracy % of the Model: 64%

Date	Open Price	High Price	Low Price	Close Price	Volume	RSI	SMA	LMA	ADX	Actual Price Movement	Predicted Price Movement	Validate
9/19/2017	2232.00	2265.00	2200.00	2241.55	566790	64.80	2132.00	2316.00	32.50	Down	Up	FALSE
9/20/2017	2250.00	2329.00	2240.00	2314.70	1333098	63.90	2147.60	2306.70	31.30	Up	Up	TRUE
9/21/2017	2346.25	2500.00	2346.25	2486.55	4057364	73.40	2164.10	2298.50	29.10	Up	Up	TRUE
9/22/2017	2494.00	2526.75	2446.55	2457.75	1991351	85.00	2186.90	2294.40	29.00	Down	Down	TRUE
9/25/2017	2464.00	2475.00	2405.00	2417.25	911049	79.60	2205.40	2289.70	29.10	Down	Down	TRUE
9/26/2017	2414.00	2415.00	2322.65	2370.75	1424436	71.70	2223.80	2284.20	28.60	Down	Up	FALSE
9/27/2017	2360.00	2384.00	2281.10	2304.70	935536	63.40	2240.20	2277.60	27.00	Down	Up	FALSE
9/28/2017	2300.00	2415.00	2272.20	2375.40	2815633	53.00	2253.50	2269.70	25.10	Up	Up	TRUE
9/29/2017	2375.40	2390.70	2320.10	2328.75	1133595	60.40	2271.20	2262.80	23.60	Down	Up	FALSE
10/3/2017	2415.00	2415.00	2320.45	2335.55	978860	53.20	2276.80	2253.90	22.20	Up	Up	TRUE
10/4/2017	2342.00	2398.40	2319.55	2387.60	1187980	52.90	2283.80	2245.60	21.20	Up	Up	TRUE
10/5/2017	2392.00	2428.40	2376.00	2389.15	759504	58.10	2292.70	2238.10	20.30	Up	Up	TRUE
10/6/2017	2395.00	2409.90	2368.00	2375.85	533762	57.70	2302.30	2232.00	19.80	Down	Down	TRUE
10/9/2017	2380.00	2433.00	2379.25	2408.95	633048	54.80	2309.90	2225.70	19.20	Up	Up	TRUE
10/10/2017	2419.70	2431.25	2401.60	2411.35	316335	58.70	2322.20	2219.70	19.00	Up	Up	TRUE
10/11/2017	2409.90	2420.05	2351.40	2372.15	450295	58.30	2334.80	2215.50	18.80	Down	Down	TRUE
10/12/2017	2397.00	2397.00	2358.50	2379.85	416094	47.60	2345.10	2213.70	17.70	Up	Up	TRUE
10/13/2017	2373.00	2379.60	2350.00	2354.45	341789	48.10	2353.20	2213.60	16.70	Down	Down	TRUE
10/16/2017	2356.00	2411.00	2356.00	2403.50	386658	43.40	2358.30	2212.00	15.50	Up	Up	TRUE
10/17/2017	2397.00	2397.00	2375.50	2381.35	307110	57.50	2367.90	2212.50	15.10	Down	Down	TRUE
10/18/2017	2386.70	2404.00	2376.00	2385.40	333114	52.60	2374.90	2213.70	14.60	Up	Up	TRUE
10/19/2017	2385.00	2387.00	2362.05	2373.05	47303	53.90	2382.10	2216.60	14.30	Down	Down	TRUE
10/23/2017	2390.10	2415.00	2352.00	2366.20	363956	48.50	2385.00	2219.90	13.70	Down	Down	TRUE
10/24/2017	2373.00	2382.30	2356.00	2363.65	194673	45.90	2379.00	2225.30	13.70	Down	Down	TRUE
10/25/2017	2366.00	2395.10	2340.20	2348.85	295696	43.20	2374.20	2231.70	13.70	Down	Down	TRUE

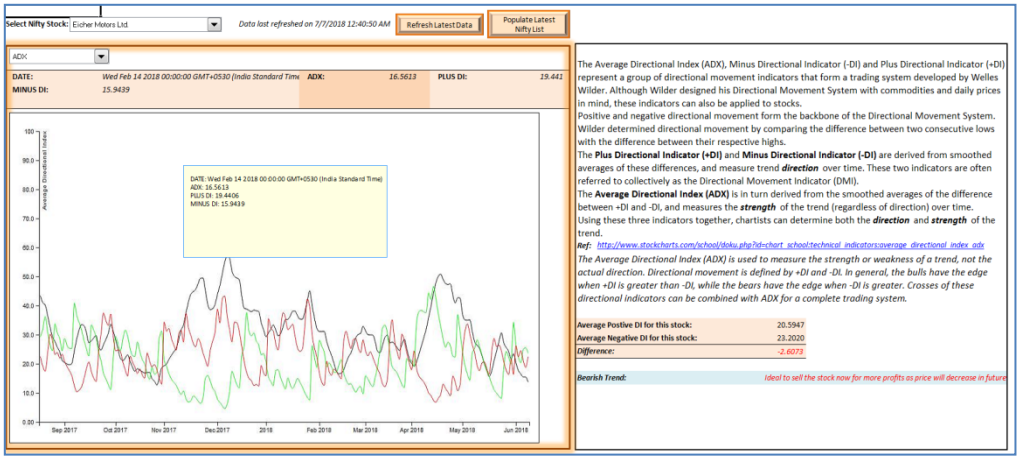
Accuracy % of the Model: 64%

Condition	True Value	False Value
Vol < 311722.5	Down	Adj < 2378.925
<le NA NA		
Adj < 2378.925	SMA >= 2287.85	Ope >= 2419.85
SMA >= 2287.85	Down	Up
<le NA NA		
<le NA NA		
Ope >= 2419.85	Down	Up
<le NA NA		
<le NA NA		

3. Predictive Indicator:

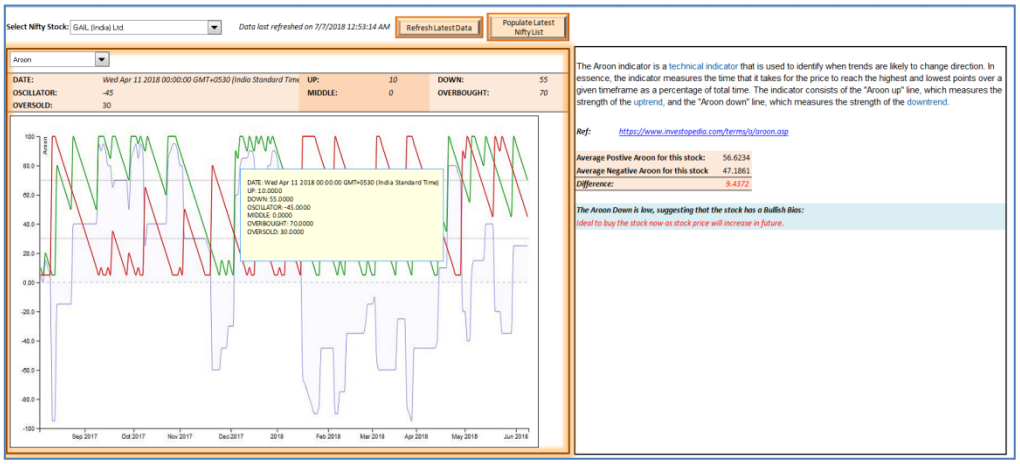
The Predictive Indicator page provides a dropdown list to the user listing all the NIFTY stock name, from which a user can select one of the stock and view the analysis related to the stock. The page will show when the data was last refreshed and at what time. If the Date of refresh exceeds from the current date, the application fetches current data from Yahoo! Finance server and re-processes the images and validation model. However, if the Date of Refresh does not exceed the current date, the application just refreshes the existing model analysis image and validation on the local drive. The page also provides two buttons. The “Refresh Latest Data” button fetches current data from Yahoo! Finance server, even if the last refreshed date does not exceed the current date. The “Populate Latest Nifty List” button fetches the current NIFTY stocks list from NSE website, as few stocks may move in or out of NSE’s Top 50 from time to time. Below the dropdown and the buttons is another dropdown which gives a list of the predictive indicators to the user. The user can choose an indicator from this dropdown, and accordingly the corresponding trend chart image for the indicator will be displayed below the dropdown. On hover over the images, the data points/labels will be displayed as per the region of the hover, both over the chart as a tooltip and values will keep changing in the cells above the chart based on the hover region. Adjacent to the Image, a short explanation of the chosen Predictive Indicator is mentioned, along with the interpretation of the trend chart and the final prediction results.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"EICHERMOT"** for the Technical Indicator: **ADX (Average Directional Index)** –



