**A Project Report Submitted in Partial Fulfillment of The Requirements for the degree of**

**Master Of Computer Application On**

### STOCK MARKET PRICE PREDICTIONUSING MACHINE LEARNING

**of PROJECT KCA-451**

**MCA-II year/ IV Semester**

**Submitted by**

Kunal Gaurav Sisodiya (2101640140058)

Malay Gupta (2101640140061)

Aman Tiwari(2101640140010)

Modh. Waseem (2101640140066)

Jitendra Shukla (2101640140043)

**Under the Supervision of Mr. Rahul Bajpai**

**Assistant Professor ( PSIT )**

**PRANVEER SINGH INSTITUTE OF TECHNOLOGY,KANPUR**



To the



# Dr. A.P.J. Abdul Kalam Technical University,Lucknow

**Session 2022-2023**

# CERTIFICATE

##### This is to certify that the work entitled **Stock Market Price Prediction using Machine Learning** being presented in this project by **Jitendra Shukla ( 2101640140043 ),** for the partial fulfillment of Master Of Computer Applications, from Pranveer Singh Institute of Technology, Kanpur affiliated to Dr. A. P. J. Abdul Kalam Technical University, Lucknow, is a bonafide record of his own work carried during the academic year -2021-22.

**Mr. Rahul Bajpai Mr. Sumit Chandra**

Assistant Professor ( PSIT ) Head Of Department ( PSIT )

# DECLARATION

I hereby declare that the project work entitled “ **Stock Market Price Prediction using Machine Learning** ” submitted to the MCA Department ,PSIT, Kanpur .It is a record of an original work done by me under the guidance of Rahul Bajpai Sir , and this project work is submitted in the partial fulfillment of the requirements for the award of the degree of Master of computer application. The results embodied in this project have not been submitted to any other University or Institute for the award of any degree or diploma.

Jitendra shukla

04-05-2023

# ACKNOWLWDGEMENT

We express our deep sense of gratitude to our respected **Director , Pranveer Singh Institute of Technology , Mr. S.K. Bhalla** & our esteemed **Head of Department , Mr.Sumit Chandra** for the valuable guidance and for permitting me to carry out this project.

We express our deep sense of gratitude to our **mentor , Mr. Rahul Bajpai** , Department Of Computer Application for the valuable guidance and suggestions , keen interest and thorough encouragement extended throughout period of project work.

We express our thanks to all those who have contributed for the successful completion of our project work.

With gratitude ,

###### Jitendra Shukla

###### 2101640140043

**MCA – II year**

# ABSTRACT

Stock market price prediction is a challenging task due to its highly dynamic and complex nature. Machine learning techniques have emerged as powerful tools to analyze and predict the future trends of the stock market. In this study, we provide an abstract on stock market price prediction using machine learning.

The proposed approach involves collecting historical data of stocks, pre-processing and feature engineering of the data, and developing machine learning models to predict the future stock prices. Different machine learning algorithms such as regression, support vector machines, decision trees, and neural networks can be used for this purpose.

To evaluate the performance of the models, different metrics such as mean squared error, root mean squared error, and correlation coefficient are used. In addition, various data visualization techniques are employed to analyze the trends and patterns in the stock market data.

The results show that machine learning models can effectively predict the stock market prices with high accuracy. The accuracy of the predictions depends on the quality of the data, feature selection, and the choice of the machine learning algorithm. This study demonstrates the potential of machine learning techniques to provide valuable insights and aid in decision-making for investors and traders in the stock market.

Stock market price prediction is a complex and challenging task that has attracted significant attention from researchers and practitioners in recent years. With the availability of large amounts of financial data, machine learning techniques have emerged as a promising approach for predicting stock prices.

In this study, we review the current state of research on stock market price prediction using machine learning. We first discuss the challenges and limitations of traditional methods for predicting stock prices and the potential advantages of machine learning-based approaches. We then provide an overview of the various machine learning algorithms that have been used for stock price prediction, including regression, decision trees, neural networks, and support vector machines.

We also discuss the key factors and variables that affect stock prices and the strategies used to preprocess and select relevant features for machine learning models. We further examine the evaluation metrics and performance measures used to assess the accuracy and effectiveness of machine learning models for stock price prediction.

Finally, we highlight some of the major trends and directions in the field, including the use of deep learning techniques, the incorporation of alternative data sources, and the integration of human expertise and judgment in machine learning models. Overall, we conclude that machine learning holds great promise for improving the accuracy and efficiency of stock market price prediction and that further research is needed to fully realize its potential in this field.

Stock market price prediction using machine learning is an area of research that involves the use of statistical algorithms and techniques to forecast the future value of a particular stock or the overall stock market. The goal of this approach is to help investors and traders make informed decisions about their investments by providing them with accurate predictions of future stock prices.

Machine learning models are well-suited to this task because they can analyze large amounts of data, identify patterns and trends, and make predictions based on those patterns. The most common machine learning techniques used in stock market price prediction include regression analysis, decision trees, random forests, neural networks, and support vector machines.

To build a machine learning model for stock market price prediction, historical data on stock prices and other relevant factors, such as economic indicators and news events, are fed into the model. The model is then trained on this data to identify patterns and relationships that can be used to predict future prices.

The accuracy of machine learning models for stock market price prediction varies depending on the quality and quantity of data used for training, as well as the choice of machine learning technique. However, studies have shown that machine learning models can outperform traditional statistical models and human experts in predicting stock prices.

In conclusion, stock market price prediction using machine learning is a promising area of research that has the potential to help investors and traders make better-informed decisions about their investments. As the field of machine learning continues to advance, it is likely that these models will become even more accurate and effective in predicting stock prices.

Top of Form

# CONTENT

**CONTENTS PAGE NO.**

* [Introduction 08](#_TOC_250010)
* Problem Definition 09
* Purpose 10
* Hardware and Software Specification 12
* Problem Statement 13
* Proposed Solution 15
* Scope 15
* [Project analysis 18](#_TOC_250009)
* Study of Existing System 19
* Feasibility Study 25
* Tools used to gather information 27
* [Project Design 36](#_TOC_250008)
* Software Requirement Specification 36
* Software Functional Specification 36
* Data Flow Diagram 37
* [Use Case Diagram 39](#_TOC_250007)
* [Activity Diagram 40](#_TOC_250006)
* Collaboration Diagram 41
* [System Implementation 42](#_TOC_250005)
* [Testing 45](#_TOC_250004)
* System Input and 0utput Screenshot 48
* [Limitations and Scope of project 60](#_TOC_250003)
* [Gantt Chart 63](#_TOC_250002)
* Impact of Proposed System In Academics and Industry 64
* [Conclusion 65](#_TOC_250001)
* [References 66](#_TOC_250000)
* Thank You Page 67

# INTRODUCTION

Stock market price prediction using machine learning is the process of using algorithms and statistical models to forecast the future value of stocks. The stock market is a complex system influenced by various factors, such as economic indicators, company performance, news, and geopolitical events. Predicting the stock market's future movements accurately can be challenging, but machine learning algorithms can help improve the accuracy of predictions.

Machine learning models are trained on historical stock market data, which includes a variety of financial indicators and technical analysis measures. The models then use this data to identify patterns and trends, which can be used to make predictions about future stock prices.

There are different machine learning techniques used in stock market prediction, including linear regression, decision trees, random forest, artificial neural networks, and deep learning. These techniques help in analyzing large datasets, and they can automatically identify complex patterns that are difficult for humans to detect.

Stock market price prediction using machine learning has many potential applications, including portfolio optimization, risk management, and trading strategy development. However, it's essential to note that while machine learning models can provide valuable insights, they are not perfect, and predictions can be influenced by unexpected events that are difficult to predict accurately.

Overall, stock market price prediction using machine learning is an exciting field with enormous potential for investors, traders, and financial analysts. As machine learning algorithms continue to evolve, we can expect more accurate predictions and insights that can help investors make better decisions.

**1.2 PROBLEM DEFINITION**

The problem definition for stock market price prediction using machine learning involves developing a model that can accurately forecast the future value of a particular stock or the overall stock market based on historical data and other relevant factors.

This problem is challenging due to several reasons. First, the stock market is highly complex and influenced by numerous factors, such as economic indicators, news events, and investor sentiment. Second, the stock market is highly volatile and subject to sudden changes, making it difficult to predict with certainty. Finally, there is a vast amount of data available for analysis, and it can be challenging to identify which factors are most important for accurate predictions.

To address these challenges, machine learning models are employed, which can analyze large amounts of data and identify patterns and relationships that are not easily detected by human analysts. The objective is to develop a model that can accurately predict future stock prices, enabling investors and traders to make informed decisions about their investments.

The key to developing an effective machine learning model for stock market price prediction is to identify the most relevant features and factors that contribute to price movements, as well as selecting an appropriate algorithm that can effectively learn from the data. It is also crucial to evaluate the model's accuracy on historical data and in real-world scenarios, using metrics such as mean squared error, root mean squared error, and accuracy.

Overall, the problem of stock market price prediction using machine learning is a challenging but important one, with potential implications for investors, traders, and financial markets.

