Indira Gandhi Delhi Technical University for Women

(Established by Govt. of Delhi vide Act 09 of 2012)

Kashmere Gate, Delhi - 110006

B.Tech. CSAI

SOFTWARE ENGINEERING

Subject Code: BIT 203

STOCK PREDICTION SYSTEM

PROJECT FILE

Submitted to: *Submitted by:*

Dr. Nisha Rathee   **03001172020 - Jhanvee Khola**

05201172020 - Khushi Punia

ACKNOWLEDGEMENT

I would like to express my sincere thanks to Ms. Nisha Rathee, for her valuable guidance and support in completing my project.

I would also like to express my gratitude towards her for giving us this great opportunity to do a project on Stock Prediction using Machine Learning. Without her support and suggestions, this project would not have been completed.

I am assuring that this project was done by me and not copied from anywhere.

DECLARATION

We, **Jhanvee Khola(03001172020) and Khushi Punia(05201172020)**, solemnly declare that the project report, **STOCK PREDICTION PROJECT,** is based on my work carried out during our study under the supervision of Asst. Prof. Nisha Rathee, Department of Information Technology. I assert the statements made and conclusions are drawn are an outcome of my research work. I further certify that:

I. The work contained in the report is original and has been done by me under the supervision of my supervisor.

II. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this university or the any other University of India or abroad.

III. We have followed the guidelines provided by the university in writing the report.

IV. Whenever we have used materials (text, data, theoretical analysis/equations, codes/program, figures, tables, pictures, text, etc.) from other sources, we have given due credit to them in the report and have also given their details in the references.

Jhanvee Khola(03001172020)

Khushi Punia(05201172020)

INDEX

LIST OF TABLES

PROBLEM STATEMENT

Stock is an unpredictable curve that has been in the picture ever since. Its essence had been ever long living and indulging. It has grown its popularity with respect to time. People are more fascinating and interested in the same things than before. Same for the case for the organisation. Organisation had created it as a better source of revenue generation rather than investing and taking a loan approval from the bank. It's way more efficient and less hectic from the firm point of view.

Stock is unpredictable and its been the same from the start. Its way of escalating and de-escalating has been a phenomenon and experiencing the same is the best integral part of it. It has its upper hand and flexibility with the changes that has the chances of uprising as well as crashing the whole market. Its easily defined in few words but making an essence and understanding the same is way more hectic and time consuming.

Simpler its sound complex is its phenomenon and integrating the same. It has its whole different sets of dependencies and integration from different agents which fluctuate the same in the market. Finding an accurate and getting the exact values out of the same is still unaligned and no particular model of the same is seen in the market value.

Finding the closest and getting an accurate proximate value out of such an unpredictability is a problem in itself. Merging of the data getting the best prediction to increase the efficiency alongside considering the different aspects of the moderator is tough and we took the same in consideration and implemented with every aspect to generate the best out of the same and get a result that can be better interrupted and the efficiency remains the same with the value of different aspects of creating an impact of reducing the risk and influencing the same over the time period to gain the most out of it.

This is totally based on Machine Learning Algorithm to proceed and provide an effective result. Getting the data and processing it and generating a forecast for three days is the problem statement that we worked on.

SOFTWARE REQUIREMENT SPECIFICATION

Table of Contents

INTRODUCTION

PURPOSE

This project helps in bridging the resources and empowering the people to know and trade the most out of stock and understand the generation and the vulnerabilities that have to be seen and predicted. The enhancement of the same is done with the resource graph which makes an user or the customer to analyse the same and take the needs and important details before dealing and consider those things for the yield that the person is willing to invest on. The predictability itself is a challenge and that's the main purpose of the report.A correct prediction of stocks can lead to huge profits for the seller and the broker. Frequently, it is brought out that prediction is chaotic rather than random, which means it can be predicted by carefully analysing the history of the respective stock market. Machine learning is an efficient way to represent such processes. It predicts a market value close to the tangible value, thereby increasing the accuracy.

SCOPE

This project requires investigation in the following areas:
Stock Market -
●   Investigating trends in stock market and factors affecting the stock prices.
Data mining techniques -
●   Investigating the available tools and techniques for data mining and then selecting those that are best fit to solve the problem.