Looking at the trend chart, the stock has a Bearish trend.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"GAIL"** for the Technical Indicator: **Aroon** –



Looking at the trend chart, the stock has a Bullish bias.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"GRASIM"** for the Technical Indicator: **ATR (Average True Range)** –

Looking at the trend chart, the stock does not have many ups and downs. The stock can be traded by conservative investors.

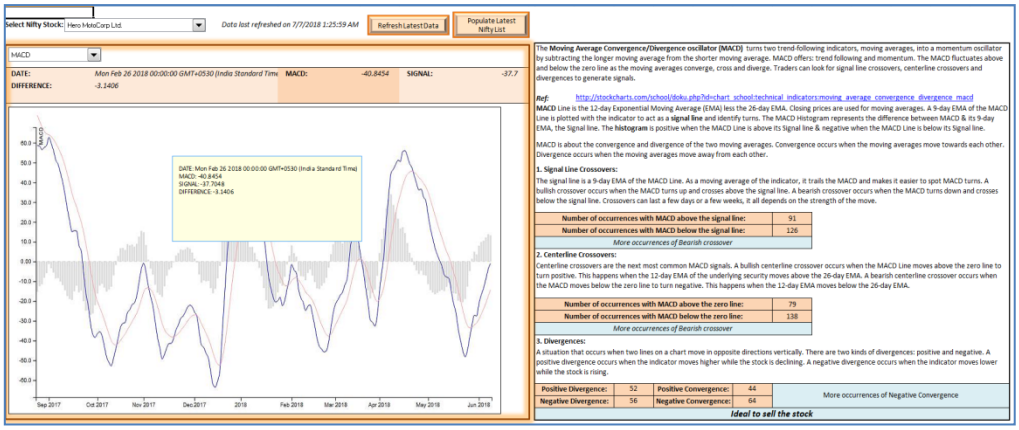
Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HCLTECH"** for the Technical Indicator: **Heikin Ashi** –

Looking at the trend chart, the stock has a Bullish trend.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HDFCBANK"** for the Technical Indicator: **Icimoku Cloud** –

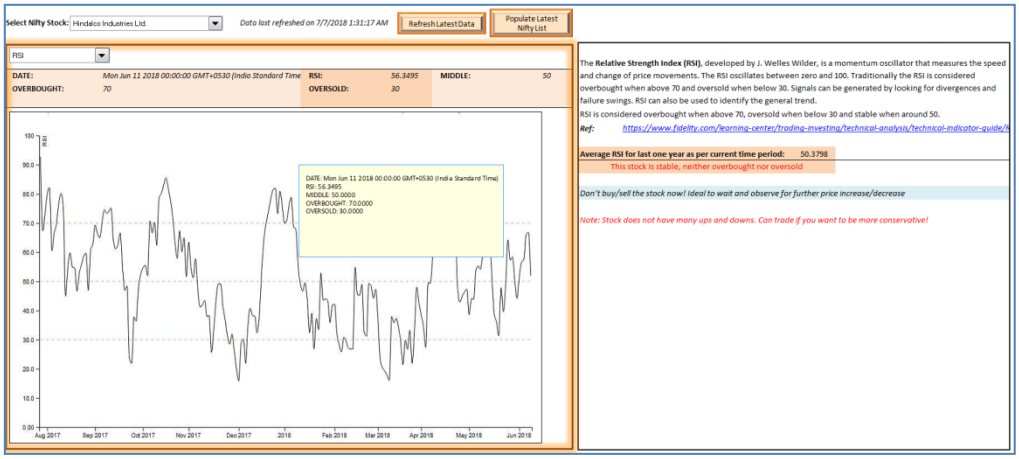
Looking at the trend chart, the stock has a Bearish trend.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HEROMOTOCO"** for the Technical Indicator: **MACD (Moving Average Convergence/Divergence Oscillator)** –



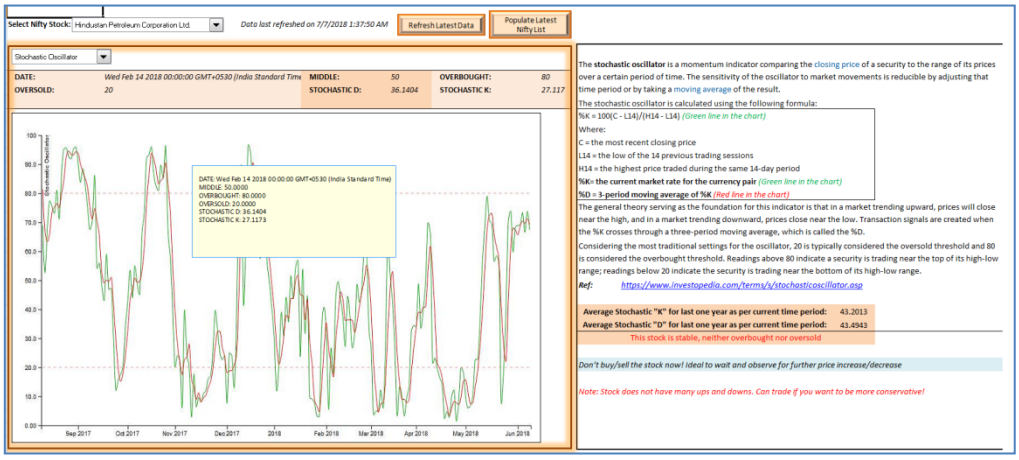
Looking at the trend chart, the stock has a Bearish trend.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HINDALCO"** for the Technical Indicator: **RSI (Relative Strength Index)** –



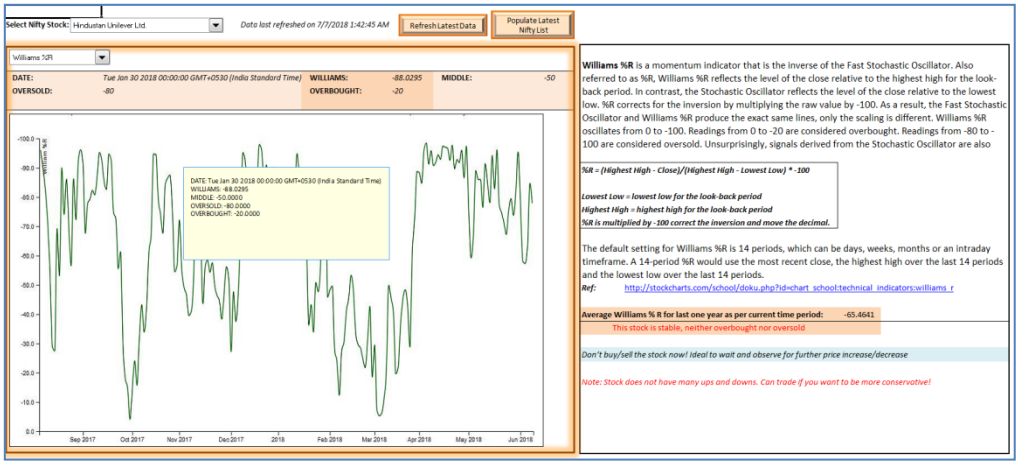
Looking at the trend chart, the stock is stable, neither overbought nor oversold.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HINDPETRO"** for the Technical Indicator: **Stochastic Oscillator** –



Looking at the trend chart, the stock is stable, neither overbought nor oversold.

