STOCKS PRICE PREDICTION

Final Year Project

Session 2020-2024

A project submitted in partial fulfilment of the degree of

BS in Computer Science

Department of Computer Science

Faculty of Computer Science & Information Technology

The Superior University, Lahore

Spring 2024

Type (Nature of project)	[✓] **D**evelopment [ ] **R**esearch [ ] **R**&**D**
Area of specialization	
FYP ID	FYP-BCSM-F23-064

Project Group Members

Sr.#	Reg. #	Student Name	Email ID	*Signature
(i)	BCSM-F20-365	Aafaq Ahmed	Bcsm-f20-365@superior.edu.pk	
(ii)	BCSM-F20-077	Mubashir Murtaza	Bcsm-f20-077@superior.edu.pk	
(iii)	BCSM-F20-305	Ali Haider	Bcsm-f20-305@superior.edu.pk	

*The candidates confirm that the work submitted is their own and appropriate credit has been given where reference has been made to work of others

Plagiarism Free Certificate

This is to certify that, I Aafaq Ahmed Son of Muhammad Ashfaq, group leader of FYP under registration no BCSM-F20-365 at Computer Science Department, The Superior University, Lahore. I declare that my FYP report is checked by my supervisor.

Date: Name of Group Leader: Aafaq Ahmed Signature: _____

Name of Supervisor: Ms. Hafiza Maria Iqbal

Designation: Lecturer

Signature: _____

HoD: Dr. Irfan ud Din

Signature: _____

Stocks Price Prediction

Change Record

Author(s)	Version	Date	Notes	Supervisor's Signature
Aafaq Ahmed	1.0	7-9-2023	Project Overview and read research papers	
Ali Haider	1.1	20-9-2023	Literature Review	
Aafaq Ahmed, Mubashir Murtaza	1.2	10-10-2023	Gantt Chart, Empathy Map	
Mubashir Murtaza	1.3	27-10-2023	Functional and non-functional requirements	
Ali Haider	1.4	8-11-2023	Use Case model and Description	
Aafaq Ahmed	2.0	20-11-2023	System design and architecture	
Aafaq Ahmed, Ali Haider, Mubashir Murtaza	2.1	2-12-2023	Implementation of the software started	
Aafaq Ahmed	2.2	18-2-2024	Testing and Evaluation of the application	
Aafaq Ahmed, Ali Haider, Mubashir Murtaza	3.0	13-03-2024	Concluding the project and summarizing info	

APPROVAL

PROJECT SUPERVISOR

Comments: _____

_____

Name: _____

Date: _____ Signature: _____

PROJECT MANAGER

Comments: _____

_____

Date: _____ Signature: _____

HEAD OF THE DEPARTMENT

Comments: _____

_____

Date: _____ Signature: _____

Dedication

This work is dedicated to my respected Parents who have been there all the times in my life whether it was good times or bad times. They also gave me motivation and provided necessary guidance to tackle any kind of problem in my life. I would also like to dedicate my work to my respected teachers who not only provided important guidance during the work but also gave us counselling that me and members might require in the near future. In the end, this is dedicated to all the people who helped us even a little bit during our entire degree journey.

Acknowledgements

We are really thankful to our supervisor who has guided us thoroughly and throughout our final year project (FYP I and II) and helped us even at the smallest obstacle that we came across various times during this journey of ours.

We would also like to take this moment to express and show our gratefulness and regard to Superior University Gold Campus Lahore, Pakistan for providing us a platform to pursue our studies and develop our final year project (FYP).

We would also like to show our thanks to all the faculty of Computer Science who actively or passively encouraged us throughout the entire process. Last but not the least, we would like to thank our respected parents and good friends for all of their moral support they gave us during this wonderful journey of ours.

Executive Summary

Predicting the value of the stocks in the stock markets have always interested not only researchers but the people also. People who wanted to buy new stocks always want that their money do not go to waste and they make a good profit out of it. Same is the case with people who already owns a single or multiple stock in the market to sell it a very reasonable price and they could also get their hands on the profit by selling the shares. Many methods have been developed and implemented regarding that. As the human made advancements in the field of technology, new and improved methods have come out and are being developed to a perfect point where it can provide some great benefits to the mankind.

The growing trend of Artificial Intelligence and Machine Learning has provided us some magnificent algorithms and other methods through which we can not only predict the future values with great accuracy in any field e.g., Stock market, but we can do a lot of other things as well. This project will mainly focus to target the stock market field in which we are going to use Machine Learning or its relevant algorithms to predict the upcoming future values of stocks and with that we will combine it with application whether web or mobile with great user interfaces.

The method will include steps like in short, the algorithm or method will be provided some dataset of different companies or we can call it as historical data. Then we will use this dataset for training a model which we will then verify it by testing the model. The accuracy will depend how well a dataset has been provided for the training of the model. It will also depend which algorithm has been picked for the whole process. LSTMs have very powerful features when it comes to sequence prediction problems because they are able to store useful information as well past things which is going to prove a valuable feature for our problem. This is important in our case because the previous price of a stock is crucial in predicting its future price.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Chapter 1: Introduction

This chapter mainly focus on the very in detail and in-depth introduction of our project which is called as “Stocks Price Prediction Using Machine Learning “. This chapter will start with brief introduction of the project. Then we will look at the motivations and challenges we faced while doing this project. Then we will look at some of the previous work done by other people on the same topic or any other work related to this field. Gantt chart and work breakdown structure has also been discussed in this chapter as well.

1.1. Background

Stock market is a collection and mixture of different public companies and exchanges where normal activities such as buying and selling shares of publicly traded company occurs. It is a market for trading shares of publicly traded companies which are being listed. The primary market is where companies sell their stock to the general public in an initial public offering (IPO). This is mainly done to raise and increase funds. People usually buy shares or stocks in the hope that their value will rise in the upcoming future. However, there is always uncertainty in the stock market, which makes people hesitant to invest their money in it. As a result, we require a technique that can forecast stock market prices so that people can invest their money. Time-series forecasting has become a widely used method in many real-world related problems and applications. It predicts the next time unit using continuous or sequential data collected over a period of time. Many time-series predictions algorithms have been demonstrated to be effective in practice. The most widely used algorithm which will help us tackle the problem of forecasting of sequential data is, Long-short Term Memory (LSTM) which is an improved and advanced variant of Recurrent Neural Network (RNN).

1.2. Motivations and Challenges

Every person or specially investors wish they somehow knew the future and if not the future, they wished they might get some details or insights about the particular stock or share they are interested in. So, this thing motivated us to develop an application where a user can interact with it and the system will give them an output.