DEFINITIONS, ABBREVIATIONS

LSTM: Long short term memory
PCA: Principal component analysis
RNN: Recurrent neural networks
*ANN:*Artificial neural network
DNN: Deep neural network
DTW: Dynamic Time Warping
RFE: Recursive feature elimination
SVM: Support vector machine
CNN: Convolutional neural network
SGD: Stochastic gradient descent
ReLU: Rectified linear unit
MLP: Multi layer perceptron

OBJECTIVE

In Stock Market Prediction, the aim is to predict the future value of the financial stocks of a company. The recent trend in stock market prediction technologies is the use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and authentic. The product focuses on the use of Regression and LSTM based Machine learning to predict stock values. Factors considered are open, close, low, high and volume.

OVERVIEW

The rest of the SRS is organised as follows:

- Section 2 is an overall description of the project.
- Section 3 cites the specific requirements

OVERALL DESCRIPTION

PRODUCT PERSPECTIVE

The prediction of a stock market direction may serve as an early recommendation system for short-term investors and as an early financial distress warning system for long-term shareholders. Forecasting accuracy is the most important factor in selecting any forecasting methods. Research efforts in improving the accuracy of forecasting models are increasing since the last decade. The appropriate stock selections those are suitable for investment is a very difficult task. The key factor for each investor is to earn maximum profits on their investments. In this paper Support Vector Machine Algorithm (SVM) is used. SVM is a very specific type of learning algorithms characterised by the capacity control of the decision function, the use of the kernel functions and the scarcity of the solution. With this product,we will investigate the predictability of financial movement with SVM. To evaluate the forecasting ability of SVM, we compare its performance with Decision trees. These methods are applied on 2 years of data retrieved from Yahoo Finance. The results will be used to analyse the stock prices and their prediction in depth in future research efforts.
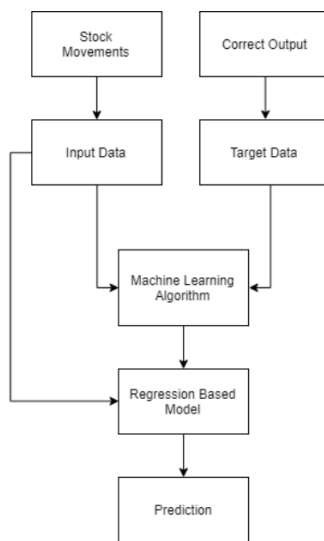
2.1 Regression Based Model
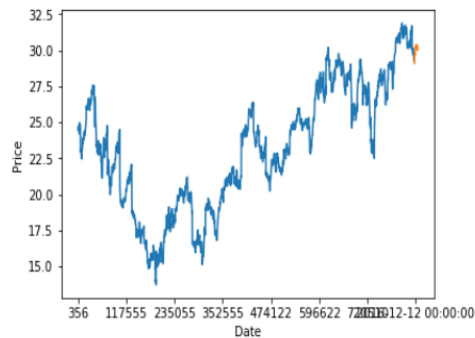

Figure 1. Flow Chart for Regression Based Model

Figure 3. Plot between Price and Date Using Regression

PRODUCT FUNCTIONS

The following list of function descriptions explains the major features of the Stock Predictor-

1. User Registration
   The registration function shall allow users to create secure accounts. The account will track the user's name, email address and password. This offers convenience so the user does not have to enter a lot of information to get started.
   For security purposes, the password of the user is encrypted so that no one can have access to the user's account, not even the administrator.

2. User Login
   The account login function shall allow account members to enter their email and password. Once verified, users will be able to access his/her account, view previous predictions and select new company stock to predict.

3. Predict Closing Stock Price of Company
   Users can predict the closing price of the top company he/she selected by choosing the option of the specific company from the options list.
   He/she can also compare and get advise on that particular stock or any other company's stock.

4. Add/Update/Delete
   Update/Delete users from database rights are only accessible to the admin.

5. Read Blogs

The users can use Stockamie to read up on top rated blogs related to stock market, financial investments and upcoming trends. This feature will also be available to guest users(who are not registered/logged in).

6. About Us
   Know more about the product and the mechanism behind it. Learn about the vision and mission and future goals of the product.

7. Account Logout
   The account logout function shall allow account members to exit their accounts for security purposes.

USER CHARACTERISTICS

Users of the website should have basic knowledge about stock markets and financial investments. They must know how to navigate a website and fill online forms.

CONSTRAINTS

One constraint of the product is that Financial/Stock markets are very unpredictable, filled with ups and lows being affected due to various factors. Therefore, disclaimer needs to be provided to not use the prediction just from word of mouth. Another constraint on the services and functions of the system is that the software product must conform to the user disclaimer policy.