Following is a snapshot of the page showing Predictive Indicator Analysis for the stock: **"HINDUNILVR"** for the Technical Indicator: **Williams %R –**



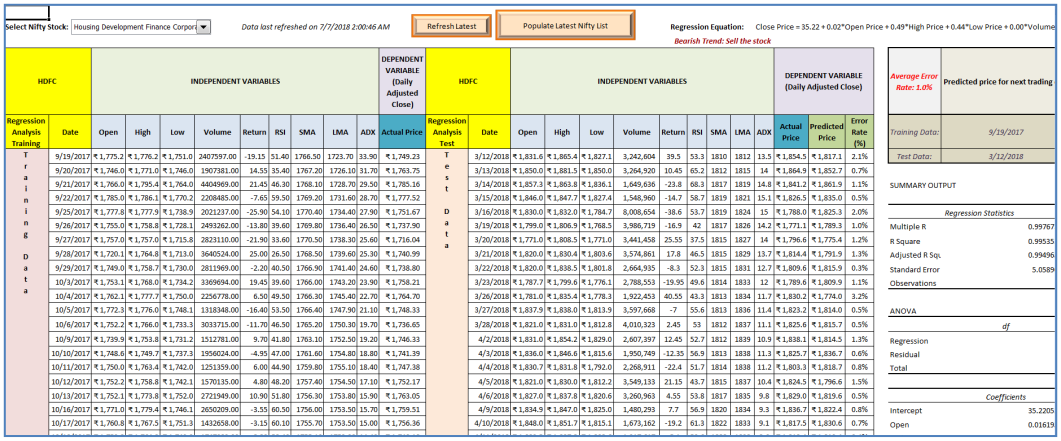
Looking at the trend chart, the stock is stable, neither overbought nor oversold.

4. Regression Analysis:

The Regression Analysis page provides a dropdown list to the user listing all the NIFTY stock name, from which a user can select one of the stock and view the analysis related to the stock. The page will show when the data was last refreshed and at what time. If the Date of refresh exceeds from the current date, the application fetches current data from Yahoo! Finance server and re-processes the images and validation model. However, if the Date of Refresh does not exceed the current date, the application just refreshes the existing model analysis image and validation on the local drive. The page also provides two buttons. The "Refresh Latest Data" button fetches current data from Yahoo! Finance server, even if the last refreshed date does not exceed the current date. The "Populate Latest Nifty List" button fetches the current NIFTY stocks list from NSE website, as few stocks may move in or out of NSE's Top 50 from time to time. Below the dropdown and the buttons are shown the Training and the Test Datasets. The Datasets show the Dependent and the Independent Variables and the Predicted Close Price of the stock derived using the Regression equation generated by the model. Adjacent to the Image, the average error rate percentage along with the date intervals of the training and the test datasets are shown. Below this is shown the Summary output of the Regression Model. On the top of the page, also is shown the Regression Equation used for calculating the Predicted close price of the stock and the prediction results.

Following is a snapshot of the page showing Regression Analysis for the stock: **"HDFC" –**

Model Validation:



Select Nifty Stock: Housing Development Finance Corpor | Data last refreshed on 7/7/2018 2:00:46 AM | Refresh Latest | Populate Latest Nifty List

Regression Equation: Close Price = 35.22 + 0.02*Open Price + 0.49*High Price + 0.44*Low Price + 0.00*Volume + 0.37*Return + 0.13*RSI + 0.05*SMA + -0.03*LMA + 0.17*ADX

Bearish Trend: Sell the stock

HDFC		INDEPENDENT VARIABLES									DEPENDENT VARIABLE (Daily Adjusted Close)
Regression Analysis Training	Date	Open	High	Low	Volume	Return	RSI	SMA	LMA	ADX	Actual Price
Training Data	9/19/2017	₹ 1,775.2	₹ 1,776.2	₹ 1,751.0	2407597.00	-19.15	51.40	1766.50	1723.70	33.90	₹ 1,749.23
	9/20/2017	₹ 1,746.0	₹ 1,771.0	₹ 1,746.0	1907381.00	14.55	35.40	1767.20	1726.10	31.70	₹ 1,763.75
	9/21/2017	₹ 1,766.0	₹ 1,795.4	₹ 1,764.0	4404969.00	21.45	46.30	1768.10	1728.70	29.50	₹ 1,785.16
	9/22/2017	₹ 1,785.0	₹ 1,786.1	₹ 1,770.2	2208485.00	-7.65	59.50	1769.20	1731.60	28.70	₹ 1,777.52
	9/25/2017	₹ 1,777.8	₹ 1,777.9	₹ 1,738.9	2021237.00	-25.90	54.10	1770.40	1734.40	27.90	₹ 1,751.67
	9/26/2017	₹ 1,755.0	₹ 1,758.8	₹ 1,728.1	2493262.00	-13.80	39.60	1769.80	1736.40	26.50	₹ 1,737.90
	9/27/2017	₹ 1,757.0	₹ 1,757.0	₹ 1,715.8	2823110.00	-21.90	33.60	1770.50	1738.30	25.60	₹ 1,716.04
	9/28/2017	₹ 1,720.1	₹ 1,764.8	₹ 1,713.0	3640524.00	25.00	26.50	1768.50	1739.60	25.30	₹ 1,740.99
	9/29/2017	₹ 1,749.0	₹ 1,758.7	₹ 1,730.0	2811969.00	-2.20	40.50	1766.90	1741.40	24.60	₹ 1,738.80
	10/3/2017	₹ 1,753.1	₹ 1,768.0	₹ 1,734.2	3369694.00	19.45	39.60	1766.00	1743.20	23.90	₹ 1,758.21
	10/4/2017	₹ 1,762.1	₹ 1,777.7	₹ 1,750.0	2256778.00	6.50	49.50	1766.30	1745.40	22.70	₹ 1,764.70
	10/5/2017	₹ 1,772.3	₹ 1,776.0	₹ 1,748.1	1318348.00	-16.40	53.50	1766.40	1747.90	21.10	₹ 1,748.33
	10/6/2017	₹ 1,752.2	₹ 1,766.0	₹ 1,733.3	3033715.00	-11.70	46.50	1765.20	1750.30	19.70	₹ 1,736.65
	10/9/2017	₹ 1,739.9	₹ 1,753.8	₹ 1,731.2	1512781.00	9.70	41.80	1763.10	1752.50	19.20	₹ 1,746.33
	10/10/2017	₹ 1,748.6	₹ 1,749.7	₹ 1,737.3	1956024.00	-4.95	47.00	1761.60	1754.80	18.80	₹ 1,741.39
	10/11/2017	₹ 1,750.0	₹ 1,763.4	₹ 1,742.0	1251359.00	6.00	44.90	1759.80	1755.10	18.40	₹ 1,747.38
	10/12/2017	₹ 1,752.2	₹ 1,758.8	₹ 1,742.1	1570135.00	4.80	48.20	1757.40	1754.50	17.10	₹ 1,752.17
	10/13/2017	₹ 1,752.1	₹ 1,773.8	₹ 1,752.0	2721949.00	10.90	51.80	1756.30	1753.80	15.90	₹ 1,763.05
	10/16/2017	₹ 1,771.0	₹ 1,779.4	₹ 1,746.1	2650209.00	-3.55	60.50	1756.00	1753.50	15.70	₹ 1,759.51
	10/17/2017	₹ 1,760.8	₹ 1,767.5	₹ 1,751.3	1432658.00	-3.15	60.10	1755.70	1753.50	15.00	₹ 1,756.36