1.3. Goals and Objectives

The goal of this project is to develop a system or we can call it an application where people or users can input a ticker symbol of a publicly traded company in the stock market, and in return they can get a forecast value of the next upcoming days. This primarily done so that people or investors can get a slight insight of how the stock or share will behave in the future.

1.4. Literature Review/Existing Solutions

Literature Survey:

Stock closing price prediction based on sentiment analysis and LSTM:

In this research paper, the researcher has used deep learning approach to predict the stock market using LSTM and sentiment analysis. First, they involved the investors or user's sentiment in their research because it can greatly improve the accuracy of a model. Then they adapted the LSTM because of its great advantages like storing and remembering past information in its cell or block which make it most suitable for time series data. Similarly, they made further changes in it by introducing attention-based mechanism so that it can pay attention more towards the more critical information. Thus, in the end, the results showed that by adopting these changes can combining them not only improves model accuracy but it can also.

Existing Solutions:

Number of platforms are available which provide forecasting on the stock market and the stocks/shares. But two of them which stand out are Altindex and candlestick ai. The second one only provides its services through mobile application and the first one has number of defects like it takes into account how many numbers of followers have increased or decreased, how many posts have made related to that particular company ticker symbol like amazon, tesla etc. It also does not take into account the sentiments of people over the various social media platforms like twitter, on how are they feeling towards that company performance.

1.5. Gap Analysis

At the moment our app can predict or forecast stock prices with just historical prices or dataset, sentiment analysis of tweets and view real time stock data but in the near future we want the application to take into account every related field or aspects that affect the stock market in one way or other.

1.6. Proposed Solution

First of all, we are going to make a web-based application and in that we will use machine learning algorithm LSTM, and any other method as deemed appropriate, and will perform sentiment analysis of tweets which can help us to have better accuracy rather than using just a simple LSTM model (because researches have showed that combining sentiment analysis with LSTM algorithm gives better RMSE value). Also, we are providing this all free of cost so that users don't have to pay anything and the users can view real time stock data.

1.7. Project Plan

We are creating a plan that will helps us in completing our project and it includes many different steps that will tell us about the accuracy of the project. And we will be able to know the time of the project completion and how it will be completed.

1.7.1. Work Breakdown Structure

Project Management

- Work Breakdown Structure (WBS)
- Roles & Responsibility Matrix
- Develop Course Plan
- Develop Schedule

Reports/Documentation

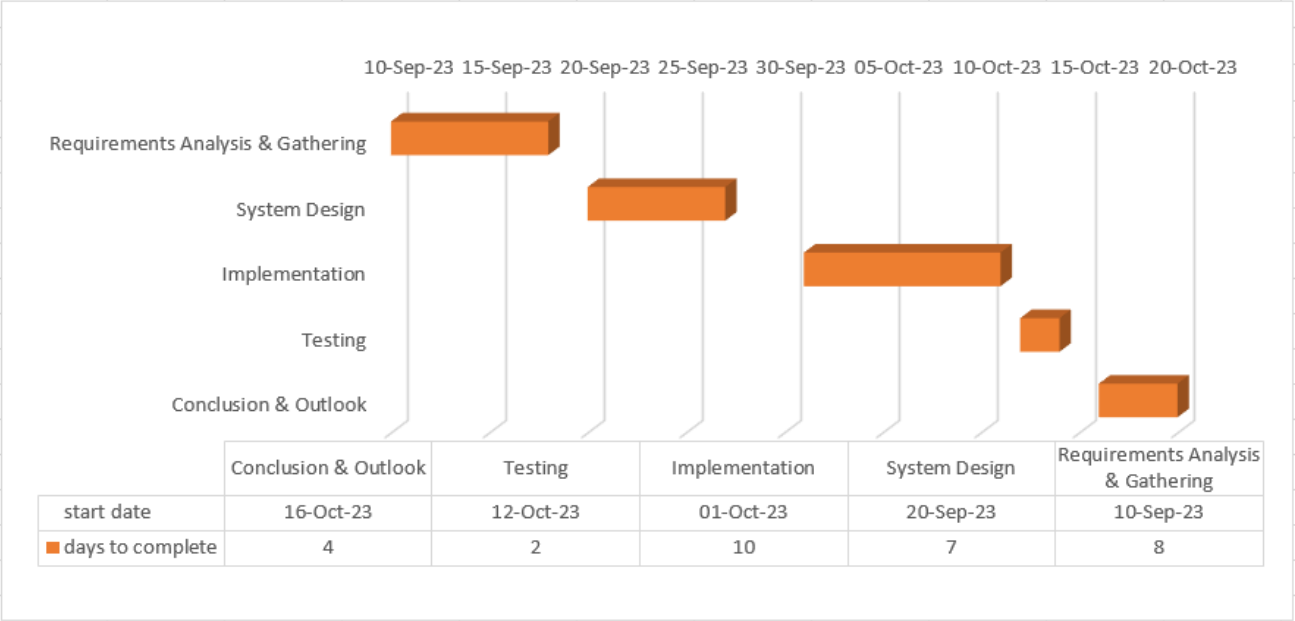
- SRS
- Requirements Analysis
- System Design
- Implementation
- Testing & Performance Evaluation

1.7.2. Roles & Responsibility Matrix

Table 1.1. Roles and Responsibility Matrix

WBS #	WBS Deliverable	Activity #	Activity to Complete the Deliverable	Duration (# of Days)	Responsible Team Member(s) & Role(s)
1	Requirements Analysis & Gathering	2	Meeting	8 Days	Aafaq & Ali
2	System Design	2	Meeting	1 Week	All group members
3	Implementation	2	Meeting	10 days	All group members
4	Testing and Performance Evaluation	2	Meeting	1-2 days	Ali Haider & Mubashir
5	Conclusion & Outlook	2	Meeting	3-4 Days	Aafaq Ahmed

1.7.3. Gantt Chart



	Conclusion & Outlook	Testing	Implementation	System Design	Requirements Analysis & Gathering
start date	16-Oct-23	12-Oct-23	01-Oct-23	20-Sep-23	10-Sep-23
days to complete	4	2	10	7	8

Figure. 1.1. Gantt Chart

1.8. Report Outline

- Introduction (Objectives, project overview etc.)
- Details of Technology used
- Features and Functionalities of app (User authentication, real-time stock prices etc.)
- Implementation and their details (Integration of yfinance API, setting up the environment etc.)
- Challenges and their solutions
- Future Work and further enhancements
- References

1.9. Empathy Map

Say:

- Users want a user-friendly web or mobile application through which they can navigate easily & perform their desired task.