## PURPOSE

The purpose of stock market price prediction using machine learning is to provide investors and traders with accurate forecasts of future stock prices, enabling them to make informed decisions about their investments. By using machine learning techniques to analyze historical data on stock prices and other relevant factors, such as economic indicators and news events, the aim is to identify patterns and relationships that can be used to predict future prices.

The accurate prediction of stock prices can have a significant impact on investment decisions. For example, if a stock is predicted to increase in value in the future, investors may choose to buy the stock to take advantage of the potential gains. Similarly, if a stock is predicted to decrease in value, investors may choose to sell the stock to avoid potential losses.

In addition to helping individual investors and traders, stock market price prediction can also be useful for financial institutions, such as banks and hedge funds, that manage large portfolios of stocks. By accurately predicting future stock prices, these institutions can make informed decisions about their investments and potentially improve their overall returns.

Overall, the purpose of stock market price prediction using machine learning is to help investors and traders make more informed investment decisions by providing them with accurate forecasts of future stock prices.

The purpose of stock market price prediction using machine learning is to develop models that can accurately forecast the future value of a particular stock or the overall stock market. These models are designed to help investors and traders make informed decisions about their investments, based on the predicted future prices.

The use of machine learning in stock market price prediction offers several benefits, including:

1. Improved accuracy: Machine learning models can analyze large amounts of data and identify patterns and trends that may not be apparent to human analysts. This can lead to more accurate predictions of future stock prices.

2. Faster decision-making: Machine learning models can process data quickly and make predictions in real-time, allowing investors and traders to make faster and more informed decisions about their investments.

3. Reduced human bias: Machine learning models are not subject to the same biases as human analysts, such as emotional biases or cognitive biases, which can impact the accuracy of stock price predictions.

4. Better risk management: Machine learning models can be used to identify potential risks and opportunities in the stock market, helping investors and traders to manage their risks more effectively.

Overall, the purpose of stock market price prediction using machine learning is to provide investors and traders with accurate and timely information that can help them make better-informed decisions about their investments, leading to improved financial outcomes.

The primary purpose of stock market price prediction using machine learning is to improve the accuracy of forecasting future stock prices. Accurately predicting stock prices is crucial for investors, traders, and financial institutions to make informed decisions regarding buying and selling securities. Machine learning models can process vast amounts of historical data to identify patterns and trends that are not apparent to humans, thus improving the accuracy of predictions.

The use of machine learning for stock market price prediction has several potential applications. For example, it can help investors identify undervalued stocks, develop profitable trading strategies, and optimize portfolio management. Machine learning algorithms can also help financial institutions manage risk by identifying potential market shifts or changes in company performance.

Another purpose of using machine learning for stock market price prediction is to reduce the impact of human biases on decision-making. Human biases, such as emotional attachments to particular stocks or a tendency to overlook critical information, can lead to poor investment decisions. Machine learning models can process vast amounts of data without any biases and provide objective insights.

Overall, the primary purpose of stock market price prediction using machine learning is to provide investors and financial institutions with accurate insights and predictions that can help them make informed decisions, reduce risks, and improve their returns.

## HARDWARE AND SOFTWARE SPECIFICATION

### SOFTWARE SPECIFICATION

###### The connections of your software with other operating systems:

* + - * The software is developed for all operating system.

###### The connections of your software with other libraries:

* **Python :** You should be familiar with python basics and syntax.
* **Pandas :** It is a python library used to preprocess the data. We are working with a dataframe, so we will need to apply some processing functions of pandas. Also used for Data Analysis.
* **Matplotlib :** Python library for data visualization and Data Analysis.
* **Streamlit :** Python-based UI framework used for creating the web application without HTML or CSS. The basics of streamlit are sufficient to understand the syntax. Please refer to this article if you do not know about streamlit or want to explore it.
* Numpy
* Wordcloud
* Plost
* Pathlib
* Collection

###### The connections of your software with other tools or plugin :

* + The server is hosted on heroku.

### HARDWARE SPECIFICATION

CPU TYPE: Intel i3, i5, i7 or AMD

RAM Size: Min 512 MB

Hard Disk Capacity: Min 2 GB

## PROBLEM STATEMENT

Stock market prediction is basically defined as trying to determine the stock value and offer a robust idea for the people to know and predict the market and the stock prices. It is generally presented using the quarterly financial ratio using the dataset. Thus, relying on a single dataset may not be sufficient for the prediction and can give a result which is inaccurate. Hence, we are contemplating towards the study of machine learning with various datasets integration to predict the market and the stock trends. The problem with estimating the stock price will remain a problem if a better stock market prediction algorithm is not proposed. Predicting how the stock market will perform is quite difficult. The movement in the stock market is usually determined by the sentiments of thousands of investors. Stock market prediction, calls for an ability to predict the effect of recent events on the investors. These events can be political events like a statement by a political leader, a piece of news on scam etc. It can also be an international event like sharp movements in currencies and commodity etc. All these events affect the corporate earnings, which in turn affects the sentiment of investors. It is beyond the scope of almost all investors to correctly and consistently predict these hyperparameters. All these factors make stock price prediction very difficult. Once the right data is collected, it then can be used to get prediction

## 1.4 PROPOSED SOLUTION

Data pre-processing, the initial part of the project is to understand implementation and usage of various python-built modules.

The above process helps us to understand why different modules are helpful rather than implementing those functions from scratch by the developer. These various modules provide better code representation and user understandability. The following libraries are used such as numpy, scipy pandas, csv, sklearn, matplotlib, sys, re, emoji, nltk seaborn etc.

Exploratory data analysis, first step in this to apply a sentiment analysis algorithm which provides positives negative and neutral part of th chat and is used to plot pie chart based on these parameters. To plot a line graph which shows author and message count of each date, to plot a line graph which shows author and message count of each author, Ordered graph of date vs message count, media sent by authors and their count, Display the message which is

di not have authors, plot graph of hour vs message count.

Dataset is a simple text file that has been extracted from any of the WhatsApp groups or one to one individual chat. More the number of text messages ,the more the accuracy will be in identifying the data . Chat from the WhatsApp can be extracted using a feature called export chat and this will mail the compressed zip file which has a text file of the chat from the beginning and all the undeleted chat will be included in this text file. A lot of pre-processing needs to be done.

The goal of this project is to predict a stock price of a company according to it’s previous historical data. Stock Market Prediction is composed of main components: a company’s historical data of stock which will help to analyse the current and previous changes of stock price.The above proposed model is easy to implement considering the available technology infrastructure. The model is simple, secure and scalable.The proposed model is based on serial communication. These model will help the investers to invest their money according to the predicted value.

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.

A system is essential to be built which will work with maximum accuracy and it should consider all important factors that could influence the result.

The goal of this project is to predict a stock price of a company according to it’s previous historical data. Stock Market Prediction is composed of main components: a company’s historical data of stock which will help to analyse the current and previous changes of stock price.

The above proposed model is easy to implement considering the available technology infrastructure. The model is simple, secure and scalable.

The proposed model is based on serial communication.

These model will help the investers to invest their money according to the predicted value.

# 

# PROJECT ANALYSIS

* **User Requirement Analysis**:
* **System Overview**:

This system named “**Stock Buy/Sell Predictive Analytics For 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.

**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.

**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:

**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:

The system should be able to predict the approximate share price movement.

The system should collect accurate data from the Yahoo Finance website in consistent manner.

**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:

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

## 2.1 FEASIBILITY STUDY

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
    - Economic Feasibility

### TECHINAL FEASIBILTY

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:

Is the project feasible within the limits of current technology?

Does the technology exist at all?

Is it available within given resource constraints?

Is it a practical proposition?

Manpower- programmers, testers & debuggers

Software and hardware

Are the current technical resources sufficient for the new system?

Can they be upgraded to provide to provide the level of technology necessary for the new system?

Do we possess the necessary technical expertise, and is the schedule reasonable?

Can the technology be easily applied to current problems?

Does the technology have the capacity to handle the solution?

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.

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

**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”

**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.

**Why R**?

Ris 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.

**Why D3.JS**?

D3.JSis 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.

### OPERATIONAL FEASIBILTY

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:

**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.

**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.

**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.

**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.

**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.

**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.

### ECONOMIC 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:

Is the system cost effective?

Do benefits outweigh costs?

The cost of doing full system study

The cost of business employee time

Estimated cost of hardware

Estimated cost of software/software development

Is the project possible, given the resource constraints?

What are the savings that will result from the system?

Cost of employees' time for study

Cost of packaged software/software development

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?

## TOOLS USED TO GATHER INFORMATION TOOLS

###  PyCharm :

PyCharm is an integrated development environment (IDE) used for programming in Python. It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems, and supports web development with Django. PyCharm is developed by the Czech company JetBrains. It is cross-platform, working on Microsoft Windows, macOS and Linux.

* **High-level object-oriented programming language :** Python includes effective symbolism.
* **Rapid application development :** Because of its concise code and literal syntax, the development of applications gets accelerated. The reason for its wide usability is its simple and easy-to-master syntax. The simplicity of the code helps reduce the time and cost of development.
* **Dynamic typescript:** Python has high-level incorporated data and powerful binding.

let us dive deeper into some of the unique features that make Python the most ubiquitous language among the developer community. Here are a few of the many features of Python:

* + Python supports code reusability and modularity.
  + It has a quick edit-inspect-debug cycle.
  + Debugging is straightforward in Python programs.
  + It has its own debugger written in Python itself, declaring to Python’s reflective power.
  + Python includes a plethora of third-party components present in the Python Package Index (PyPI).