ASSUMPTIONS AND DEPENDENCIES

Stockamie website is only available on the internet, so it is automatically assumed on the user end that they have good internet connectivity on their end. Availability of web browser, hardware requirements and website navigation knowledge is also expected from the user.

APPORTIONING OF REQUIREMENTS

There are no requirements that may be delayed until future versions of the system. Users will be notified of any major updates to the product.

SPECIFIC REQUIREMENTS

EXTERNAL INTERFACE REQUIREMENTS

SYSTEM INTERFACE

1) About us: It will be showcasing the purpose of the website and major idea and vision about the product.
2) Registration: Users have to register with a name and email id in order to use the system.
3) Predicting Page : After registration the user will enter the page where he will be asked for which company he wants to see the prediction.Then closing prices will be predicted for that company on that day.
4) Blogs: There will be some related blogs to give suggestions to the user and aware him about the latest on going trends in the market.
5) Trend Analysis : It will present graphs of closing prices in the previous years of different companies.This will give some insight about how the company's closing prices have risen or dropped in the previous years and how it will be in upcoming years.

USER INTERFACE

The system will provide the users to predict closing prices of stocks and have insight about the ongoing trend of stocks of various companies online. There will be two different
user interfaces that will accompany this website: one for the
user, and the administrators.

●   Users will be able to predict the stock prices only after login or registration. They will be able to access the blogs and the trend analysis graphs without login.

●   Administrators will be required to log in at all times. However, they will have access via the web interface to be able to add or delete users from the database or add more companies in the prediction option list.

HARDWARE INTERFACE

There are no special hardware interfaces requirements.

SOFTWARE INTERFACE

Software required to make the working of the product is:-
●   Front end - HTML
●   Back end - Flask
●   Database - SQLite

- Machine Learning - Python

There are no special communication interfaces requirements.

FUNCTIONAL REQUIREMENTS

1. **Stimulus: Click "Register" Button: User Registration**
   a. The system shall allow a non-registered user to create a secure account.
   b. The system shall require the following information from the user: Name, email id , Password.
   c. The system shall confirm the email and password are acceptable.
   d. The system shall store the information in the database.

2. **Stimulus: Click "Login" Button: User Login**
   a. The system shall allow a registered user to log in to their account.
   b. The system shall require an email id and password from the user.
   c. The system will verify the email id and password, and the user will be logged-in.

3. **Stimulus: Click "Predict Now" Button: To see list of companies available**
   a. The system shall allow users to see available choices of companies on clicking the predict now button.

4. **Stimulus: Click "Select Company" Button: Choose Company to predict stock price**
   a. The system shall allow the user to select the company for which the user wants to predict the prices.

5. **Stimulus: Click "About Us" Button: Know more about the product**
   a. The system shall allow the user to see the motive and vision behind the product.
   b. The user shall be able to see various platforms where they can contact the administrator.

6. **Stimulus: Click "Blogs" Button: Top rated blogs to improve financial literacy**
   a. The user shall be able to see a variety of blogs to improve their knowledge on stocks and finance.

7. **Stimulus: Click "Logout" Button: Account Logout**
   a. The system shall allow the registered and logged-in user to exit his/her account so that access to operations requiring a user to be logged in is now disabled.

PERFORMANCE REQUIREMENTS

The performance requirements/aims are as follows-

1. Final prediction of stock and rendering of trend analysis time does not exceed 5 seconds
2. User Registration/Login should be completed in maximum 10 seconds

LOGICAL STRUCTURE OF DATA

TYPES OF INFORMATION USED

FUNCTION	TYPES OF INFORMATION USED
Account Registration	User information (name, email address,and password)
Account Login	User information (email id, password)
Predict price	Stock dataset of companies in previous years(Opening prices,high,low,closing prices,volume)
About us	Administrator information(email id, linkedin)

DATA MODEL

USER	
Name	String
E-mail	Unique, Valid, String
Password	String, Min 8 characters long

COMPANY STOCK	
Name	String
Opening Price	Float
Closing Price	Float

Low	Float
High	Float
Volume	Float

DESIGN CONSTRAINTS

The Stock Predictor shall conform to the following design constraints:
● Able to support Windows/MAC/Android.
● The system logs out when the user clicks the logout button.
● Website is compatible with all types of web browsers.

STANDARDS COMPLIANCE

The Stock Predictor will follow existing standards and regulations, which are stated in the user policy/ Terms & Conditions.

SOFTWARE SYSTEM ATTRIBUTES

AVAILABILITY

The Stock Predictor shall be available to users 24 hours a day, 7 days a week, except being down for maintenance no more than one hour a week. If the system crashes, it should be back up within one hour.