Do:

- We will develop an application that will have functions as per gathered requirements and will make sure that it meets the standards.

Think:

- Users might think that developing such an application will require big budget but our team will try to keep it in the range of people.

Feel:

- By using the app users will find It friendly and easy to use and they will feel more secure because all their expectations have been met.

Chapter 2

Software Requirement Specifications

Chapter 2: Software Requirement Specifications

2.1. Introduction

This chapter mainly focus on the system requirements which must be specified while making this application. This helps the user as well as the reader to understand what type of specifications they are going to get if they use this product. Also, it provides a deep detail on the product features and how they interact with one another.

2.1.1. Purpose

The purpose of this SRS to define every minor detail possible of the product we are going to make in this project. Like it will include what type of features it should have and what don't. What types of requirements it has to fulfill in order to run properly and smoothly? What type of OS and what things a user needs to have in order to run this application in their own device etc.

2.1.2. Document Conventions

Our project document will follow the standard conventions so that it gets easier for the user or reader to read properly and it will help them to understand it properly. For example, for Formatting we will be using the below methods:

- Fonts: The documentation of our project uses Calibri (Body) font for the main text and Calibri light (Heading) for headings.
- Special Importance: We have also used bold or colored text for section headings and their content will follow simple formatting as defined above.

2.1.3. Intended Audience and Reading Suggestions

This document is intended for developers, project managers, users, testers, and documentation writers. Sequence for reading this document is to begin with the executive summary and chapter number 1 for introduction of the project and proceed to the sections such as chapter 2 for understanding on which tools and software this project was build.

2.1.4. Product Scope

Defining the product scope is very important because it helps designers, developers and other users in understanding what type of functionalities it will include and what features it will not include or what constraints are there. Our application includes user friendly interface through which users can navigate to the services the application will offer, user authentication so that only genuine users can use the features of the website. We will be building our application's front-end as well back-end so that the user has full developed applications waiting for them to get used.

2.1.5. References

- Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.
- Moritz, B., & Zimmermann, T. (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. Available at SSRN 2740751.

2.2. Overall Description

2.2.1. Product Perspective

This product being specified in this SRS is a follow-on member of a product family or we can say a new, self-contained product. Many researches have been conducted on this type of topic and yet it still requires a lot of more improvements in this specific research domain.

2.2.2. Product Functions

User Accounts: Our application users need to create an account or log in to use the app and our users can also update their account information.

Real-Time Stock Info and Details: application users can look up specific stocks by their ticker symbol. Show the latest stock prices and other details like trading volume and market value.

Compare Stock Performance: Users of app can pick multiple stocks to compare how they've been doing over time. Visual tools to show the performance of different stocks side by side.

Price Predictions: we are using the LSTM model to forecast stock prices for the next 30 days and show the predicted prices in easy-to-read charts and tables.

2.2.3. User Classes and Characteristics

Various user classes will use this product of our project such as user class working in finance, user class of undergraduates that might use or seek this product as a reference to develop their own project related to this field. The pertinent characteristics of each user class includes that each of them wants to test and try for themselves and rather for anyone else which means they have the same objective which match with each other.

2.2.4. Operating Environment

Environment in which the product of our project will run includes web because it is a web-based application, so the users must navigate to their browsers to open this app in their browsers to use it efficiently. They should also have an internet connection. Users can use their devices such as phones, laptops and PC etc.

2.2.5. Design and Implementation Constraints

Design and implementation constraints are those which can affect how a system works overall. If we want our system or application to run without any effect, we must know the limitations or constraints our system might or might not have. Some of the constraints are discussed below:

Implementation Constraints: Some constraints of our project are explained below;

Ensuring Availability and Quality of Data: First of all, in our project the accuracy and reliability of stock price predictions are very much highly dependent on the quality and availability of historical data of stock market because if a user wants to view the prediction about a particular company, there is chance that the dataset fetched by yahoo finance API might contain improper data or uncleaned data which might prove not good for our project.

Having Computational Resources: Our project's application computational resources are also one of the constraints for our application because the resources required for training and running the LSTM model may or may not be equally important, as it might potentially will require high-performance computing resources, especially for large datasets.

Design Constraints: Some of the constraints of our application are listed below;

Scalability of an app: If our application start getting traffic, it might become a constraint and it will become challenging to accommodate the increasing users on our application.

Project's app user Interface: Designing an interactive and user-friendly interface for our project app that provides relevant information to the users visiting or using our application features without any complexity can prove to us be very challenging in the near future or in present.

2.2.6. User Documentation

- **Login/Signup page**: This is the first page user will see. They have to login or signup first to access the main page.
- **Main Page**: This is the main page of our application. The users can view real-time stock data information, compare performances of more than one stock data, and in the end, users can predict future prices of a particular stock.

2.2.7. Assumptions and Dependencies

Two important terms known as assumptions and dependencies are used while doing any project management and planning. Assumptions are just statements that we consider true but they are always guaranteed to be proven and on the other hand dependencies are just relationships which exist between different tasks. Some of them are listed below to get more insights about them:

Using Data Availability in app: In our project, one of the assumptions includes that historical stock price data is consistently available for the selected stocks, allowing for reliable model training and prediction.

Ensuring good model Training: Similarly, an assumption of project's application is there which is the LSTM model training process is effective in capturing relevant patterns in the historical data, which will lead our application to accurate predictions.

Having Real-time Data Feeds: Real-time market data feeds of our project, are accessible to provide users with up-to-date information and will help enhance the accuracy of predictions also falls in assumptions.

User Engagement: Users will actively engage with our project's application, providing meaningful feedback that can help the system to work more efficiently in one way or other.

Use of good web Infrastructure: The web infrastructure of our application or product, including hosting services and server resources, can adequately handle the expected user load & data processing.

Data preprocessing and API for obtaining datasets: For our project, there are some of the dependencies one of which includes is that the training of our LSTM model depends on the data preprocessing to make the data suitable to give it as an input to the model training. Similarly, to do preprocessing is depends on the availability of the API like yahoo finance to get desired dataset to make predictions.

2.3. External Interface Requirements

2.3.1. User Interfaces

Dashboard of our project app: The main interface our project application, where users can view the main page of the website and other relevant features. In this feature, our application page will contain information and navigation bar the application from where users can navigate through different pages of the website.