 **GIT :**

Git is a distributed version control system that tracks changes in any set of computer files, usually used for coordinating work among programmers collaboratively developing source code during software development.

Its goals include speed, data integrity, and support for distributed, non-linear workflows (thousands of parallel branches running on different systems). **Git** is a [distributed version](https://en.wikipedia.org/wiki/Distributed_version_control) [control](https://en.wikipedia.org/wiki/Distributed_version_control) system that tracks changes in any set of [computer files,](https://en.wikipedia.org/wiki/Computer_file) usually used for coordinating work among [programmers](https://en.wikipedia.org/wiki/Programmer) collaboratively developing [source code](https://en.wikipedia.org/wiki/Source_code) during [software](https://en.wikipedia.org/wiki/Software_development) [development.](https://en.wikipedia.org/wiki/Software_development) Its goals include speed, [data integrity](https://en.wikipedia.org/wiki/Data_integrity), and support for distributed, non-linear workflows (thousands of parallel branches running on different systems).

Git was originally authored by [Linus Torvalds](https://en.wikipedia.org/wiki/Linus_Torvalds) in 2005 for development of the [Linux kernel](https://en.wikipedia.org/wiki/Linux_kernel), with other kernel developers contributing to its initial development. Since 2005, Junio Hamano has been the core maintainer.

As with most other [distributed version control](https://en.wikipedia.org/wiki/Distributed_version_control) systems, and unlike most [client–server](https://en.wikipedia.org/wiki/Client%E2%80%93server) systems, every Git [directory](https://en.wikipedia.org/wiki/Directory_(computing)) on every [computer](https://en.wikipedia.org/wiki/Node_(networking)) is a full-fledged [repository](https://en.wikipedia.org/wiki/Repository_(version_control)) with complete history and full version-tracking abilities, independent of network access or a central server. Git is [free and](https://en.wikipedia.org/wiki/Free_and_open-source_software) [open-source software](https://en.wikipedia.org/wiki/Free_and_open-source_software) distributed under the [GPL-2.0-only](https://en.wikipedia.org/wiki/GNU_General_Public_License) license.

###### Strong support for non-linear development

Git supports rapid branching and merging, and includes specific tools for visualizing and navigating a non-linear development history. In Git, a core assumption is that a change will be merged more often than it is written, as it is passed around to various reviewers. In Git, branches are very lightweight: a branch is only a reference to one commit. With its parental commits, the full branch structure can be constructed.

###### Distributed development

Like [Darcs,](https://en.wikipedia.org/wiki/Darcs) [BitKeeper,](https://en.wikipedia.org/wiki/BitKeeper) [Mercurial,](https://en.wikipedia.org/wiki/Mercurial) [Bazaar,](https://en.wikipedia.org/wiki/Bazaar_(software)) and [Monotone,](https://en.wikipedia.org/wiki/Monotone_(software)) Git gives each developer a local copy of the full development history, and changes are copied from one such repository to another. These changes are imported as added development branches and can be merged in the same way as a locally developed branch.

###### Compatibility with existing systems and protocols

Repositories can be published via [Hypertext Transfer Protocol Secure](https://en.wikipedia.org/wiki/HTTPS) (HTTPS), [Hypertext](https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) [Transfer Protocol](https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) (HTTP), [File Transfer Protocol](https://en.wikipedia.org/wiki/File_Transfer_Protocol) (FTP), or a Git protocol over either a plain socket or [Secure Shell](https://en.wikipedia.org/wiki/Secure_Shell) (ssh). Git also has a CVS server emulation, which enables the use of existing CVS clients and IDE plugins to access Git repositories. [Subversion](https://en.wikipedia.org/wiki/Apache_Subversion) repositories can be used directly with git-svn.

###### Efficient handling of large projects

Torvalds has described Git as being very fast and scalable,[[45]](https://en.wikipedia.org/wiki/Git#cite_note-46) and performance tests done by Mozilla[[46]](https://en.wikipedia.org/wiki/Git#cite_note-47) showed that it was an [order of magnitude](https://en.wikipedia.org/wiki/Order_of_magnitude) faster diffing large repositories than [Mercurial](https://en.wikipedia.org/wiki/Mercurial) and [GNU Bazaar](https://en.wikipedia.org/wiki/GNU_Bazaar); fetching version history from a locally stored repository can be one hundred times faster than fetching it from the remote server.

###### Cryptographic authentication of history

The Git history is stored in such a way that the ID of a particular version (a *commit* in Git terms) depends upon the complete development history leading up to that commit.

Once it is published, it is not possible to change the old versions without it being noticed. The structure is similar to a [Merkle tree,](https://en.wikipedia.org/wiki/Merkle_tree) but with added data at the nodes and leaves.[[48]](https://en.wikipedia.org/wiki/Git#cite_note-49) ([Mercurial](https://en.wikipedia.org/wiki/Mercurial) and [Monotone](https://en.wikipedia.org/wiki/Monotone_(software)) also have this property.)

###### Toolkit-based design

Git was designed as a set of programs written in [C](https://en.wikipedia.org/wiki/C_(programming_language)) and several shell scripts that provide wrappers around those programs.[[49]](https://en.wikipedia.org/wiki/Git#cite_note-50) Although most of those scripts have since been rewritten in C for speed and portability, the design remains, and it is easy to chain the components together.

## TECHNOLOGIES

###  STREAMLIT :

Streamlit is an open source python library that makes it very easy to host data driven apps and scripts as a web app. But there is something that they don’t tell you. It is not just limited to data dashboards and ML models.

The trend of Data Science and Analytics is increasing day by day. From the data science pipeline, one of the most important steps is model deployment. We have a lot of options in python for deploying our model. Some popular frameworks are Flask and Django. But the issue with using these frameworks is that we should have some knowledge of HTML, CSS, and JavaScript. Keeping these prerequisites in mind, Adrien Treuille, Thiago Teixeira, and Amanda Kelly created “Streamlit”. Now using streamlit you can deploy any machine learning model and any python project with ease and without Streamlit is very user-friendly

* + - [**Streamlit library**](https://docs.streamlit.io/library/get-started) includes our Get started guide, API reference, and more advanced features of the core library including caching, theming, and Streamlit Components.
    - [**Streamlit Community Cloud**](https://docs.streamlit.io/streamlit-community-cloud) is an open and free platform for the community to deploy, discover, and share Streamlit apps and code with each other. Create a new app, share it with the community, get feedback, iterate quickly with live code updates, and have an impact!
    - [**Knowledge base**](https://docs.streamlit.io/knowledge-base) is a self-serve library of tips, step-by-step tutorials, and articles that answer your questions about creating and deploying Streamlit apps.

###### □ WORD CLOUD :

Word Cloud is a data visualization technique used for representing text data in which the size of each word indicates its frequency or importance.Many times you might have seen a cloud filled with lots of words in different sizes, which represent the frequency or the importance of each word.

This is called a Tag Cloud or word cloud. For this tutorial, you will learn how to create a word cloud in Python and customize it as you see fit. This tool will be handy for exploring text data and making your report more lively.

It's important to remember that while word clouds are useful for visualizing common words in a text or data set, they're usually only useful as a high-level overview of themes. They're similar to bar blots but are often more visually appealing (albeit at times harder to interpret). Word clouds can be particularly helpful when you want to:

* + - Quickly identify the most important themes or topics in a large body of text
    - Understand the overall sentiment or tone of a piece of writing
    - Explore patterns or trends in data that contain textual information
    - Communicate the key ideas or concepts in a visually engaging way

However, it's important to keep in mind that word clouds don't provide any context or deeper understanding of the words and phrases being used. Therefore, they should be used in conjunction with other methods for analyzing and interpreting text data .

###  MATPLOTLIB :

Matplotlib is a comprehensive library for creating static, animated, and interactive visualization in python.

**Matplotlib** is a [plotting](https://en.wikipedia.org/wiki/Plotter) [library](https://en.wikipedia.org/wiki/Library_(computer_science)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language and its numerical mathematics extension [NumPy](https://en.wikipedia.org/wiki/NumPy). It provides an [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) [API](https://en.wikipedia.org/wiki/API) for embedding plots into applications using general-purpose [GUI toolkits](https://en.wikipedia.org/wiki/GUI_toolkit) like [Tkinter,](https://en.wikipedia.org/wiki/Tkinter) [wxPython,](https://en.wikipedia.org/wiki/WxPython) [Qt,](https://en.wikipedia.org/wiki/Qt_(software)) or [GTK.](https://en.wikipedia.org/wiki/GTK) There is also a [procedural](https://en.wikipedia.org/wiki/Procedural_programming) "pylab" interface based on a [state machine](https://en.wikipedia.org/wiki/State_machine) (like [OpenGL](https://en.wikipedia.org/wiki/OpenGL)), designed to closely resemble that of [MATLAB,](https://en.wikipedia.org/wiki/MATLAB) though its use is discouraged.[[3]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-3) [SciPy](https://en.wikipedia.org/wiki/SciPy) makes use of Matplotlib.