SECURITY

Users will be able to access only their personal information and not that of other users. Users will not be able to see the search history of other users to maintain privacy of other users.
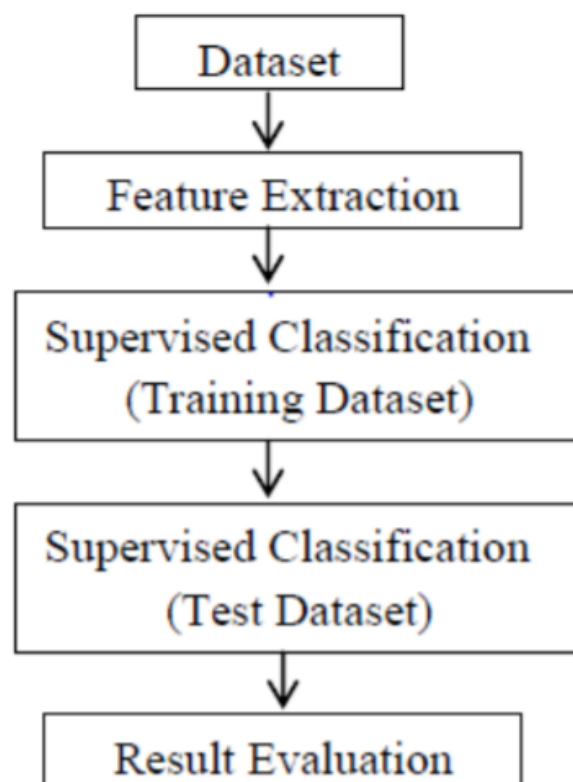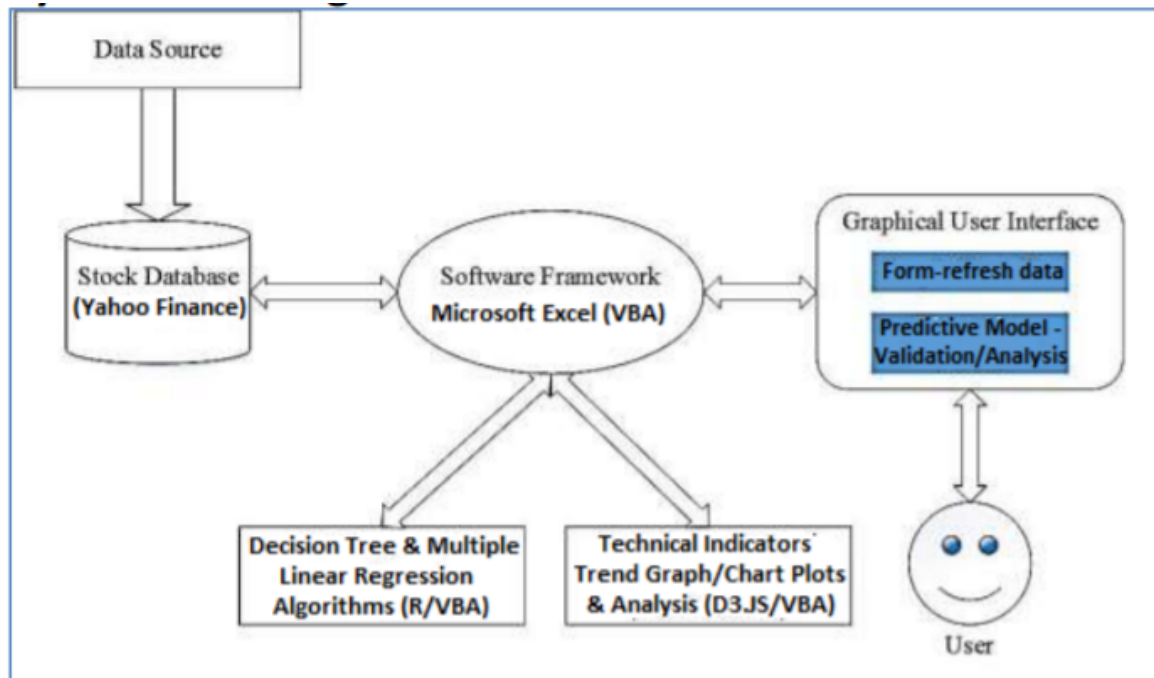
MAINTAINABILITY