Using Historical or Prediction Data Visualization: This interface of our project application, presents historical stock price data and relevant information through simple and easy to interact with charts or graphs. In this, project app will use historical dataset of companies or stock prices which will then be displayed during the forecasting or predictions of any company as input by the use of our application. This will increase and enhance the user experience.

2.3.2. Hardware Interfaces

Hardware:

- Installed Memory (RAM): 8.00 GB
- System type: 32-bit or 64-bit Operating System and all other devices that can access these features through their smartphones.

2.3.3. Software Interfaces

- **Back-end Data Service Provider**: Yahoo! Finance or another provider as deemed appropriate.
- **Operating System:** Windows 7 and so on.
- **Front-end Tool:** stream lit

2.3.4. Communications Interfaces

Our Web User Interface (UI): This interface of our project app involves user inputs such as entering a ticker symbol and displays predictions and relevant information about that particular stock or ticker symbol.

API for our project app: This interface is very much important in our project because it will allow or enables the data exchange and prediction requests between the web application and external systems in the project. For instance, as the user of our application, inputs the ticker symbol, the application will pick it up and send it to the already present trained LSTM model where all the important steps will be performed and then the output generated by the model will be sent back to our main application where it will be displayed to the user.

Using Database Interface: The main and important duty of our project application database interface is to retrieve and store data for model training, as well as accessing profiles of our application users such as their login and signup details.

Project's app ML Model Interface: An important communication interface of our project application because our most of the work will be done through this and it Involves sending historical data for training, sending prediction results to the application and displaying it to the user on the application screen.

2.4. System Features

Some of the system feature include friendly user interface. Login and registration facility. Main or Home page through which our application user can navigate to other services or features. Prediction chart will be shown after a user has input a specific ticker symbol, which will then show the original and predicted price calculated by the LSTM model of our project.

2.4.1. System Feature 1

Login and Registration on the web application.

2.4.1.1. Description and Priority

This is an important feature allowing only the authentic users to use the service and its priority is 7.

2.4.1.2. Stimulus/Response Sequences

Once the user login signup for the first time, the system will take response by directing them to home page from where they can user other features.

2.4.2. System Feature 2

LSTM model as our main feature.

2.4.2.1. Description and Priority

The main feature of our project is a page where a user input a ticker symbol of a company and then our model shows the predicted price and the original price as a comparison. Its priority is 9.

2.4.2.2. Stimulus/Response Sequences

User inputs a ticker symbol in the field bar, the system or we can say that yahoo finance fetches the specific dataset from the website and performs its operations including all important data pre-processing steps on it and show the results in a form of a graph.

2.4.2.3. Functional Requirements

REQ-SF2-1: Login and Registration

REQ-SF2-2: Input a Specific ticker symbol of a company

REQ-SF2-3: ML model should be ready beforehand

2.5. Non-functional Requirements

Non-functional requirements are the requirements which are not directly concerned with the specific function delivered by the system. They specify the criteria that can be used to judge the operation of a system. Non-functional requirements of our project application are specified as:

- Data should be taken from authorized sources
- Data retrieved through API should be accurate
- Predictions made should be useful and worthwhile

2.5.1. Performance Requirements

Some of the performance requirements of our project application are discussed below: -

Good Response Time: An important requirement of our application because the web application or system of our project should provide stock price predictions within a maximum response time. It is crucial because making or designing our system in a way that responds to user requests quickly, will help increase the user experience of our app.

Having Throughput: Another important requirement for our project includes that the system should be able to handle simultaneous requests for stock price predictions in large/bulk amounts. The ability to handle concurrent users are crucial for a web application.

Keeping in mind the Scalability: Our project should also have the ability to scale and accommodate in case of an increasing number of users without a decrease in response time of our application or prediction accuracy of the pre trained LSTM model. Once number of users of our project application increase in size, it should be able to handle the demand without sacrificing or any alteration in performance.

2.5.2. Safety Requirements

Safety requirements are important as they are designed and developed to ensure the working environment of any application is secure and operational just like we mentioned below:

Ensuring Security of data: We will be going to Implement security mechanisms to secure sensitive user data, such as login credentials because it is very important to protect user data and is essential to prevent unauthorized access of someone else data and ensure privacy.

2.5.3. Security Requirements

Using Encryption for data: Similarly, for our project application it is necessary to Implement end-to-end encryption for all communication between our app's users and server, especially for sensitive data such as user credentials and other information because encryption can help us and the system to protect sensitive information from unauthorized access during its use.

Using Authentication and Authorization: In our project application, it is very important to implement and develop an authentication system, using secure password storage mechanisms etc. Enforce proper authorization mechanisms to control user access levels. We are doing this because the authentication and authorization is important for ensuring that only authorized users can access and modify their data.

2.5.4. Software Quality Attributes

Software quality attributes means that there are characteristics of a software that determine how well an application can performs and meets the needs the users of the applications over a period of time. These includes such as security which is important and means how well our software application will protect against unauthorized access and other type of cyber-attacks on our web applications. Also includes how well our application is easily useable and how friendly the interface is.

2.5.5. Business Rules

While a writing a report for a software business rules are the conditions and guidelines that helps in understanding how the application will interact with the users of a project software. Such as user first have to authenticate themselves before using the application. Similarly, the data of the stocks should be collected through the yahoo finance and it should be updated on regular basis and interface should user friendly and should show all the information as required.

2.5.6. Usability Requirements

Once again for our system and project application usability requirements are important as they define the easiness with which a user can interact with our application and get what they want form our website. For example, our application will have a user-friendly interface for smoother experience and we will provide relevant and necessary information on how to use the application so that a user doesn't get caught anywhere and can use it very easily.

2.5.7. Reliability Requirements

Just like any other requirements, reliability requirements are important for our project application as they will help us to understand what does the system need to run its desired features smoothly without encountering any kind of failure. Our system or app should be operational and accessible at all time. For instance, our project application will be made to go through error handling which help the application to prevent any crashes and should manage any other error gracefully.

2.5.8. Maintainability/Supportability Requirements

For maintaining our project application, number of things would be considered to ensure the applications does not encounter any type of error. A reliable database will be used for maintaining the relevant information of our application and on the side the application server will take care of the website. In case of any type of failure, the system will try to run and launch himself again to ensure our user does not experience any difficulty. Also, the software design is being done with modularity in mind so that maintainability can be done easily.

2.5.9. Portability Requirements

The application should work across various operating systems, browsers and different kinds of smart phones, laptops, tablets etc.

2.5.10. Efficiency Requirements

The system should deliver stock predictions within an acceptable time frame.