HDFC		INDEPENDENT VARIABLES									DEPENDENT VARIABLE (Daily Adjusted Close)		
Regression Analysis Test	Date	Open	High	Low	Volume	Return	RSI	SMA	LMA	ADX	Actual Price	Predicted Price	Error Rate (%)
Test Data	3/12/2018	₹ 1,831.6	₹ 1,865.4	₹ 1,827.1	3,242,604	39.5	53.3	1810	1812	13.5	₹ 1,854.5	₹ 1,817.1	2.1%
	3/13/2018	₹ 1,850.0	₹ 1,881.5	₹ 1,850.0	3,264,920	10.45	65.2	1812	1815	14	₹ 1,864.9	₹ 1,852.7	0.7%
	3/14/2018	₹ 1,857.3	₹ 1,863.8	₹ 1,836.1	1,649,636	-23.8	68.3	1817	1819	14.8	₹ 1,841.2	₹ 1,861.9	1.1%
	3/15/2018	₹ 1,846.0	₹ 1,847.7	₹ 1,827.4	1,548,960	-14.7	58.7	1819	1821	15.1	₹ 1,826.5	₹ 1,835.0	0.5%
	3/16/2018	₹ 1,830.0	₹ 1,832.0	₹ 1,784.7	8,008,654	-38.6	53.7	1819	1824	15	₹ 1,788.0	₹ 1,825.3	2.0%
	3/19/2018	₹ 1,799.0	₹ 1,806.9	₹ 1,768.5	3,986,719	-16.9	42	1817	1826	14.2	₹ 1,771.1	₹ 1,789.3	1.0%
	3/20/2018	₹ 1,771.0	₹ 1,808.5	₹ 1,771.0	3,441,458	25.55	37.5	1815	1827	14	₹ 1,796.6	₹ 1,775.4	1.2%
	3/21/2018	₹ 1,820.0	₹ 1,830.4	₹ 1,803.6	3,574,861	17.8	46.5	1815	1829	13.7	₹ 1,814.4	₹ 1,791.9	1.3%
	3/22/2018	₹ 1,820.0	₹ 1,838.5	₹ 1,801.8	2,664,935	-8.3	52.3	1815	1831	12.7	₹ 1,809.6	₹ 1,815.9	0.3%
	3/23/2018	₹ 1,787.7	₹ 1,799.6	₹ 1,776.1	2,788,553	-19.95	49.6	1814	1833	12	₹ 1,789.6	₹ 1,809.9	1.1%
	3/26/2018	₹ 1,781.0	₹ 1,835.4	₹ 1,778.3	1,922,453	40.55	43.3	1813	1834	11.7	₹ 1,830.2	₹ 1,774.0	3.2%
	3/27/2018	₹ 1,837.9	₹ 1,838.0	₹ 1,813.9	3,597,668	-7	55.6	1813	1836	11.4	₹ 1,823.2	₹ 1,814.0	0.5%
	3/28/2018	₹ 1,821.0	₹ 1,831.0	₹ 1,812.8	4,010,323	2.45	53	1812	1837	11.1	₹ 1,825.6	₹ 1,815.7	0.5%
	4/2/2018	₹ 1,831.0	₹ 1,854.2	₹ 1,829.0	2,607,397	12.45	52.7	1812	1839	10.9	₹ 1,838.1	₹ 1,814.5	1.3%
	4/3/2018	₹ 1,836.0	₹ 1,846.6	₹ 1,815.6	1,950,749	-12.35	56.9	1813	1838	11.3	₹ 1,825.7	₹ 1,836.7	0.6%
	4/4/2018	₹ 1,830.7	₹ 1,831.8	₹ 1,792.0	2,268,911	-22.4	51.7	1814	1838	11.2	₹ 1,803.3	₹ 1,818.7	0.8%
	4/5/2018	₹ 1,821.0	₹ 1,830.0	₹ 1,812.2	3,549,133	21.15	43.7	1815	1837	10.4	₹ 1,824.5	₹ 1,796.6	1.5%
	4/6/2018	₹ 1,827.0	₹ 1,837.8	₹ 1,820.6	3,260,963	4.55	53.8	1817	1835	9.8	₹ 1,829.0	₹ 1,819.6	0.5%
	4/9/2018	₹ 1,834.9	₹ 1,847.0	₹ 1,825.0	1,480,293	7.7	56.9	1820	1834	9.3	₹ 1,836.7	₹ 1,822.4	0.8%
	4/10/2018	₹ 1,848.0	₹ 1,851.7	₹ 1,815.1	1,673,162	-19.2	61.3	1822	1833	9.1	₹ 1,817.5	₹ 1,830.6	0.7%

Average Error Rate: 1.0%	Predicted price for next trading
Training Data:	9/19/2017
Test Data:	3/12/2018

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99767
R Square	0.99535
Adjusted R Sqι	0.99496
Standard Error	5.0589
Observations	

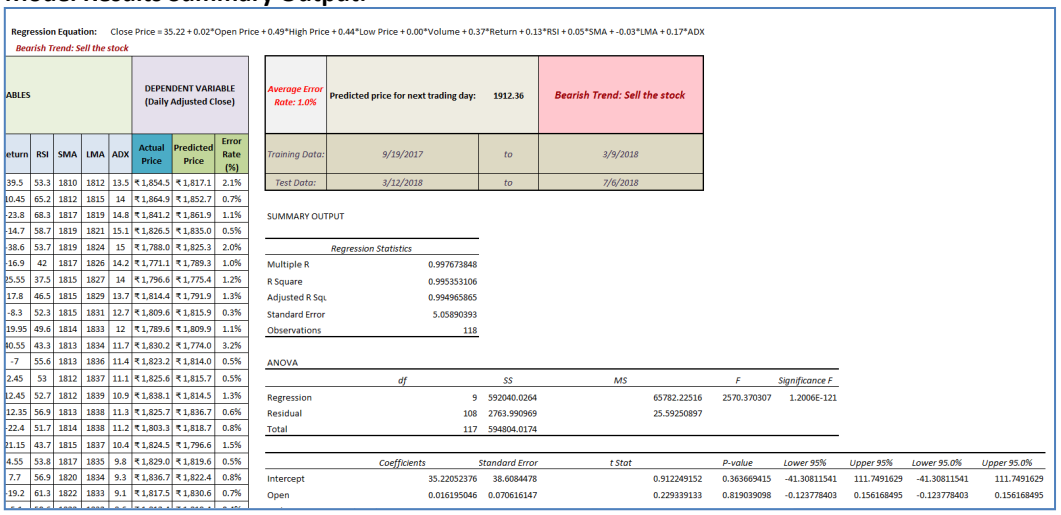
ANOVA

	df
Regression	
Residual	
Total	

	Coefficients
Intercept	35.2205
Open	0.01619

As per the model, price of the stock will go Down on the next trading day. However, the average error is very high: +1.0%.