Any updates or defect fixes shall be able to be made on server-side computers only without any patches required by the user.
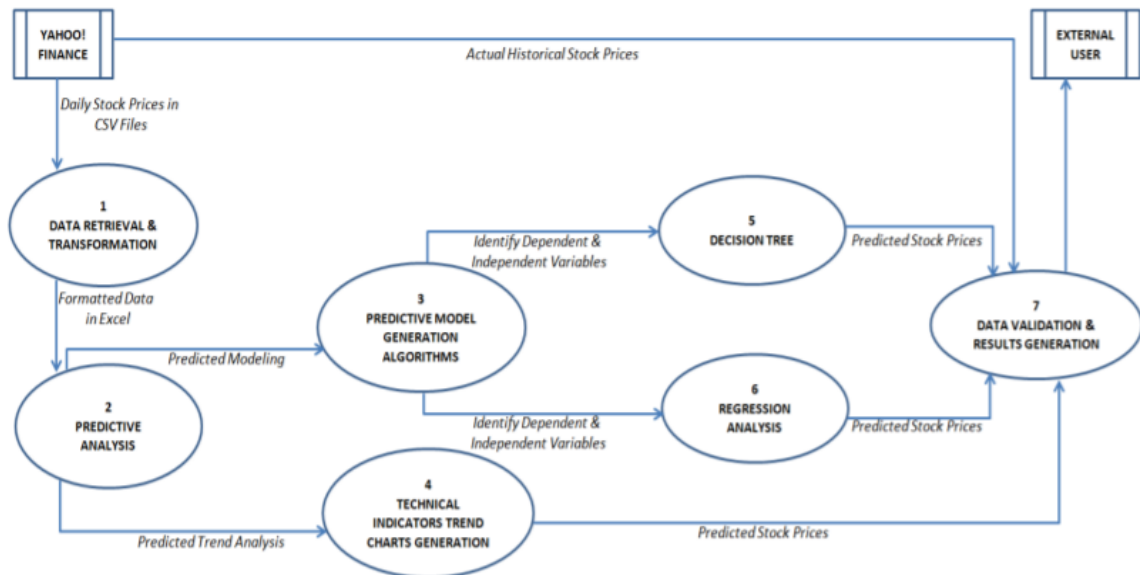
PORTABILITY