2.6. Domain Requirements

Domain requirements like any other requirement are important and are those which are specific to an application domain in which the system operates. Some of the requirements are discussed below:

Having Historical Data Access: Our project application must have access to reliable and up-to-date historical stock price and market data because accurate predictions will depend on the availability of up to date and reliable historical data for training our application LSTM model.

Using Real-Time Data Feeds: Our application will also integrate real-time data feeds to provide users with the most current stock prices and market information so that we can provide users of our project application with best experience while they are on our website. If we provide details on time and update the information, it will become important for making investment decisions.

Chapter 3

Use Case Analysis

Chapter 3: Use Case Analysis

In this chapter we will look at the use-case model and use cases description. So, a use case model is a model of how different types of users interact with the system to solve a problem. As such, it describes the goals of the users, the interactions between the users and the system, and the required behaviour of the system in satisfying these goals. Then we will look at the use cases description which A use case description is a detailed document that provides more information about a specific use case identified in the use case model.

3.1. Use Case Model

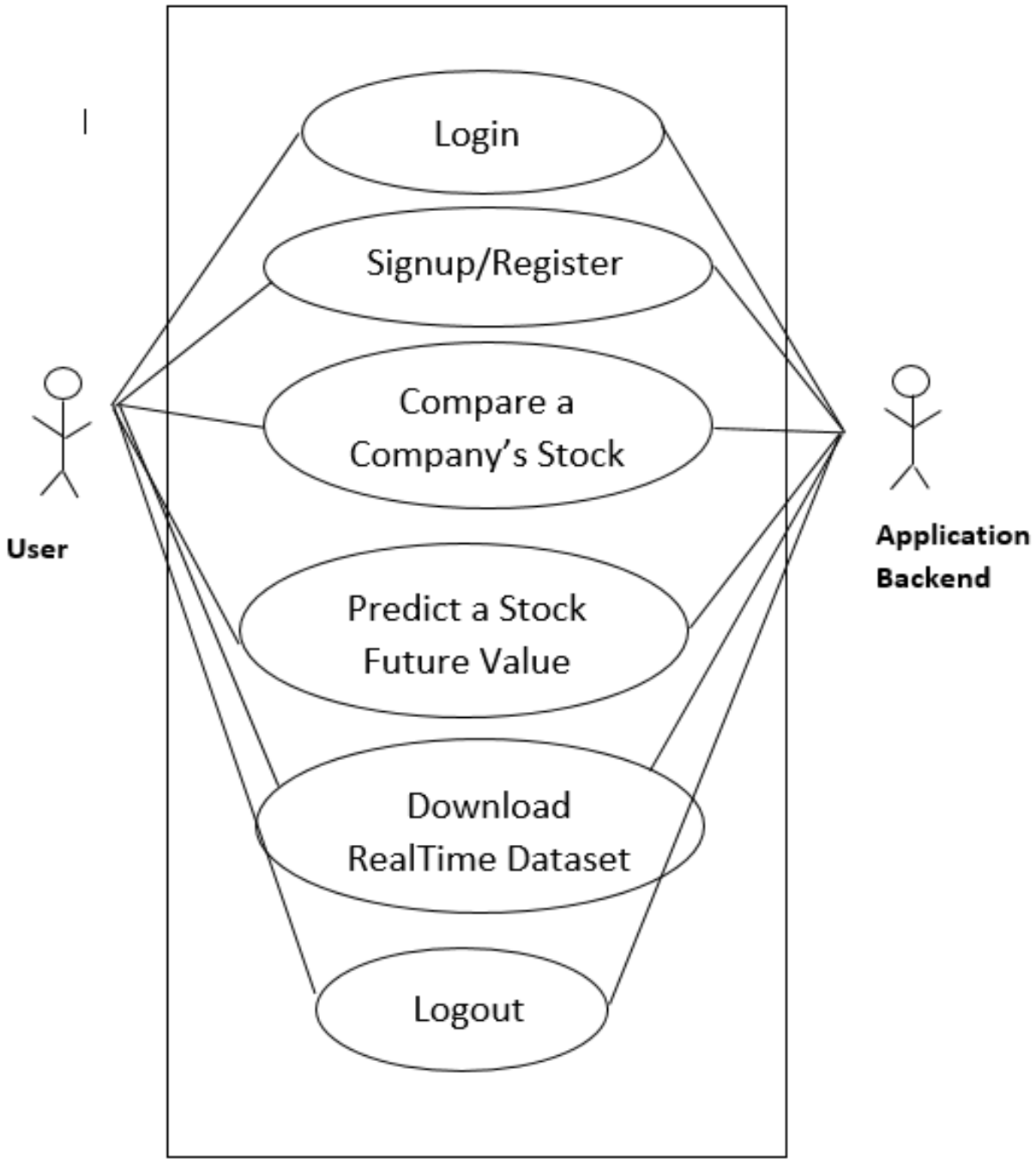


Figure. 3.1. Use Case Model

3.2. Use Cases Description

Use case ID: 1 Use case name: Login: Users can Login through their credentials which they provided during their signup process.

Use case ID: 2 Use case name: Signup/Register: Users can register or signup through user friendly interface and once they do this, they can login and see the main page where they can browse through different pages of the website. Users can also easily reset their password if they accidently forgot it. They can simply enter their email, temporary password will be sent and then using temporary password they can reset their it by creating a new password for their account.

Use case ID: 4 Use case name: Compare a company stock: Users can easily compare and visualize stock data of two companies by easily entering their stock symbol.

Use case ID: 5 Use case name: Compute result and performance Description: Prediction result will be handled and generated by Automated User Interface Application Backend. The system will be built, through which the result of prediction and system performance will be analysed.

Use case ID: 6 Use case name: Predict a stock value: With this user can easily see the predictions of their desired company's stocks. The result will be displayed in the form of a graph.

Use case ID: 7 Use case name: Download Realtime Dataset: Users can easily download a dataset of any company (historical) in a csv format.

Use Case ID: 8 Use Case Name: logout: Users can logout of their account whenever they want to.

Chapter 4

System Design

Chapter 4: System Design

In this chapter we will various kinds of designs of our product which will in turn help us to understand about the project working and how the process actually works.