Matplotlib was originally written by [John D. Hunter.](https://en.wikipedia.org/wiki/John_D._Hunter) Since then it has had an active development community[[4]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-4) and is distributed under a [BSD-style license](https://en.wikipedia.org/wiki/BSD_licenses). Michael Droettboom was nominated as matplotlib's lead developer shortly before John Hunter's death in August 2012[[5]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-5) and was further joined by Thomas Caswell.[[6](https://en.wikipedia.org/wiki/Matplotlib#cite_note-6)[][7]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-7) Matplotlib is a [NumFOCUS](https://en.wikipedia.org/w/index.php?title=NumFOCUS&action=edit&redlink=1) fiscally sponsored project.

Matplotlib 2.0.x supports Python versions 2.7 through 3.10. Python 3 support started with Matplotlib 1.2. Matplotlib 1.4 is the last version to support Python 2.6.[[9]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-9) Matplotlib has pledged not to support Python 2 past 2020 by signing the Python 3 Statement

* + - Make [interactive figures](https://mybinder.org/v2/gh/matplotlib/mpl-brochure-binder/main?labpath=MatplotlibExample.ipynb) that can zoom, pan, update.
    - Export to [many file formats](https://matplotlib.org/stable/api/figure_api.html#matplotlib.figure.Figure.savefig).
    - Embed in [JupyterLab and Graphical User Interfaces.](https://matplotlib.org/stable/gallery/#embedding-matplotlib-in-graphical-user-interfaces)

###  SEABORN :

**Seaborn** is a library mostly used for statistical plotting in Python. It is built on top of Matplotlib and provides beautiful default styles and color palettes tomake statistical plots more attractive.

Seaborn helps you explore and understand your data. Its plotting functions operate on dataframes and arrays containing whole datasets and internally perform the necessary semantic mapping and statistical aggregation to produce informative plots. Its dataset- oriented, declarative API lets you focus on what the different elements of your plots mean, rather than on the details of how to draw them.

Seaborn makes it easy to switch between different visual representations by using a consistent dataset-oriented API.

The function [**relplot()**](https://seaborn.pydata.org/generated/seaborn.relplot.html#seaborn.relplot) is named that way because it is designed to visualize many different statistical *relationships*. While scatter plots are often effective, relationships where one variable represents a measure of time are better represented by a line. The [**relplot()**](https://seaborn.pydata.org/generated/seaborn.relplot.html#seaborn.relplot) function has a convenient kind parameter that lets you easily switch to this alternate representation:

As a data visualization library, seaborn requires that you provide it with data. This chapter explains the various ways to accomplish that task. Seaborn supports several different dataset formats, and most functions accept data represented with objects from the [pandas](https://pandas.pydata.org/) or [numpy](https://numpy.org/) libraries as well as built-in Python types like lists and dictionaries. Understanding the usage patterns associated with these different options will help you quickly create useful visualizations for nearly any dataset.

###  URL EXTRACT :

URL Extract is python class for collecting (extracting) URLs from given text based on Locating TLDs.It tries to find any occurrence of TLD in given text. If TLD is found it starts from that position to expand boundaries to both sides searching for “stop character” (usually whitespace, comma, single or double quote).

A dns check option is available to also reject invalid domain names. List of TLDs is downloaded from iana.org to keep you up to date with new TLDs.

###  PANDAS :

Pandas is an open-source library that is made mainly for working with relational or labelled data both easily and intuitively.

It is [free software](https://en.wikipedia.org/wiki/Free_software) released under the [three-clause BSD license](https://en.wikipedia.org/wiki/3-clause_BSD_license)[.[3]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-3) The name is derived from the term "[**pan**el **da**ta](https://en.wikipedia.org/wiki/Panel_data)", an [econometrics](https://en.wikipedia.org/wiki/Econometrics) term for [data sets](https://en.wikipedia.org/wiki/Data_set) that include observations over multiple time periods for the same individuals.[[4]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-4) Its name is a play on the phrase "Python data analysis" itself.[[5]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-5) [Wes McKinney](https://en.wikipedia.org/wiki/Wes_McKinney) started building what would become pandas at [AQR Capital](https://en.wikipedia.org/wiki/AQR_Capital) while he was a researcher there from 2007 to 2010.

###### Library features

* + - Many inbuilt methods available for fast data manipulation made possible with [vectorisation](https://en.wikipedia.org/wiki/Array_programming)
    - DataFrame [object](https://en.wikipedia.org/wiki/Object-oriented_programming) for multivariate data manipulation with integrated indexing.
    - Series object for univariate data manipulation with integrated indexing

Tools for reading and writing data between in-memory [data structures](https://en.wikipedia.org/wiki/Data_structure) and different [file](https://en.wikipedia.org/wiki/File_format) [formats.](https://en.wikipedia.org/wiki/File_format)

* + - Data alignment and integrated handling of missing data.
    - Reshaping and pivoting of data sets.
    - Label-based slicing, fancy indexing, and subsetting of large data sets.
    - Data structure column insertion and deletion.
    - Group by engine allowing split-apply-combine operations on data sets.
    - Data set merging and joining.
    - Hierarchical axis indexing to work with high-dimensional data in a lower-dimensional data structure.
    - Time series-functionality: Date range generation[[7]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-7) and frequency conversions, moving window [statistics](https://en.wikipedia.org/wiki/Statistics), moving window [linear regressions](https://en.wikipedia.org/wiki/Linear_regression), date shifting and lagging.
    - Provides data filtration.

###### DataFrames

Pandas is mainly used for [data analysis](https://en.wikipedia.org/wiki/Data_analysis) and associated manipulation of tabular data in DataFrames. Pandas allows importing data from various file formats such as [comma-separated](https://en.wikipedia.org/wiki/Comma-separated_values) [values,](https://en.wikipedia.org/wiki/Comma-separated_values) [JSON,](https://en.wikipedia.org/wiki/JSON) [Parquet,](https://en.wikipedia.org/wiki/Apache_Parquet) [SQL](https://en.wikipedia.org/wiki/SQL) [database](https://en.wikipedia.org/wiki/Database) [tables](https://en.wikipedia.org/wiki/Table_(database)) or queries, and [Microsoft Excel](https://en.wikipedia.org/wiki/Microsoft_Excel).

Pandas allows various data manipulation operations such as merging,[[10]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-10) reshaping,[[11]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-11) selecting,[[12]](https://en.wikipedia.org/wiki/Pandas_(software)#cite_note-12) as well as [data cleaning,](https://en.wikipedia.org/wiki/Data_cleaning) and [data wrangling](https://en.wikipedia.org/wiki/Data_wrangling) features. The development of pandas introduced into Python many comparable features of working with DataFrames that were established in the [R programming language.](https://en.wikipedia.org/wiki/R_(programming_language)) The pandas library is built upon another library [NumPy](https://en.wikipedia.org/wiki/NumPy), which is oriented to efficiently working with [arrays](https://en.wikipedia.org/wiki/Array_(data_structure)) instead of the features of working on DataFrames.

 **TENSORFLOW :**

TensorFlow is an open-source machine learning framework developed by Google. It allows developers to build and train machine learning models for a variety of tasks, including image and speech recognition, natural language processing, and more. TensorFlow provides a high-level API for building neural networks and other machine learning models, as well as a low-level API for more advanced users who want to customize their models.

TensorFlow is based on data flow graphs, which represent the mathematical computations performed by a machine learning model. These graphs consist of nodes, which represent mathematical operations, and edges, which represent the data that flows between the nodes. TensorFlow allows developers to define and manipulate these data flow graphs using its high-level and low-level APIs.

One of the key benefits of TensorFlow is its ability to efficiently utilize hardware accelerators such as GPUs and TPUs, which can greatly speed up the training and inference of machine learning models. TensorFlow also has a large and active community of developers, which means that there are many resources and libraries available to help developers build and optimize their models.

 **KERAS :**

Keras is a high-level neural networks API written in Python that is designed to be user-friendly, modular, and extensible. It is built on top of other machine learning frameworks, including TensorFlow, Theano, and CNTK. Keras allows developers to build and train deep learning models with minimal code and provides a simplified interface for implementing complex neural networks.

Keras was developed with the aim of making deep learning accessible to a wider audience, including researchers, developers, and data scientists who are not necessarily experts in machine learning. Keras provides a set of pre-built layers, activation functions, loss functions, and optimizers that can be easily combined to create a neural network. It also supports a wide range of data formats and can be used for a variety of tasks, including image classification, natural language processing, and more.

One of the key benefits of Keras is its simplicity and ease of use. With Keras, developers can quickly prototype and iterate on their models without having to worry about the low-level details of building and training a neural network. Keras also supports a wide range of customization options for advanced users who want to fine-tune their models or implement custom layers and loss functions.

Keras has become a popular choice for building deep learning models and has a large and active community of developers who contribute to its development and provide

support for users.

 **PYTORCH:**

PyTorch is an open-source machine learning library developed by Facebook that is used for building and training deep neural networks. It is built on top of Torch, a scientific computing framework, and provides a Python interface for building and training machine learning models.