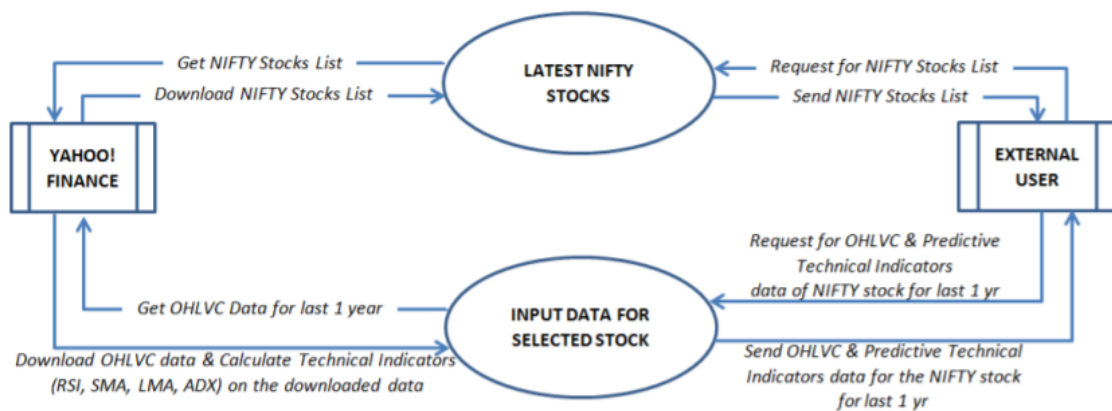Nothing required

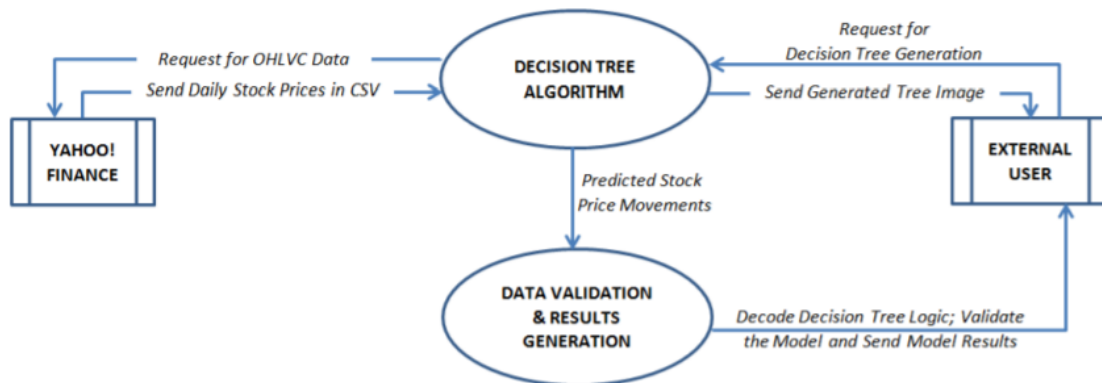# DATA FLOW DIAGRAM

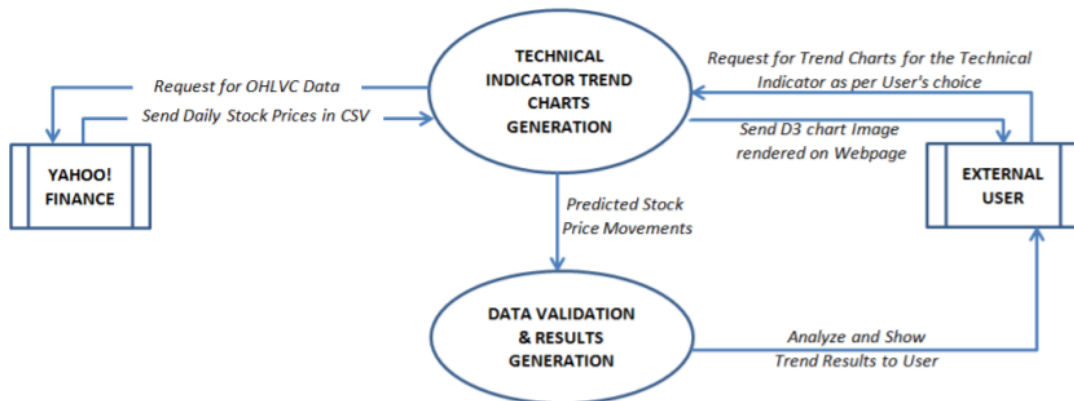## SYSTEM FLOW DIAGRAM

LEVEL 0 DFD

LEVEL 1 DFD