4.1. Architecture Diagram

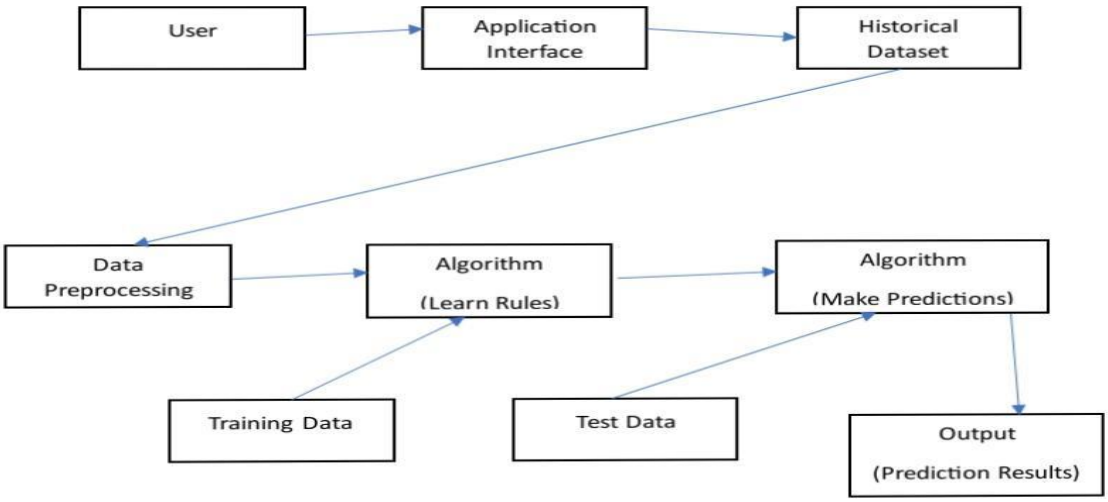


Figure. 4.1. Architecture Diagram

4.2. Domain Model

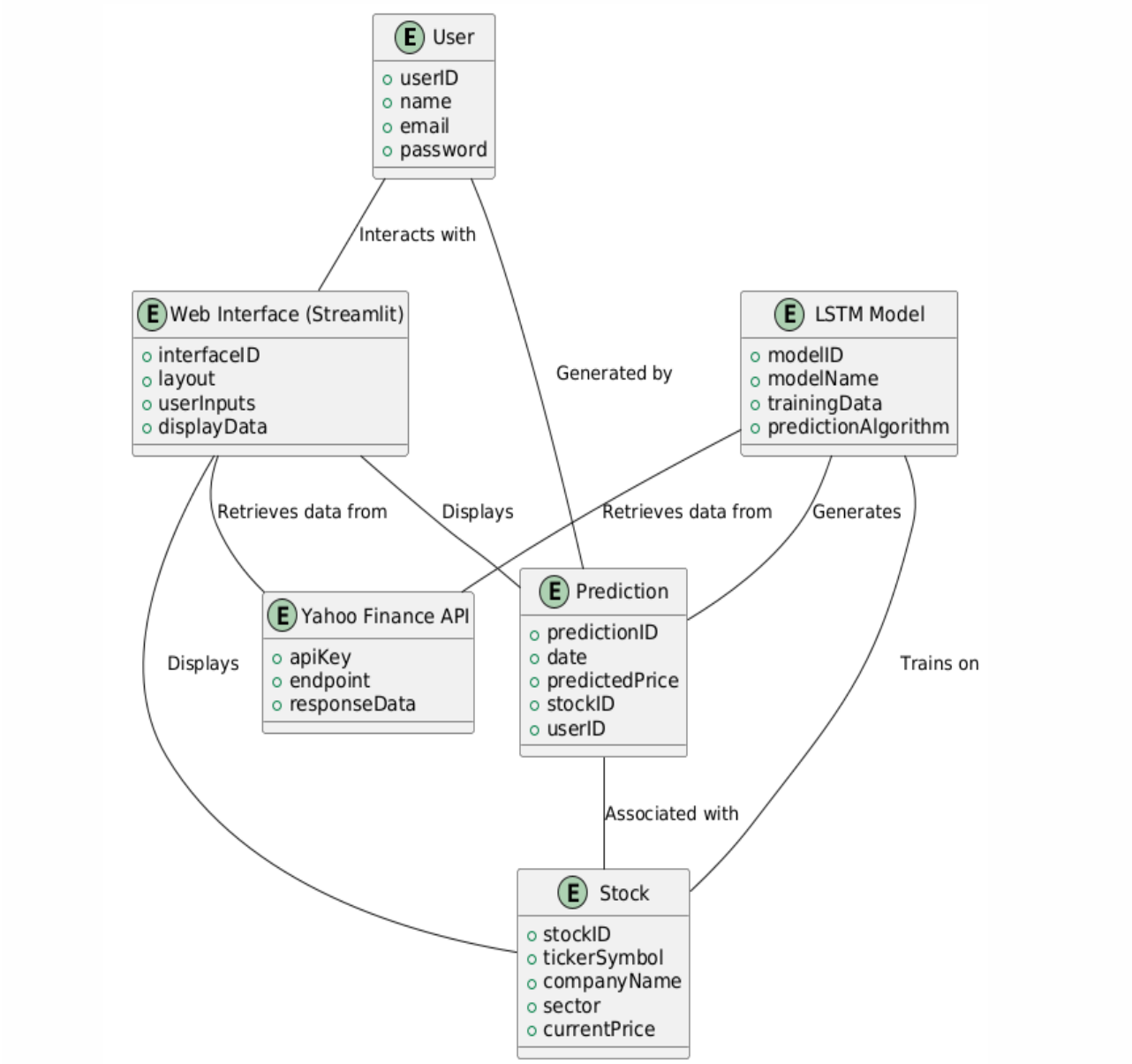


Figure. 4.2. Domain Model

4.3. Entity Relationship Diagram with data dictionary

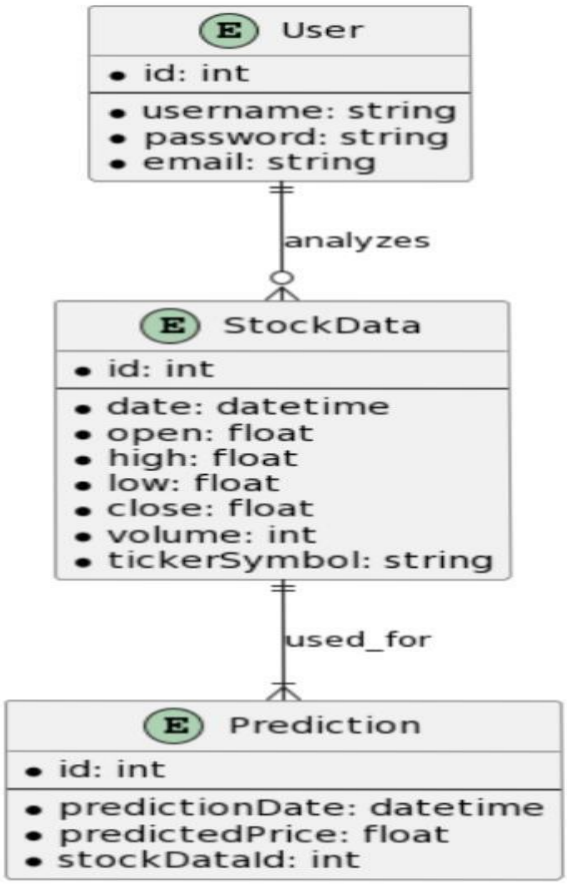


Figure. 4.3. Entity Relationship diagram

4.4. Class Diagram