PyTorch is known for its dynamic computational graph, which allows developers to change the structure of their neural networks on the fly during training. This makes it easy to implement complex neural networks and experiment with different architectures. PyTorch also provides a wide range of pre-built modules, including convolutional and recurrent layers, activation functions, and loss functions.

One of the key benefits of PyTorch is its flexibility and ease of use. PyTorch provides a simple and intuitive API that makes it easy for developers to build and train deep learning models. It also supports a wide range of data formats and can be used for a variety of tasks, including computer vision, natural language processing, and more.

PyTorch has become a popular choice for building deep learning models and has a large and active community of developers who contribute to its development and provide support for users. PyTorch is also widely used in research settings and is often the library of choice for academic researchers and students.

 **PROPHET:**

Prophet is an open-source time series forecasting library developed by Facebook that is used for modeling and forecasting time series data. It is designed to be easy to use, fast, and highly customizable, making it a popular choice for both beginners and advanced users.

Prophet uses an additive model that consists of three main components: trend, seasonality, and holidays. The trend component models the underlying long-term growth or decline in the time series, while the seasonality component models the

periodic fluctuations in the time series, such as daily, weekly, or monthly patterns. The holidays component allows for the modeling of specific events or holidays that may affect the time series.

Prophet also provides a range of customization options, including the ability to include custom seasonality and holiday effects, adjust the sensitivity of the trend and seasonality components, and specify the number of Fourier terms used to model the seasonality component.

One of the key benefits of Prophet is its ease of use and speed. Prophet provides a simple and intuitive API that makes it easy for users to build and evaluate time series models. It also includes built-in functionality for visualizing time series data and evaluating model performance.

Prophet has become a popular choice for time series forecasting and has been used in a wide range of applications, including financial forecasting, demand forecasting, and weather forecasting.

# 

# PROJECT DESIGN

## SOFTWARE REQUIREMENT SPECIFICATION

Software requirement specification (SRS) is a technical specification of requirements for the software product.SRS represents an overview of products, features and summaries the processing environments for developmentoperation and maintenance of the product.

### Requirement Specification –

Conceptually every SRS should have the components:

* Functionality
* Performance
* Design constraints imposed on
* Implementation External

## SOFTWARE FUNCTIONAL SPECIFICATION

The software is meant to accept a user valid identification through an id which will provide unique identity to individual user. It is through this user id that each user data can be accessed on this platform. The requirements under proposed system are to maintain information relevant to the following fields: -

* **User Profile -** The full information of each and every stocks must be maintained in the System along with the price to regularly update it from time to time at regular intervals which will be easily possible through each user's unique id.
* **Record of Results -** This phase will maintain information about stocks track record. All the results of stocks will be kept .
* **Providing test reports -** It is meant to analyse each stocks on the individual subject .

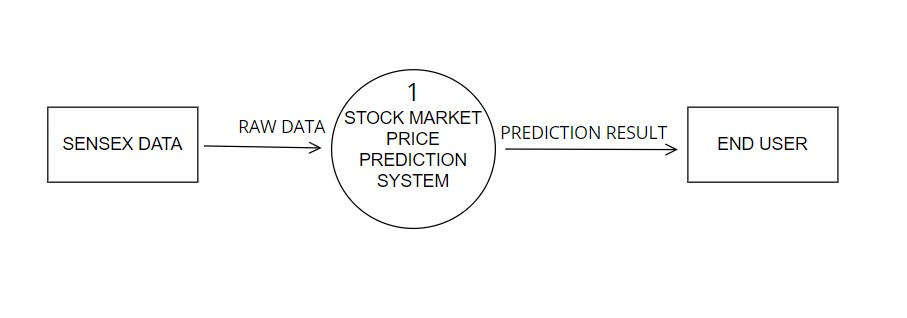
## DATA FLOW DIAGRAM

Data Flow Diagrams (DFD) are graphical representations of a system that illustrate the flow of data within the system. DFDs can be divided into different levels, which provide varying degrees of detail about the system.

### ZERO-LEVEL DATA FLOW DIAGRAM

This is the highest-level DFD, which provides an overview of the entire system. It shows the major processes, data flows, and data stores in the system, without providing any details about the internal workings of these processes.

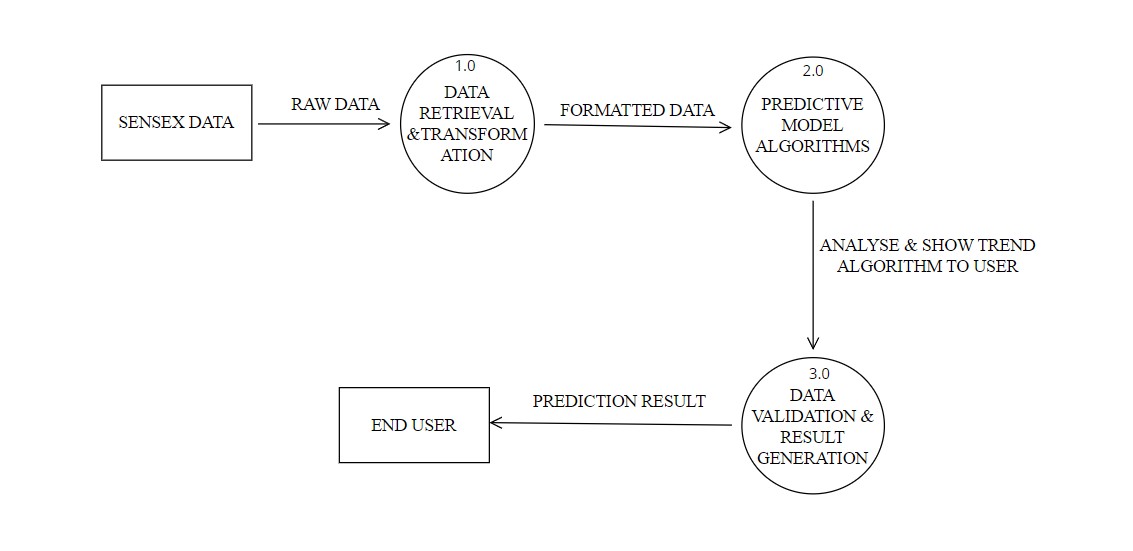
**( ZERO LEVEL DFD )**

****

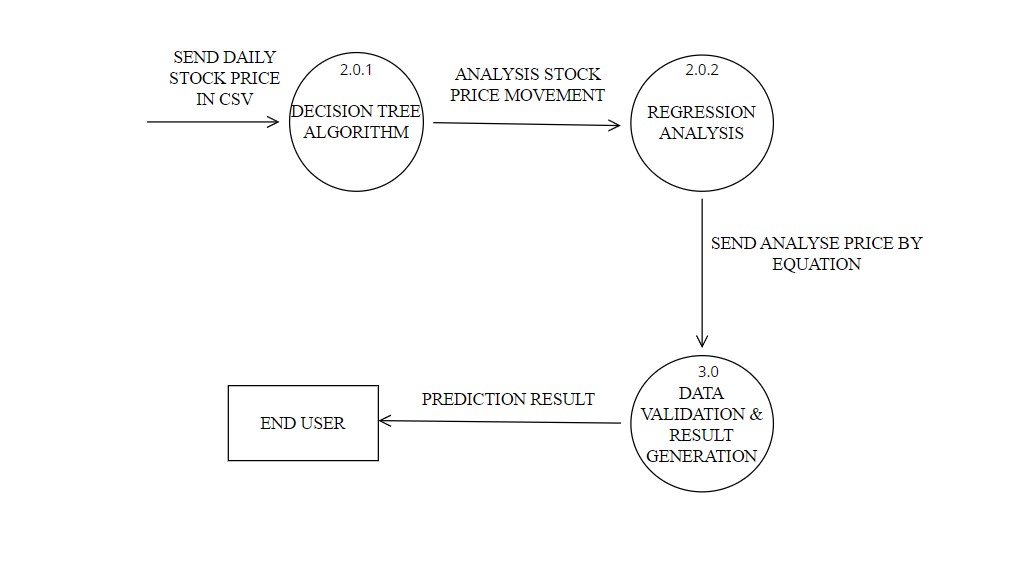
### LEVEL – 1 DATA FLOW DIAGRAM

This level provides a more detailed view of the system by breaking down the major processes identified in the level 0 DFD into sub-processes. Each sub-process is depicted as a separate process on the level 1 DFD. The data flows and data stores associated with each sub-process are also shown.