Level – 1 DFD (Data Retrieval & Transformation):
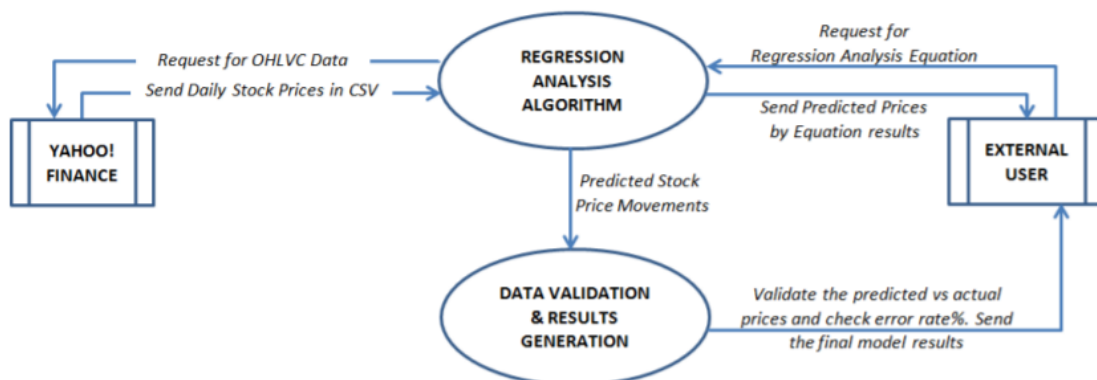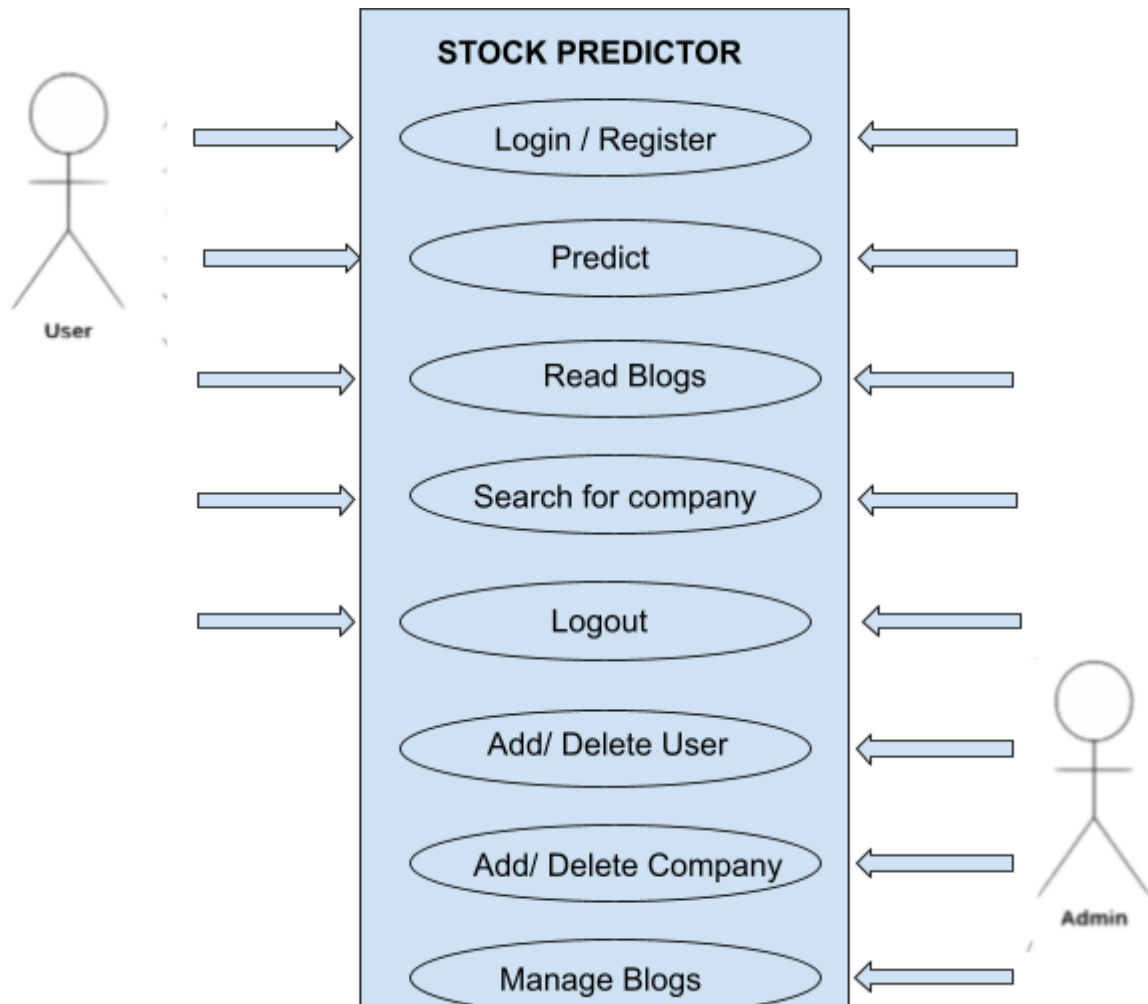
Level – 1 DFD (Decision Tree):

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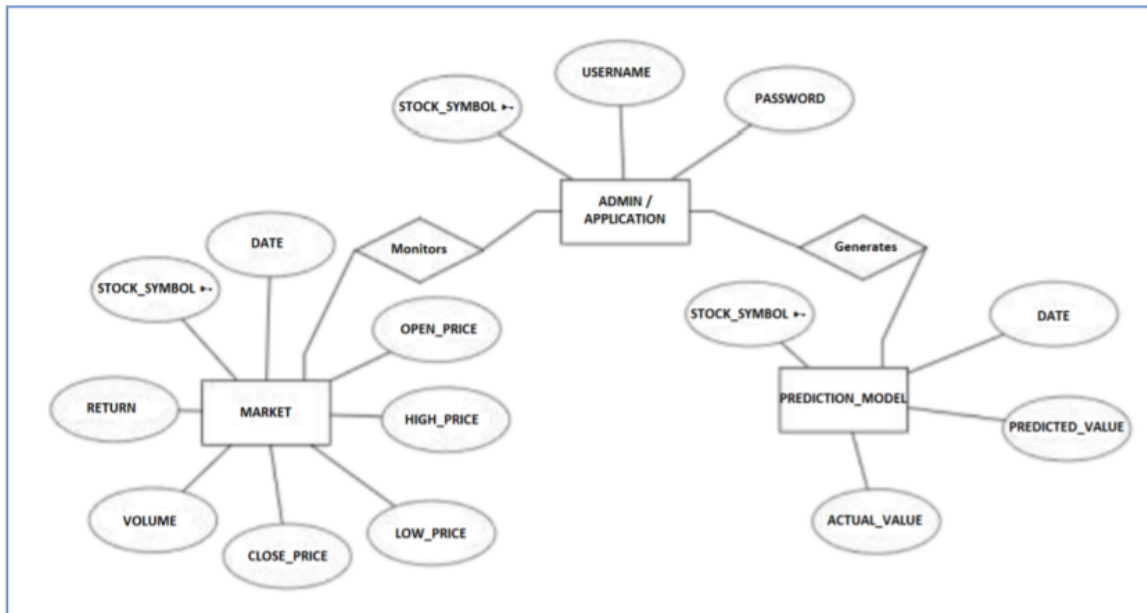