Model Results Summary Output:



Regression Equation: Close Price = 35.22 + 0.02*Open Price + 0.49*High Price + 0.44*Low Price + 0.00*Volume + 0.37*Return + 0.13*RSI + 0.05*SMA + -0.03*LMA + 0.17*ADX

Bearish Trend: Sell the stock

ABLES					DEPENDENT VARIABLE (Daily Adjusted Close)		
eturn	RSI	SMA	LMA	ADX	Actual Price	Predicted Price	Error Rate (%)
39.5	53.3	1810	1812	13.5	₹ 1,854.5	₹ 1,817.1	2.1%
10.45	65.2	1812	1815	14	₹ 1,864.9	₹ 1,852.7	0.7%
-23.8	68.3	1817	1819	14.8	₹ 1,841.2	₹ 1,861.9	1.1%
-14.7	58.7	1819	1821	15.1	₹ 1,826.5	₹ 1,835.0	0.5%
-38.6	53.7	1819	1824	15	₹ 1,788.0	₹ 1,825.3	2.0%
-16.9	42	1817	1826	14.2	₹ 1,771.1	₹ 1,789.3	1.0%
25.55	37.5	1815	1827	14	₹ 1,796.6	₹ 1,775.4	1.2%
17.8	46.5	1815	1829	13.7	₹ 1,814.4	₹ 1,791.9	1.3%
-8.3	52.3	1815	1831	12.7	₹ 1,809.6	₹ 1,815.9	0.3%
19.95	49.6	1814	1833	12	₹ 1,789.6	₹ 1,809.9	1.1%
40.55	43.3	1813	1834	11.7	₹ 1,830.2	₹ 1,774.0	3.2%
-7	55.6	1813	1836	11.4	₹ 1,823.2	₹ 1,814.0	0.5%
2.45	53	1812	1837	11.1	₹ 1,825.6	₹ 1,815.7	0.5%
12.45	52.7	1812	1839	10.9	₹ 1,838.1	₹ 1,814.5	1.3%
12.35	56.9	1813	1838	11.3	₹ 1,825.7	₹ 1,836.7	0.6%
-22.4	51.7	1814	1838	11.2	₹ 1,803.3	₹ 1,818.7	0.8%
21.15	43.7	1815	1837	10.4	₹ 1,824.5	₹ 1,796.6	1.5%
4.55	53.8	1817	1835	9.8	₹ 1,829.0	₹ 1,819.6	0.5%
7.7	56.9	1820	1834	9.3	₹ 1,836.7	₹ 1,822.4	0.8%
-19.2	61.3	1822	1833	9.1	₹ 1,817.5	₹ 1,830.6	0.7%

Average Error Rate: 1.0%	Predicted price for next trading day: 1912.36	Bearish Trend: Sell the stock
Training Data:	9/19/2017 to	3/9/2018
Test Data:	3/12/2018 to	7/6/2018

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.997673848
R Square	0.995353106
Adjusted R Sqι	0.994965865
Standard Error	5.05890393
Observations	118

ANOVA

	df	SS	MS	F	Significance F
Regression	9	592040.0264	65782.22516	2570.370307	1.2006E-121
Residual	108	2763.990969	25.59250897		
Total	117	594804.0174			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	35.22052376	38.6084478	0.912249152	0.363669415	-41.30811541	111.7491629	-41.30811541	111.7491629
Open	0.016195046	0.070616147	0.229339133	0.819039098	-0.123778403	0.156168495	-0.123778403	0.156168495

5. Summary_Conclusion:

This is the final page which gives the analysis output of each model/indicator in a summarized form. The page also gives a final decision to the user as to whether to buy or sell the stock, and if the stock is stable or risky, considering the following:

a) If the Accuracy % of the Decision Tree Model is more than 69%, then the predictive result of the Decision Tree Analysis is considered as the final result

b) Else, If the Error Rate % of the Regression Model is less than 0.2%, then the predictive results of the Regression Analysis is considered as the final result

c) Else, All the results are considered, and the final results is based on majority of similar results.

Following is a snapshot of the page showing Summary of Predictive Analysis for the stock: **"ITC"** –

I T C LTD. - STOCK PREDICTIVE ANALYSIS SUMMARY

1. OHLVC (Open, High, Low, Volume, Close) data downloaded from Yahoo Finance for 12 months till date: July 6, 2018

2. Predictive Analysis Results from the "Decision Tree" *(Performs a set of recursive actions before it arrives at the end result and the visual looks like a big tree):*

DECISION TREE ANALYSIS FOR STOCK: ITC

"ITC" Stock Details as on 06th Jul, 2018:

Open Price:	*$272.75*	*RSI (Relative Strength Index):*	*66.80*
High Price:	*$274.75*	*SMA (Simple Moving Average):*	*265.50*
Low Price:	*$270.15*	*LMA (Linearly Weighted Moving Average):*	*272.90*
Adjusted Close Price:	*$272.65*	*ADX (Average Directional Index):*	*18.40*
Volume:	*12,936,762*	*Actual Price Movement:*	*Up*

Accuracy % of the Decision Tree Model for this Stock is: ***67%***

There is a 60% probability that the price of the stock will go Up as per the Decision Tree Analysis

3. Predictive Analysis Results from the following "Predictive Indicator" trends:

a. ADX (Average Directional Index) *(Used to measure the strength or weakness of a trend):*

The average difference between the Postive Directional Index and Negative Directional index for the stock is ***-1.4348*** *Bulls have the edge when +DI is greater than -DI, while Bears have, the edge when -DI is greater. So, the stock has a* ***Bearish Trend,*** *as per the ADX indicator.*

b. Aroon *(Used to identify when trends are likely to change direction):*

The average difference between the Postive Aroon and Negative Aroon is : ***-17.2294*** *The Aroon Up is low, suggesting that the stock has a* ***Bearish Bias.***

c. ATR (Average True Range) *(Volatility measure of a stock's performance):*

The average Volatility ratio is ***0.0191*** *which means the stock is stable.*

The average Target Price is: ***$286.97*** *& average Stop Loss Price is:* ***$266.21***

Plan your buying/selling of the stock accordingly, making maximum profit at the target price and exit the trade when the price touches the Stop Loss point.

d. Heikin Ashi *(Uses the open-close data from the prior period and the open-high-low-close data from the current period to create a combo candlestick):*

Looking at the candlestick bars for the given time frame, Strong Advances are ***less than*** *Strong Declines.*

e. Ichimoku Cloud *(Defines support and resistance, identifies trend direction, gauges momentum and provides trading signals):*

The average Chikou Span for the stock is 267.0602, with a Bearish Trend (Average value less than most of the values)

The average Kijun Sen for the stock is 268.1842, with a Bearish Trend (Average value less than most of the values)

The average Senkou Span A for the stock is 267.7913, with a Bearish Trend (Average value less than most of the values)

The average Senkou Span B for the stock is 269.5065, with a Bullish Trend (Average value more than most of the values)

The average Tenkan Sen for the stock is 267.3984, with a Bearish Trend (Average value less than most of the values)

f. MACD (Moving Average Convergence/Divergence Oscillator) *(Convergence and Divergence of Moving Averages):*