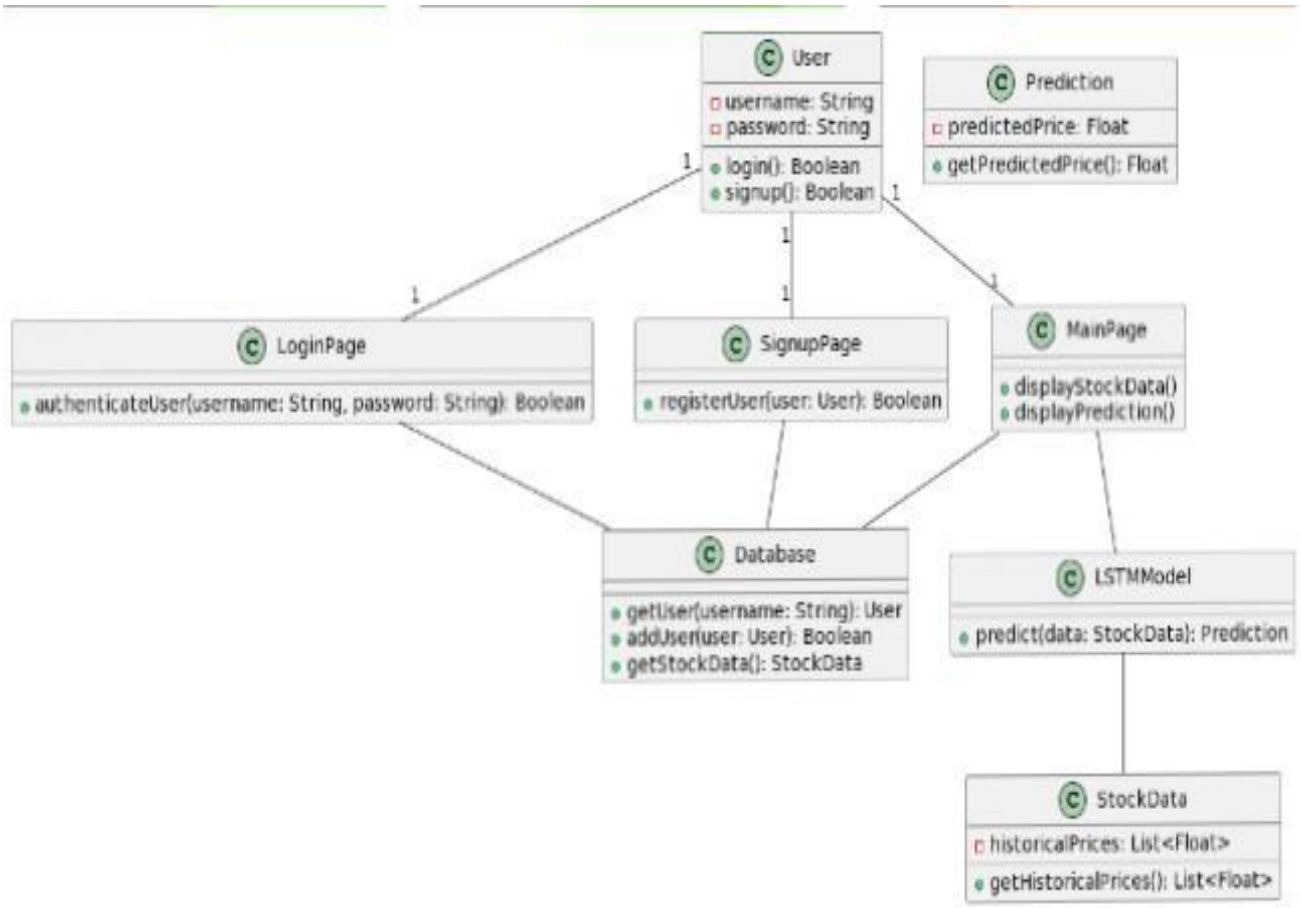


Figure. 4.4. Class Diagram

4.5. Sequence / Collaboration Diagram

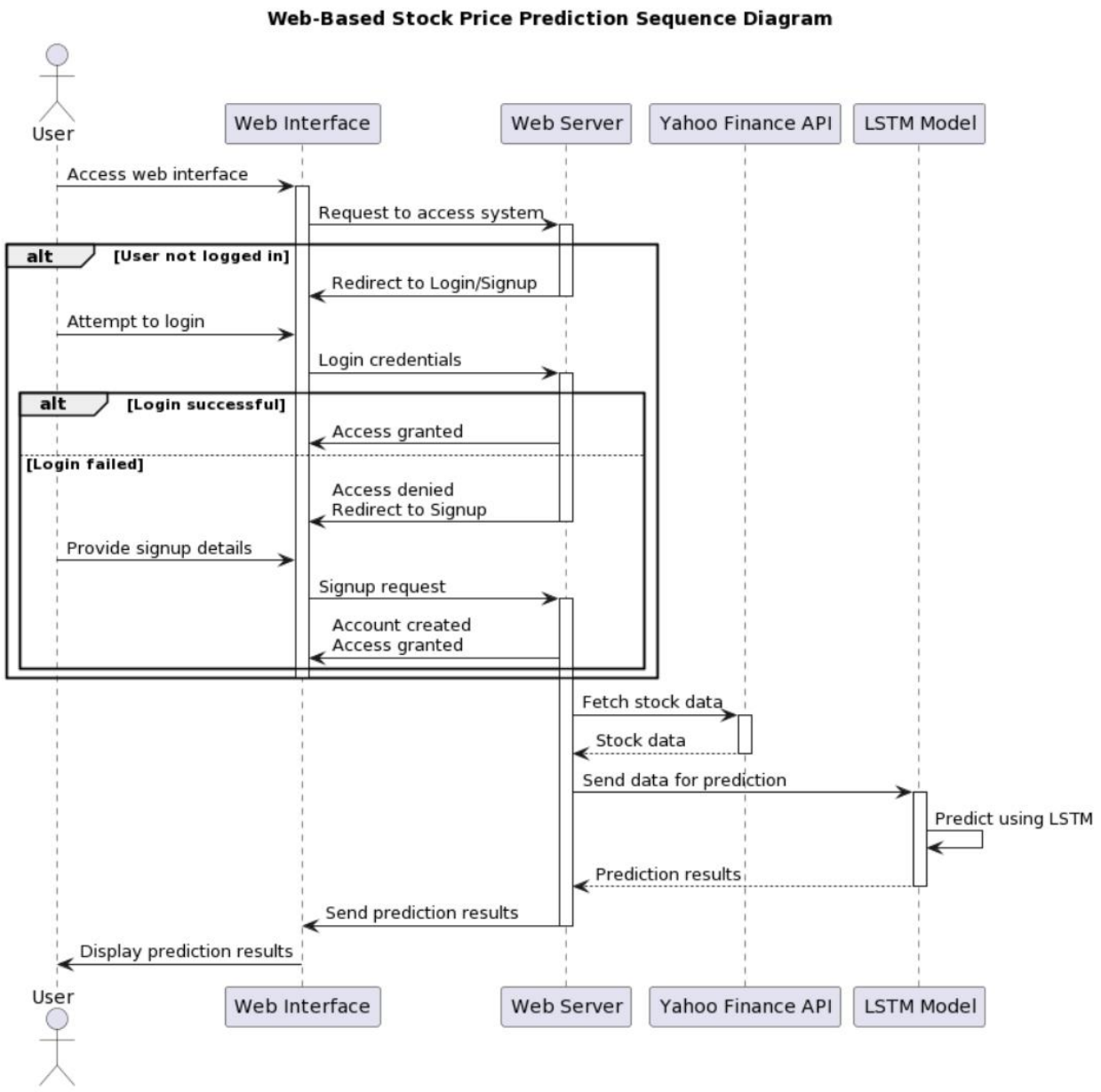


Figure. 4.5. Sequence Diagram

4.6. Operation contracts

Operation contracts is a concept of software engineering that highlights the obligations and benefits which are associated with