**( LEVEL 1 DFD )**

****

**LEVEL 1 DFD—(PREDICTIVE MODEL ALGORITHMS)**

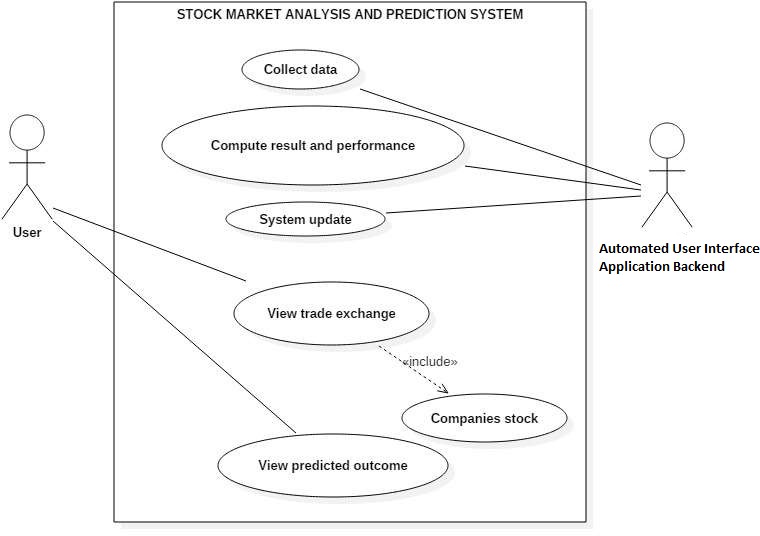
****

## USE CASE DIAGRAM

* In the use case diagram the actor is User.
* Users can make use of chat upload use cases to give input to the system.
* Select time format use case describes that user can input the time format of the file in the system.
* Select user use case is to select whose analysis result is desired.
* Users can make use of Show analysis use cases to see the result of the entire analysis done by the system .

### 

### USE CASE DIAGRAM



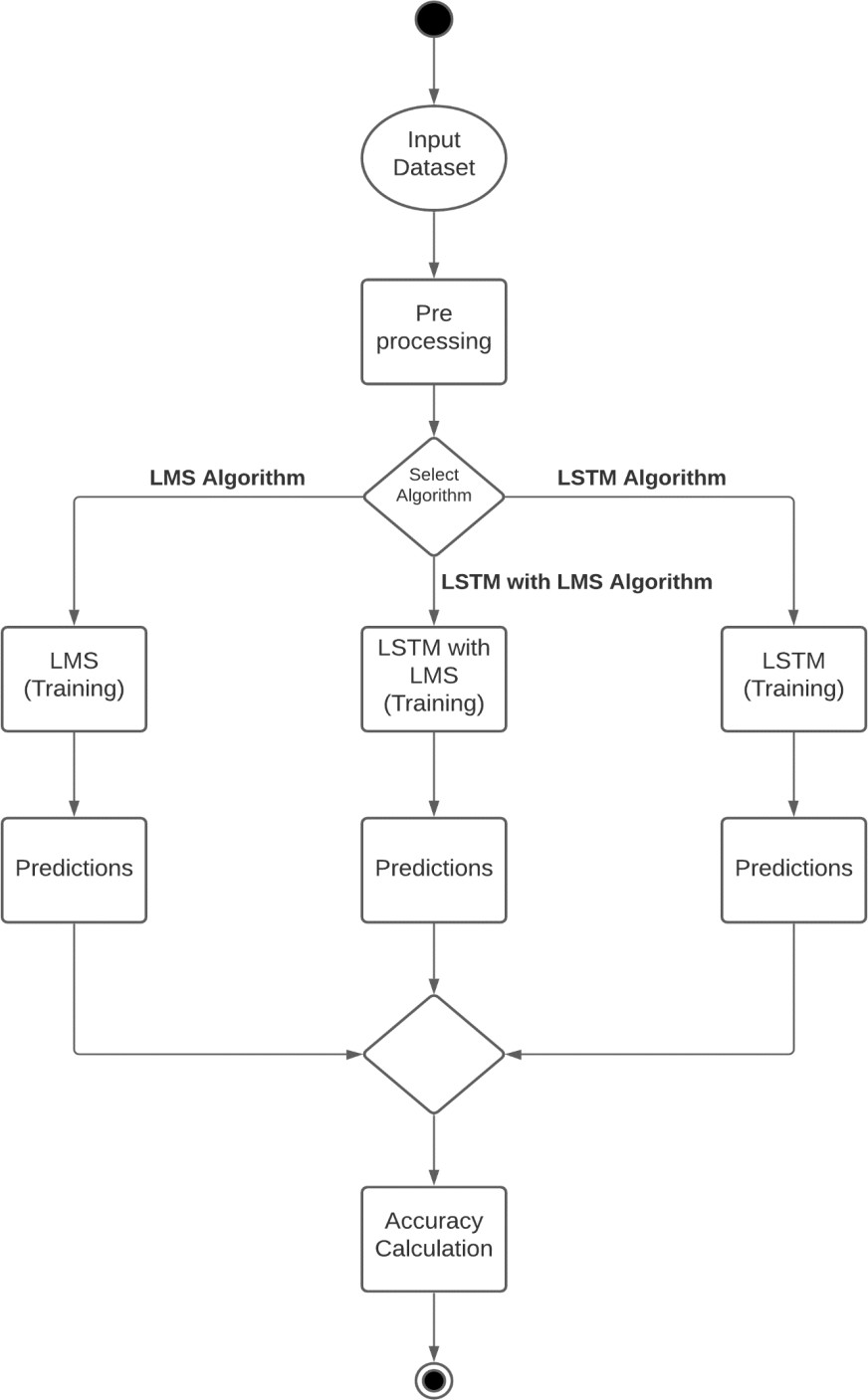
## ACTIVITY DIAGRAM

In the activity diagram as the initial activity starts user will upload the file as input which is action and in the next action time format will be selected.

* The decision box check chat format represents the validity of the time format of the file.
* If the time format is correct then analysis will be done and process will end.
* If the time format is wrong user will have to again check for the correct format.

### 

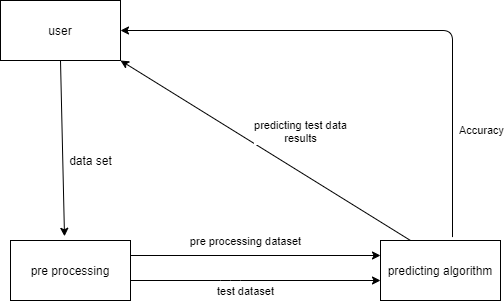
### ACTIVITY DIAGRAM



## COLLABORATION DIAGRAM

The collaboration diagram is used to show the relationship between the objects in a system. Both the sequence and the collaboration diagrams represent the same information but differently. Instead of showing the flow of messages, it depicts the architecture of the object residing in the system as it is based on object-oriented programming. An object consists of several features. Multiple objects present in the system are connected to each other. The collaboration diagram, which is also known as a communication diagram, is used to portray the object's architecture in the system.

* This collaboration diagram shows the relationship between the objects in a system.
* An object consists of several features. Multiple objects present in the system are connected to each other.



# SYSTEM IMPLEMENTATION

### PYTHON :

It is an interpreted, high-level general-purpose programming language. Created by Guido Van Rossum and first released in 1991. Its language constructs and objects-oriented approach aim to help programmer with clear, logical code for small and large-scale tools.

Python is used for web development (server-side), software development, mathematics, it can be used alongside software to create workflows, it can connect to database systems, it can also read and modify files, it can be used to handle big data and perform complex mathematics and can be used for rapid prototyping, or for production-ready software development.

Python is a high-level, general-purpose and a very popular programming language. Python programminglanguage (latest Python 3) is being used in web development, Machine Learning applications, along with allcutting-edge technology in the Software Industry. Python Programming Language is very well suited forBeginners.

* + 1. Python is currently the most widely used multi-purpose, high-level programming language.
    2. Python allows programming in Object-Oriented and Procedural paradigms.
    3. Python programs generally are smaller than other programming languages like Java. Programmers have totype relatively less and the indentation requirement of the language makes them readable all the time.
    4. Python language is being used by almost all tech-giant companies like – Google, Amazon, Facebook,Instagram, Dropbox, Uber… etc.

### JSON :

Java Script Object Notation is an open standard file format, and data interchange format, that uses human- readable text to store and transmit data objects consisting of attribute-value pairs and array data types. It is very common data format, with diverse range of applications. Such as serving as a replacement for xml in ajax systems.

Json is a language-independent data format. It was derived from JavaScript, but many modern programming languages include code to generate and parse JSON-format data. The official Internet media type for Json is application/json. Json filenames use the extension (.json). When exchanging data between a browser and a server, the data can only be text. Json is text, and we can convert any JavaScript object into json and json to the server. We can also convert any json received from the server into JavaScript objects. This way we work with the data as JavaScript objects, with no complicated parsing and transactions.

### DART :

It is a client-Optimized programming language for apps on multiple platforms. It is developed by google and is used top build mobile, desktop, server, and web applications. Dart is an object- oriented, class-based, garbage-collected language with C-style syntax. Dart can complete to either native code or JavaScript. It supports interfaces, mix-ins, abstract-classes, refined generics and type inference.

To run in mainstream web browsers, Dart relies on source-to-source compiler to JavaScript. According to the tool site. Dart was designed to be easy to write development tools for, well- suited to modern app development, and capable of high-performance implementations. When

running dart code in a web browser the code is precompiled into JavaScript using dart2.js compiler. Compiled as JavaScript, Dart code is compatible with all major browsers with no need for browsers to adapt dart.

Though optimizing the compiled JavaScript output to avoid expensive checks operations, code written in dart can, in some cases, run faster than equivalent code hand-written using JavaScript idioms.