The stock indicates More occurrences of Bullish signal line crossovers. *And, the stock also indicates More occurrences of Bearish center line crossovers.*

The stock also indicates More occurrences of Bullish Divergences

g. RSI (Relative Strength Index) *(Momentum oscillator that measures the speed and change of price movements):*

The Average RSI for last one year as per current time period is ***46.7484*** *This stock is stable, neither overbought nor oversold, as per the RSI indicator.*

h. Stochastic Oscillator *(Momentum indicator comparing the closing price of a security to the range of its prices over a certain period of time):*

The Average Stochastic "K" for last one year as per current time period is ***43.5687*** *The Average Stochastic "D" for last one year as per current time period is* ***43.1980***

This stock is stable, neither overbought nor oversold, as per the Stochastic indicator.

i. Williams %R *(Inverse of the Fast Stochastic Oscillator):*

The Average Williams % R for last one year as per current time period is ***-43.0581*** *This stock is stable, neither overbought nor oversold as per the Williams %R indicator.*

4. Predictive Analysis Results from "Regression Analysis" *(An approach for modeling the relationship between a scalar dependent variable y (**in this case price of stock**) and one or more explanatory variables (or independent variables) denoted X (**in this case date of stock**):*

Average Error rate of the Regression Model for this stock is: ***1.4%***

The price of stock as on July 6, 2018 was : *$272.65*

The price of stock on the next trading day is predicted to be : *$266.40*

So, the stock has a Bearish Trend, as per Regression Analysis

CONCLUSION:

Considering the findings of most of the predictive indicators and the accuracy level of the predictive models, we conclude that:

1. *The stock has a Bearish trend. Ideal to sell the stock before further price downfall*
2. *The stock is stable, neither overbought nor oversold. Suitable for conservative trading.*

d) System Integration & Testing: Test Report

Testing Plan:

Type of Test	Will Test be Performed	Comments/Explanation	Software Component
Requirements Testing	Yes	Needs to be done to cope up with changing environment	Fluctuation in the share market
Unit	Yes	Maximum number of defects are found. Each component of code was tested or analyzed accordingly not only to ensure the best quality of the developed software but also to make sure that code behaves in the same way as it was intended to. Unit testing was performed as and when the component was developed.	• User Interface Code • Decision Tree Code • Regression Analysis Code • Predictive Indicators Trend Charts Generation Code • NIFTY stocks list and stock latest data download Code • Prediction of Day Code
Integration	Yes	All the well-developed sub-system are integrated together and tested called as integration testing.	• Decision Tree Algorithm in R • Regression Analysis Algorithm in Excel • Predictive Indicator Charts Algorithm in D3.JS
Performance	Yes	Performance is the major criteria for evaluating any type of the system. It holds importance and is tested likewise.	Performance of different models and algorithms is measured in combination using the following two approaches: • **True Rates for Individual Approach** *(prediction model trained on the first 60 percent and test on the rest 40 percent)* • **Statistical Testing** *(Model validations: Accuracy % of Decision Tree Model and Error rate of Regression Model)*
Stress	No	-	-
Compliance	No	-	-
Security	No	-	-

Test Environment:

Software Items:

- Windows 7
- IIS Server
- Internet connection
- Microsoft Excel
- R for Windows 3.4.3

Hardware Items:

- Personal Computer/Laptop
- Wireless connection or connecting cable

Component Decomposition & Identification of Tests Required:

S.No	Components that require Testing	Type of Testing Required	Technique for writing Test Case
1	Decision Tree Code	Unit Testing	White Box Testing
2	Regression Analysis Code	Unit Testing	White Box Testing
3	Predictive Indicators Trend Charts Generation Code	Unit Testing	White Box Testing
4	NIFTY Stocks List & Stock Latest Data Download Code	Unit Testing	White Box Testing
5	Prediction of Day Code	Unit Testing	White Box Testing
6	User Interface Code	Performance Testing	Black Box Testing
7	Destination	System Testing	Black Box Testing
8	Source	System Testing	Black Box Testing

Test Cases:

Test ID	T1
Input	Enter the Stock Symbol to update the data
Expected Output	Data fetched from Yahoo! Finance
Status	Pass

Test ID	T2
Input	Predict the stock rate for the very next trading day
Expected Output	We get Close Price and Price movement for the next trading day
Status	Pass

Test ID	T3
Input	Check the precision of output by predicting the data on a date whose values are already known
Expected Output	Outputs are partially precise
Status	Pass

Error & Exception Handling:

Test Case ID	Test Case	Debugging Technique
T1	Fetching data from Yahoo! Finance	Debug the R Code and check for errors due to format change in Yahoo! Finance website or incorrect date calculations

Limitation of the Solution:

- The precision of the output sometimes is not even near to the actual value
- System might stop running in between if connection to the internet is lost, as data is fetched from Yahoo! Finance

e) Deployment of System: User/Operational Manual

Pre-requisites before running the Application:

1. **Setting up IIS Local host Server on your system:** *(This is required to generate the D3 charts)*

   a. Go To Windows Start and type "Internet Information Services" in the search box

   b. Click on "Internet Information Services (IIS) Manager"

   c. Click on the arrow besides the system user name under "Connections" to expand the list

   d. Right-click on "Sites" and click on "Add Web Site"

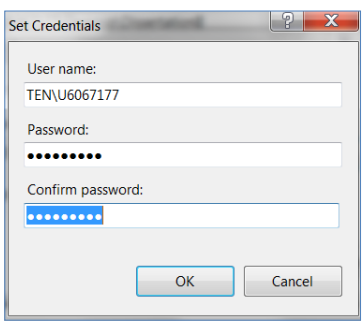
   e. Put a Site Name of your choice and set the Physical path to the location of the "DissertationII" folder (DissertationII folder available in the Soft Copy – to be extracted, unzipped and saved the in the local drive)

   f. Click on "Pass through Authentication"

   g. Click on "Specific User:"

   h. Click on "Set"

   i. Enter your system user name (along with workgroup name, if any) and your system password

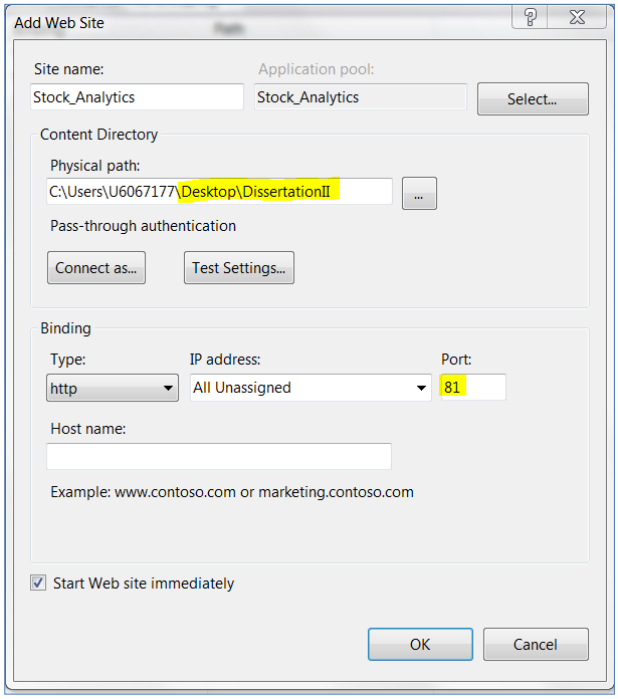


j. Click "OK"

k. Click "OK" again

l. Click on "Test Settings" and ensure the connection is successful before proceeding

m. Assign Port number as "81" and click on OK



n. Now go to your browser (preferably Chrome) and type "localhost:81" in the address bar

o. Ensure you do not get any error message & all sub-folders in the DissertationII folder are listed on the page