Level – 1 DFD (Regression Analysis):

USE CASE DIAGRAM

STOCK PREDICTOR

User

- Login / Register
- Predict
- Read Blogs
- Search for company
- Logout
- Add/ Delete User
- Add/ Delete Company
- Manage Blogs

Admin

E-R DIAGRAM

TOOLCHAIN

Technologies that have been used in this project have been selected from a large array of technologies by keeping in mind the following characteristics:
- Data validation
- Performance
- Reliability
- Scalability
- Security
- Portability
- Performance
- Time constraint

The toolchain used in developing this project is as follows:
- Client-Side Scripting: HTML, CSS, Javascript
- Server Side Scripting: Flask, Jinja
- Database tool: SQLite
- Machine Learning: Python

Literature about above-given technologies:-
- HTML (Hypertext Markup Language) - It is the language that tells your web browser what each component of a website is. So, using HTML, you can define headers, paragraphs, links, images, and other elements so that your browser understands how to organise the web page you're viewing.

- CSS (Cascading Style Sheets) - Cascading Style Sheets (CSS) is a style sheet language that is used to describe the display of a document produced in a markup language like HTML.

- Flask - Flask is a Python-based micro web framework. It is used to build the web application's backend. In the flask, we build models and routes.

- Javascript - JavaScript is a computer language that is largely utilised by Web browsers to provide users with a dynamic and interactive experience.

- Jinja - Jinja is a Python programming language web template engine.

- SQLite - It is used to keep the web application's database up to date. A database is a structured collection of information or data that is often stored electronically in a computer system.

- Python - Python is an interpreted high-level general-purpose programming language.

WORKING

STOCKAMIE

Manage your finances, get latest investment tips and news, make intelligent predictions of those stocks and grow your portfolio tenfolds...all with Stockamie

Learn more

The finance and investment world maybe seem a bit intimidating but with right guidance you'll be investing with ease in no time.

Learn more

Stockamie uses various Machine Learning Algorithms to help you make the right choices when chosing your next stock to invest in.

FUTURE SCOPE

- The historical dataset available on the company's website consists of only a few features like high, low, open, close, adjacent close value of stock prices, volume of shares traded etc., which are not sufficient enough. To obtain higher accuracy in the predicted price value new variables have been created using the existing variables.
- For future work, deep learning models could be developed which consider financial news articles along with financial parameters such as a closing price, traded volume, profit and loss statements etc., for possibly better results.
- In future we plan to provide the user to make stock investments directly from the website and maintain an investor's portfolio.
- By developing more advanced machine learning models we aim to give better trend analysis over the years and between the companies to make it easier for the user to choose the area of his/her investment.

CONCLUSION

Two techniques have been utilised in this paper: LSTM and Regression, on the Yahoo finance dataset. Both the techniques have shown an improvement in the accuracy of predictions, thereby yielding positive results. Use of recently introduced machine learning techniques in the prediction of stocks have yielded promising results and thereby marked the use of them in profitable exchange schemes. It has led to the conclusion that it is possible to predict stock market with more accuracy and efficiency using machine learning techniques. In the future, the stock market prediction system can be further improved by utilising a much bigger dataset than the one being utilized currently. This would help to increase the accuracy of our prediction models. Furthermore, other models of Machine Learning could also be studied to check for the accuracy rate resulted by them.

REFERENCES AND BIBLIOGRAPHY

[1] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock Market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.

[2] K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017 pp. 1-1.

[3] H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.

[4] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.

[5] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1-5.

[6] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.
[7] S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modeling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications ICIEA), Wuhan, 2018, pp. 2787-2790.

[8] T. Gao, Y. Chai and Y. Liu, "Applying long short term momory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.