### Monthly Chats Timeline

We will display the line chart to showcase the number of active chats per month for which year. For this, we count the messages by grouping them according to month and year columns. To plot the chart, we combine month and year columns.

### Daily Timeline

Similarly, we can create a daily timeline where you must group the data according to date and count the number of messages. To display this analysis Line chart is perfect.

# TESTING

Testing is the major quality control that can be used during software development. Its basic function is to detect the errors in the software. During requirement analysis and design, the output is a document that is usually textual and non-executable. After the coding phase, a computer program is available that can be executed for testing purposes.

## TESTING OBJECTIVES

* To check if the application is working as expected.
* To check the errors of different scenarios by using different test cases.

## TESTING METHODS & STRATEGIES USED ALONG WITH TEST DATA

### Software Testing Strategies-

Software testing is defined as an activity to check whether the actual resultsmatch the expected results and to ensure that the software system is Defect free.

It involves execution of asoftware component or system component to evaluate one or more properties of interest. Software testing alsohelps to identify errors, gaps or missing requirements in contrast to the actual requirements. It can be eitherdone manually or using automated tools.

In simple terms, Software Testing means Verification of Application under Test (AUT).

1. Functional Testing

## Functional Testing

Functional Testing is defined as a type of testing which verifies that each function of the software application operates in conformance with the requirement specification. This testing involves checking of User Interface ,APIs, Database, security, client/ server applications and functionality of the Application under Test. The testing can be done either manually or using automation.

**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

# SYSTEM INPUT AND OUTPUT SCREENSHOTS

## INPUT SCREENSHOTS

## APP.PY

* from flask import Flask, render\_template, request
* import utils
* app = Flask(\_\_name\_\_)
* @app.route('/')
* def main():
* return render\_template('index.html')
* @app.route('/predict/', methods = ['GET', 'POST'])
* def predict():
* if request.method == 'POST':
* Open = request.form.get('Open')
* High = request.form.get('High')
* Low = request.form.get('Low')
* Volume = request.form.get('Volume')
* prediction = utils.preprocess(Open, High, Low, Volume)
* return render\_template('prediction.html', prediction = prediction)
* # @app.errorhandler(404)
* # def error(e):
* #     return render\_template('404.html')
* if \_\_name\_\_ == "\_\_main\_\_":
* app.run(debug=True)

**TEMPLATES:**

**404 ERROR.HTML:**

<!DOCTYPE html>

<html lang="en">

<head>

    <meta charset="UTF-8">

    <meta http-equiv="X-UA-Compatible" content="IE=edge">

    <meta name="viewport" content="width=device-width, initial-scale=1.0">

    <title>404</title>

    <style>

        body{

            font-size: 2rem;

            margin: 0;

            padding: 0;

            display: flex;

            justify-content: center;

            align-items: center;

            background-color: black;

            color: antiquewhite;

            font-family: 'Gill Sans', 'Gill Sans MT', Calibri, 'Trebuchet MS', sans-serif;

        }

        div{

            width: 100%;

            height: 100%;

            display: flex;

            align-items: center;

            justify-content: center;

        }

    </style>

</head>

<body>

    <div>

        <h1>404 Error</h1>

    </div>

</body>

</html>

**INDEX.HTML:**

<!DOCTYPE html>

<html lang="en">

<head>

    <meta charset="UTF-8" />

    <meta http-equiv="X-UA-Compatible" content="IE=edge" />

    <meta name="viewport" content="width=device-width, initial-scale=1.0" />

    <title>Predictor</title>

    <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css" rel="stylesheet"

        integrity="sha384-1BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrCEXSU1oBoqyl2QvZ6jIW3" crossorigin="anonymous" />

    <style>

        body,

        html {

            margin: 0;

            padding: 0;

            height: 100%;

            /\* background: #60a3bc !important; \*/

            background: #0000FF         !important;

        }

        .user\_card {

            height: 400px;

            width: 350px;

            margin-top: auto;

            margin-bottom: auto;

            background: #f39c12;

            position: relative;

            display: flex;

            justify-content: center;

            flex-direction: column;

            padding: 10px;

            box-shadow: 0 4px 8px 0 rgba(0, 0, 0, 0.2), 0 6px 20px 0 rgba(0, 0, 0, 0.19);

            -webkit-box-shadow: 0 4px 8px 0 rgba(0, 0, 0, 0.2), 0 6px 20px 0 rgba(0, 0, 0, 0.19);

            -moz-box-shadow: 0 4px 8px 0 rgba(0, 0, 0, 0.2), 0 6px 20px 0 rgba(0, 0, 0, 0.19);

            border-radius: 5px;

        }

        .brand\_logo\_container {

            position: absolute;

            height: 170px;

            width: 170px;

            top: -75px;

            border-radius: 50%;

            background: #60a3bc;

            padding: 10px;

            text-align: center;

        }

        .brand\_logo {

            height: 150px;

            width: 150px;

            border-radius: 100%;

            border: 2px solid white;

        }

        .form\_container {

            margin-top: 100px;

        }

        .login\_btn {

            width: 100%;

            background: #c0392b !important;

            color: white !important;

        }

        .login\_btn:focus {

            box-shadow: none !important;

            outline: 0px !important;

        }

        .login\_container {

            padding: 0 2rem;

        }

        .input-group-text {

            background: #c0392b !important;

            color: white !important;

            border: 0 !important;

            border-radius: 0.25rem 0 0 0.25rem !important;

        }

        .input\_user,

        .input\_pass:focus {

            box-shadow: none !important;

            outline: 0px !important;

        }

    </style>

</head>

<body>

    <!-- <div class="container">

        <div class="d-flex align-items-center">

            <h1>Stock Price Prediction</h1>

        </div>

        <form action="/predict/" method="post">

            <div class="mb-3 from-group">

                <label for="Open" class="form-label">Open</label>

                <input type="text" class="form-control" id="Open" name="Open">

            </div>

            <div class="mb-3 from-group">

                <label for="High" class="form-label">High</label>

                <input type="text" class="form-control" id="High" name="High">

            </div>

            <div class="mb-3 from-group">

                <label for="Low" class="form-label">Low</label>

                <input type="text" class="form-control" id="Low" name="Low">

            </div>

            <div class="mb-3 from-group">

                <label for="Volume" class="form-label">Volume</label>

                <input type="text" class="form-control" id="Volume" name="Volume">

            </div>

            <button type="submit" class="btn btn-primary">Submit</button>

        </form>

    </div> -->

    <div class="container h-100">

        <div class="d-flex justify-content-center h-100">

            <div class="user\_card">

                <div class="d-flex justify-content-center">

                    <div class="brand\_logo\_container">

                        <img src="../static/aa.jpeg" class="brand\_logo" alt="Logo">

                    </div>

                </div>

                <div class="d-flex justify-content-center form\_container">

                    <form action="/predict/" method="post">

                        <div class="input-group mb-3">

                            <input type="text" name="Open" id="Open" class="form-control input\_user"  placeholder="Open Price">

                        </div>

                        <div class="input-group mb-2">

                            <input type="text" name="High" id="High" class="form-control input\_pass"  placeholder="High Price">

                        </div>

                        <div class="input-group mb-2">

                            <input type="text" name="Low" id="Low" class="form-control input\_pass" placeholder="Low Price">

                        </div>

                        <div class="input-group mb-2">

                            <input type="text" name="Volume" id="Volume" class="form-control input\_pass"  placeholder="Stock Volume">

                        </div>

                            <div class="d-flex justify-content-center mt-3 login\_container">

                    <button type="submit" class="btn login\_btn">Predict Close Price</button>

                   </div>

                    </form>

                </div>

            </div>

        </div>

    </div>

</body>

</html>

**PRERDICTION.HTML:**

<!DOCTYPE html>

<html lang="en">

<head>

    <meta charset="UTF-8">

    <meta http-equiv="X-UA-Compatible" content="IE=edge">

    <meta name="viewport" content="width=device-width, initial-scale=1.0">

    <title>Close Price Predict</title>

    <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css" rel="stylesheet"

        integrity="sha384-1BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrCEXSU1oBoqyl2QvZ6jIW3" crossorigin="anonymous" />

    <style>

        body{

            display: flex;

            align-items: center;

            justify-content: center;

            color: aliceblue;

            background: #001e29 !important;

        }

        .row, h1,h2{

            display: flex;

            align-items: center;

            justify-content: center;

            width: 600px;

            height: 130px;

            color: aliceblue;

        }

    </style>

</head>

<body>

    {% block content %}

    <div class="row justify-content-md-center mb-4">

        <div class="text-primary">

            <h1>Closing Price</h1>

            <h2> Prediction is {{ prediction }}</h2>

        </div>

    </div>

    {% endblock %}

</body>

</html>

**UTILITY.PY:**

import joblib

import numpy as np

def preprocess(Open, High, Low, Volume):