2. **Installing R for Windows 3.4.3:** *(This is required to generate Decision Tree model and Web scraping of Yahoo! Finance)*

   a. Download the setup of "R for Windows 3.4.3" from the following link: https://cran.r-project.org/bin/windows/base/old/3.4.3/

   b. Install R by opening the setup file and following the instructions

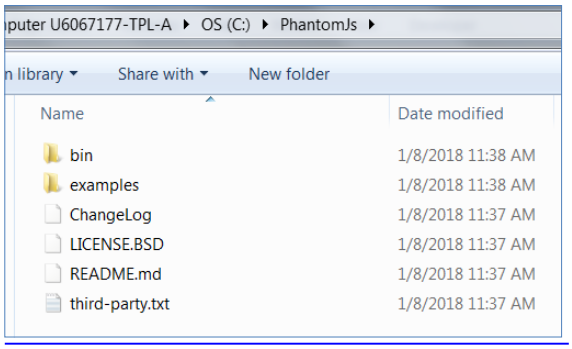
3. **Setting up PhantomJS:** *(This is required to integrate D3 and R with Excel VBA)*

   a. Download the "phantomjs.exe" file for windows from the following link: http://phantomjs.org/download.html

   b. Extract the contents of the downloaded zip file

   c. Create a folder in C:\ drive called "PhantomJs"

d. Copy the contents of the downloaded extracted folder in the C:\PhantomJs folder:



4. **Setting up Environment Variables:** *(This is required to run R and PhantomJS scripts using Excel VBA Shell Scripts)*

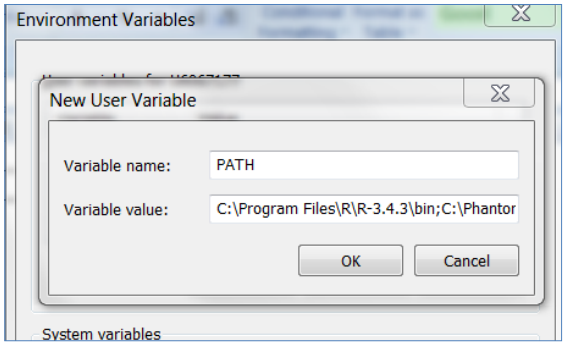
   a. Go To Windows Start and type "Environment Variables" in the search box

   b. Click on "Edit environment variables for your account"

   c. Click on the "New" button in the "Environment Variables" Dialog box

   d. Mention "Variable Name" as: "PATH"

   e. Mention "Variable Value" as: the path of the exe files for R and PhantomJS like this: *C:\Program Files\R\R-3.4.3\bin;C:\PhantomJs\bin*



5. **Open the "Stock Analysis.xlsm – Shortcut" file available in the soft copy to use the Application**

f) System Maintenance & Evaluation:

System Upgrades:

The current application uses the following software versions:

- Microsoft Office 2007
- R for Windows 3.4.3

Upgrade of Microsoft Office would not give any problems in the future as Microsoft Office 2007 is compatible with the higher versions of Microsoft Office.

However, if R for Windows is upgraded to a higher version, the admin might need to test if the system is still functioning properly and accordingly make changes to the R Scripts. This is because R might change certain packages and the coding syntax in the higher versions.

Limitations of the Application:

- The results of the application might not be 100% accurate as the Decision Tree model shows an average accuracy % of 65 to 70%, and the Regression Model has quite a high error rate of 1% for many stocks. Also, the trend analysis is just a cognitive analysis of more Upward vs More Downward Price movements

- The Independent variables used for Prediction can be increased which have a significant impact and relation to the price fluctuations. Example: Competitors, Mergers and Acquisitions etc. However, as this is just a dummy application, such real time variable are not readily available, and so haven't been used

- A piece of data which is **missing in this project is the intraday prices**, i.e. the prices minute by minute. However, intraday prices are not as freely available as interday prices and are considered a commodity in themselves. To get hold of such a dataset would incur a large cost, one that is not within the budget of a project such as this.

- Another important piece of **missing data is the order book**. The order book is a record of live buy and sell orders for a particular stock. Successful orders are matched off against the order book by the exchange. It is easy to imagine that the order book contains useful data. For instance, the weighted average of orders might be predictive of the price. However access to this data is extremely costly and far beyond what most casual investors can afford, let alone the budget for this project.

- We are trying to quantify the true value of a company when almost every company has in some way or another some purely qualitative value. Fundamental Analysis methods do not attempt to capture these qualitative values. For example, it is not possible to quantify the value of a brand, the size of its customer base, or a competitive advantage. Until these values are quantified, it leaves a large gap in what an algorithmic style approach can achieve. For instance, what algorithm, would have valued WhatsApp at $22 billion while still making an year-on-year loss? In fundamental analysis we are **limited to purely quantitative company metrics**.

Future Enhancements

- More mutually affecting Independent variables can be added for better prediction results

- More advanced prediction algorithms such as Logistic Regression, Neural Networks, etc. can be use to enhance the quality of the Prediction

- Portfolio management can be added to our existing analysis. Portfolio management is largely an extra step done after an investor has made a prediction on which direction any particular stock will move. For instance, the investor may choose not to invest all of their funds into a single company lest that company takes unexpected turn.

FINDINGS & CONCLUSION

To summarize, in this project, we attempt to build an automated trading system based on Machine Learning algorithms. Based on historical price information, the machine learning models will forecast next day returns of the target stock. A customized trading strategy will then take the model prediction as input and generate actual buy/sell orders and send them to a market simulator where the orders are executed. After training on available data at a particular time interval, our application will back test on out of sample data at a future time interval.

Following are some of the important **Findings** that were discovered after building this project:

- We found that only looking at a company's past stock price itself is not sufficient enough to predict its future returns. Better ways to do so is to look at the entire sector which the target company is part of, and use historical price information of all companies within the sector to predict the target's next day return.

- The Decision Tree model has achieved approximately 66 – 70 percent accuracy for most of the stocks with statistical significance.

- The Regression Model has achieved a high error rate close to 1% for many stocks, and so steps should be taken in the real time environment to increase the Independent variables for this analysis For future works, Variables about company fundamentals such as revenues and earnings and about macroeconomic issues such as interest rates, exchange rates and unemployment reports should also help predicting stock prices.

- Automated trading should not be just about algorithms, programming and mathematics: an awareness of fundamental market and macroeconomic issues is also needed to help us decide whether the back test is predictive and the automated trading system will continue to be predictive.

APPENDIX

a)Brief Background of my Organization:

Thomson Reuters is the world's leading source of intelligent information for businesses and professionals. Intelligent information is a unique synthesis of human intelligence, industry expertise and innovative technology that provides decision-makers with the knowledge to act, enabling them to make better decisions faster. Through our Financial & Risk (F&R) unit, we provide solutions to the global financial community - delivering critical news, information and analytics, enabling transactions and connecting communities of trading, investing, financial and corporate professionals.

The Customer Analytics function sits within the F&R business unit and provides actionable insights about our business and customers. Our goal is to develop analytical frameworks that enable our business to achieve its growth, customer engagement and retention objectives. The hired person will not only develop the frameworks but also play a key role in collaborating with business partners and stakeholders to translate the insight into actionable strategies and initiatives.

My role in the Organization:

Job Role: *Senior Customer Analyst*
Role Definition: *Eikon Platform Customer Predictive Analytics*
Responsibilities:

- Learn and understand Thomson Reuters' business strategy – our customers, products, competition, market position and the key challenges facing the organization

- Mine customer and market data to generate actionable insights that are used by our internal business partners to improve business results

    - Complex data extraction, integration & manipulation using Oracle and R
    - Advanced statistics, data science & algorithm development; predictive analytics, etc.
    - Analytical techniques include logistic regression, clustering, etc.
    - Data visualization & insight generation using Tableau and Excel
    - Consultative approach to informing decision making
    - Operationalization of data driven strategies

- Demonstrate thought leadership when understanding end-to-end business objectives and develop appropriate analytical frameworks to facilitate achievement of objectives

- Gain insight or develop analytical frameworks from large, disparate datasets

b) Data Dictionary:

NIFTY Stocks List:

Field Name	Data Type	Data Format	Field Size	Description	Example
Serial Num	Integer	NNNNNN	6	Unique number ID for stock	24
Company	String	-	100	Name of the stock's Company	ICICI Bank Ltd.
Industry	String	-	100	Industry of stock's company	Financial
Symbol	String	-	100	Stock's NSE Trading Symbol	ICICIBANK
Series	String	-	2	Type of Trading: Equity	EQ
ISIN Code	String	-	30	International Securities Identification Number	INE090A01021

Stock Daily Trading Data:

Field Name	Data Type	Data Format	Field Size	Description	Example
Stock Symbol	String	-	100	Stock's NSE Trading Symbol	ICICIBANK
Date	Date/Time	DD-MMM-YY	9	Date of Trading	31-May-18
Open Price	Floating (Currency)	$NNN.NN	7	Open Price of the stock at the beginning of the trading day	$285.00
High Price	Floating (Currency)	$NNN.NN	7	Highest Price of the stock on the trading day	$288.95
Low Price	Floating (Currency)	$NNN.NN	7	Lowest Price of the stock on the trading day	$279.30
Close Price	Floating (Currency)	$NNN.NN	7	Close Price of the stock at the end of the trading day	$285.80
Adjusted Close Price	Floating (Currency)	$NNN.NN	7	*Adjusted Close Price of stock at the end of the trading day	$285.80
Volume	Integer	NNNNNNNN	8	Trading Volume for the trading day	23698686
Return	Floating (Number)	N.N	3	Returns (Increase/Decrease in Close Price) for the trading day	0.9
Class	String	-	4	Class (Upward/Downward movement of Close Price) for the trading day	Up
RSI	Floating (Number)	NN.N	4	Relative Strength Index on the trading day	31.2
SMA	Floating (Number)	NNN.N	5	Simple Moving Average on the trading day	296.3
LMA	Floating (Number)	NNN.N	5	Linearly Weighted Moving Average on the trading day	288.1
ADX	Floating (Number)	NN.N	4	Average Directional Index on the trading day	15.0

** The adjusted closing price uses the closing price as a starting point, but it takes into account factors such as dividends, stock splits and new stock offerings to determine a value.*

c) List of Abbreviations, Figures, Tables:

List of Abbreviations:

S.No.	Abbreviation	Expansion
1	+DI	Plus Directional Indicator
2	ADX	Average Directional Index
3	ATR	Average True Range
4	CPM	Critical Path Method
5	CPU	Central Processing Unit
6	CRAN	Comprehensive R Archive Network
7	CSS	Cascading Style sheets
8	CSV	Comma Separated Values
9	DFD	Data Flow Diagram
10	-DI	Minus Directional Indicator
11	F&R	Finance & Risk
12	GB	Gigabytes
13	GHz	Gigahertz
14	HFT	High Frequency Trading
15	HTML	Hypertext Markup Language
16	IIS	Internet Information Services
17	ISIN	International Securities Identification Number
18	JS	JavaScript
19	LMA	Linearly Weighted Moving Average
20	MACD	Moving Average Convergence/Divergence Oscillator
21	NIFTY	National Stock Exchange's Top 50 Stocks
22	NSE	National Stock Exchange of India Ltd.
23	OHLC	Open, High, Low, Close Price of stock
24	OHLVC	Open, High, Low, Volume, Close
25	PERT	Program Evaluation & Review Technique
26	PIECES Framework	Performance, Information, Economy, Control, Efficiency, Services Framework
27	RAM	Random Access Memory
28	RSI	Relative Strength Index
29	SMA	Simple Moving Average
30	SP500	Standard & Poor's 500
31	SVG	Scalable Vector Graphics
32	VBA	Visual Basic for Applications
33	XHTML	Extensible Hypertext Markup Language

List of Figures:

List of Tables:

d) References & Bibliography:

- Kimoto T, Asakawa K, Yoda M and Takeoka M (1990), Stock market prediction system with modular neural networks, Proc. International Joint Conference on Neural Networks, San Diego, Vol. 1, pp. 1-6.

- https://www.quantinsti.com/blog/predictive-modeling-algorithmic-trading/
- https://www.quantinsti.com/blog/free-resources-learn-machine-learning-trading/
- https://www.quantinsti.com/blog/use-decision-trees-machine-learning-predict-stock-movements/
- Abraham A, Nath B and Mahanti P K (2001), Hybrid Intelligent Systems for Stock Market Analysis,Proceedings of the International Conference on Computational Science. Springer, pp. 337-345
- https://www.quora.com/In-the-form-of-case-studies-how-do-financial-companies-use-machine-learning
- https://www.quantinsti.com/blog/machine-learning-application-forex-markets-working-models/
- https://www.quantinsti.com/machine-learning-webinar-15-june-2017/
- K. Tsai and J. Wang, "External technology sourcing and innovation performance in LMT sectors", Research Policy, vol. 38, no. 3, pp. 518-526, 2009.
- https://www.quora.com/How-do-trading-algorithms-work?ref=huffpo&rel_pos=1
- https://www.techemergence.com/machine-learning-in-finance/
- https://www.datacamp.com/community/tutorials/r-trading-tutorial#gs.iC7p0yo
- K. Han and J. Kim, "Genetic quantum algorithm and its application to combinatorial optimization problem", Evolutionary Computation, 2000, vol. 2, pp. 1354-1360, 2000.
- https://www.quantinsti.com/blog/machine-learning-trading-predict-stock-prices-regression/
- https://www.nseindia.com/products/content/equities/indices/nifty_50.htm
- https://in.finance.yahoo.com/quote/%5ENSEI/history?p=%5ENSEI
- Market Capitalization Definition from Financial Times Lexicon
- https://in.finance.yahoo.com/quote/ADANIPORTS.NS/history?p=ADANIPORTS.NS
- https://github.com/andredumas/techan.js/wiki/Gallery

- https://plot.ly/javascript/candlestick-charts/
- "MarketTrak's Forecast Model Overview", Markettrak.com, 2016
- https://austinschaub.com/2016/11/02/r-tutorial-scrape-html-table-using-hadleys-rvest-package/
- http://www.acit2k.org/ACIT/2013Proceedings/163.pdf
- https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Classification/Decision_Trees
- S. Hannon, "5 Rules For Prediciting Stock Market Trends - StockTrader.com", StockTrader.com, 2016
- https://arxiv.org/pdf/1301.4944.pdf
- https://www.quantinsti.com/blog/rexcel-tutorial-using-r-excel/
- SOFTWARE ENGINEERING (third edition) by K.K Agrawal, Yogesh Singh.