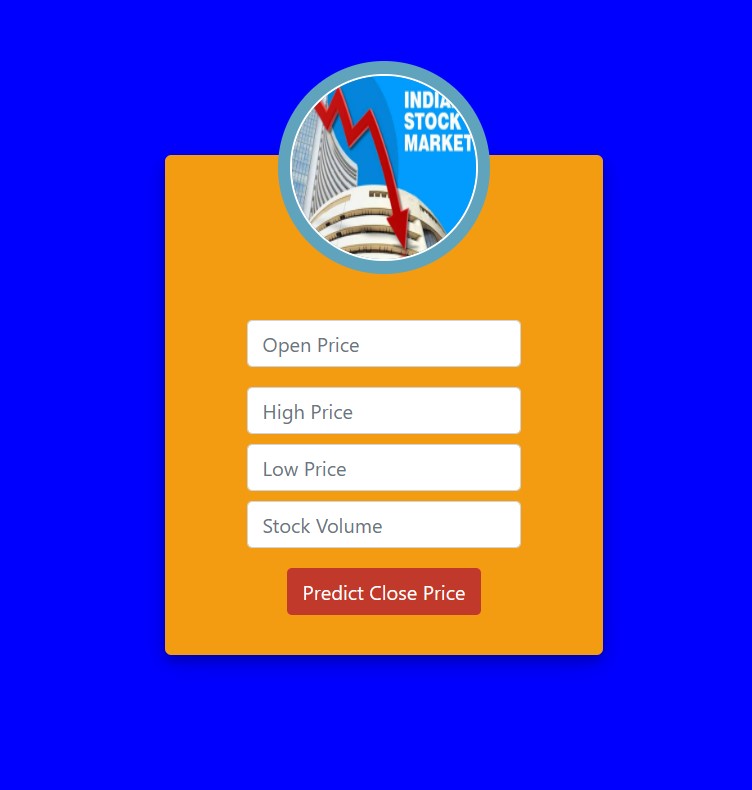
    test\_data = [[float(Open), float(High), float(Low), float(Volume)]]

    trained\_model = joblib.load('model.pkl')

    prediction = trained\_model.predict(test\_data)

    return prediction

## OUTPUT SCREENSHOTS



# 

# LIMITATIONS AND SCOPE OF PROJECT

## LIMITATIONS OF PROJECT

While machine learning can be a useful tool for predicting stock market prices, there are several limitations to using it as the sole method for making investment decisions. Here are some of the main limitations:

1. Limited historical data: Machine learning algorithms require a large amount of historical data to make accurate predictions. However, stock market data can be limited, especially for companies that have only been publicly traded for a short period of time.

2. Non-stationary data: Stock market data is often non-stationary, meaning that the statistical properties of the data change over time. This can make it difficult to build accurate machine learning models that can adapt to changing market conditions.

3. Complexity of market dynamics: The stock market is influenced by a wide range of factors, including economic indicators, geopolitical events, and investor sentiment. Machine learning models may struggle to capture the complexity of these market dynamics, which can lead to inaccurate predictions.

4. Unforeseeable events: The stock market can be unpredictable, and unexpected events such as natural disasters, political upheavals, or corporate scandals can significantly impact stock prices. Machine learning models may not be able to account for such unforeseeable events, which can make their predictions unreliable.

5. Lack of interpretability: Machine learning models can be difficult to interpret, especially for complex models such as neural networks. This can make it challenging for investors to understand why a model is making a particular prediction, which can make it difficult to trust and act on the model's recommendations.

Given these limitations, it is important to use machine learning as one of several tools for predicting stock market prices, rather than relying on it exclusively. It is also important to incorporate fundamental analysis and other forms of market analysis into the decision-making process.

# SCOPE OF PROJECT

The scope of using machine learning for predicting stock market prices is quite broad, and it has the potential to be a valuable tool for investors and traders. Here are some of the key areas where machine learning can be applied:

1. Predictive modeling: Machine learning algorithms can be used to build predictive models that analyze historical stock market data to identify patterns and make predictions about future stock prices. These models can be used for a variety of applications, such as portfolio optimization, risk management, and trading strategies.

2. Sentiment analysis: Machine learning algorithms can be used to analyze social media sentiment and news articles to gauge investor sentiment and identify potential market trends. This information can be used to inform trading decisions and develop investment strategies.

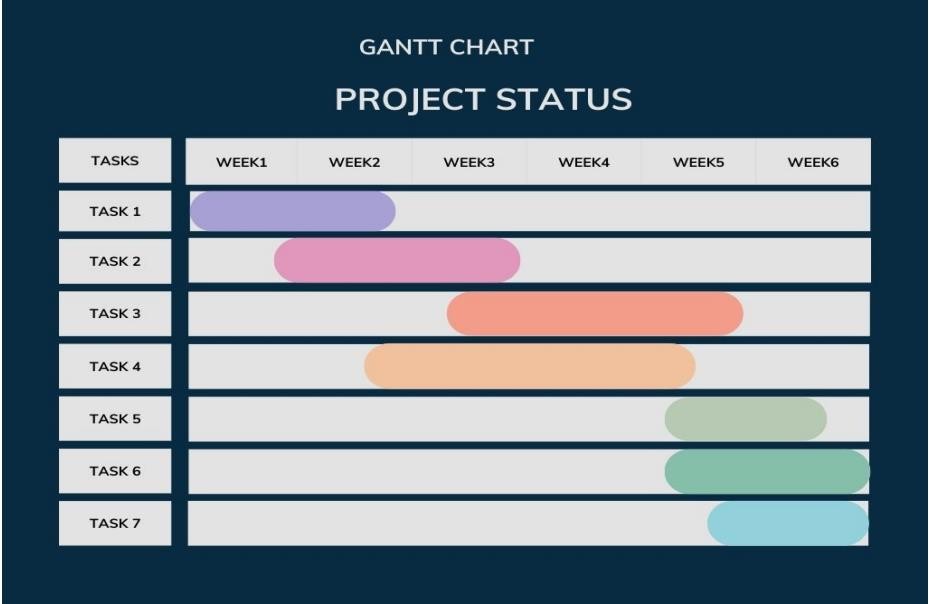
3. Fraud detection: Machine learning algorithms can be used to detect fraudulent trading activity in the stock market. This can help to prevent insider trading and other illegal activities, which can have a significant impact on stock prices.

4. Portfolio management: Machine learning algorithms can be used to optimize investment portfolios by identifying the most promising stocks and minimizing risk. This can help investors to achieve better returns and reduce losses.

5. Trading algorithms: Machine learning algorithms can be used to develop automated trading algorithms that can execute trades based on predetermined rules and market conditions. These algorithms can help to eliminate human bias and emotion from trading decisions, resulting in more consistent and profitable trading strategies.

Overall, the scope of using machine learning for predicting stock market prices is quite broad, and it has the potential to significantly improve the accuracy and efficiency of investment strategies. However, it is important to remember that machine learning should be used as one tool among many, and that human judgment and expertise remain critical for successful investing.

# GANTT CHART



**IMPACT OF PROPOSED SYSTEM IN**

# ACADEMICS AND INDUSTRY

1. Conducting a stock price prediction project can help students develop a range of skills, including data analysis, statistical modelling, programming, and problem-solving. These skills are highly valuable in a variety of fields, particularly in finance and related industries.
2. Undertaking a stock price prediction project provides students with an opportunity to apply theoretical concepts learned in class to real-world scenarios. This can help students to gain a better understanding of how financial markets work and the challenges associated with predicting stock prices.
3. Stock price prediction projects can involve collaboration between students and faculty members, as well as with industry professionals. This can help students to develop professional networks and gain exposure to different perspectives and approaches.
4. Stock price prediction can help investors make informed investment decisions. By analysing trends and patterns in stock prices, investors can decide whether to buy, hold or sell shares of a particular company. Accurate predictions can result in better investment outcomes and higher returns on investment.
5. Predicting stock prices can help companies manage their risk exposure by identifying potential risks and opportunities. For example, if a company predicts a decline in its stock price, it may decide to sell its shares before the price drops further, thereby reducing its exposure to risk.
6. Companies that are able to accurately predict stock prices may have a competitive advantage over their rivals. By making better investment decisions and managing risk effectively, these companies can outperform their competitors in the stock market.

# 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.

It has the potential to improve the accuracy and efficiency of investment strategies, and it can be applied in a wide range of areas, including predictive modeling, sentiment analysis, fraud detection, portfolio management, and trading algorithms.

However, it is important to recognize the limitations of machine learning for predicting stock market prices. The stock market is complex and influenced by a wide range of factors, and machine learning models may struggle to account for all of these factors. It is important to use machine learning as one tool among many, and to incorporate fundamental analysis and other forms of market analysis into the decision-making process.

Additionally, it is important to approach machine learning with a critical eye and to recognize the potential for bias and inaccuracies in the data and models. Machine learning models should be constantly evaluated and updated to ensure that they remain accurate and effective.

Overall, machine learning has the potential to significantly improve the accuracy and efficiency of investment strategies, but it should be used as part of a comprehensive and well-informed investment approach.

# REFERENCES

* P. Sailapathi Sekar, “Financial stock market forecastusing data mining Techniques".
* Tiffany Hui-Kung, “A Neural network-based fuzzy
  + Time.
* Rasheed Khaled. “Extracting the best features for predicting stock prices using machine learning.”
* "Automated news reading: Stock price prediction based on financial news using
  + context- specific features." InSystem Science.
* GOOGLE

THANK YOU