the operations or functions of the system. They define the behaviour of a system operations and the outcomes that a user can expect when interacting with a particular operations or functions of the developed system.

4.7. Activity Diagram

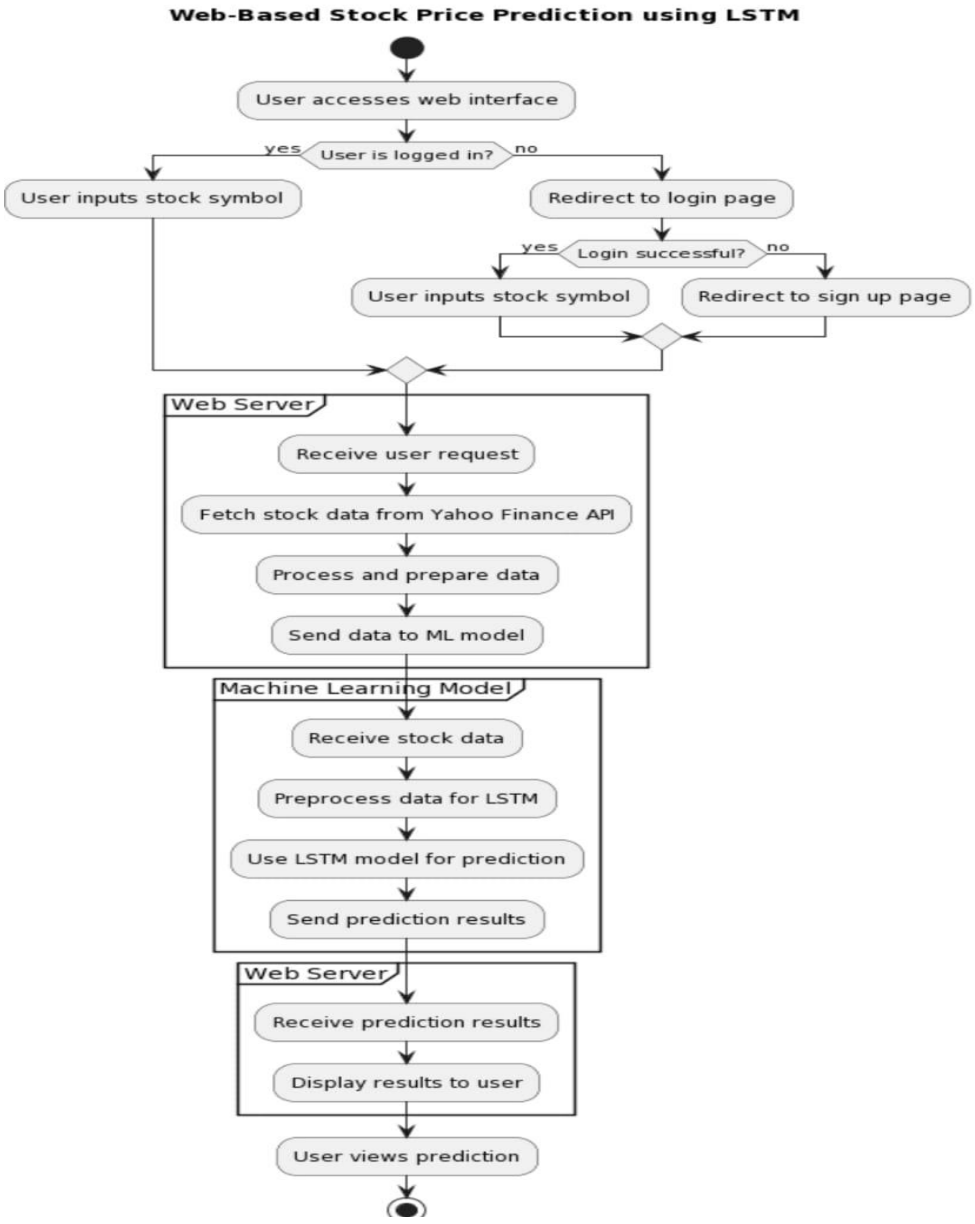


Figure. 4.7. Activity Diagram

4.8. State Transition Diagram

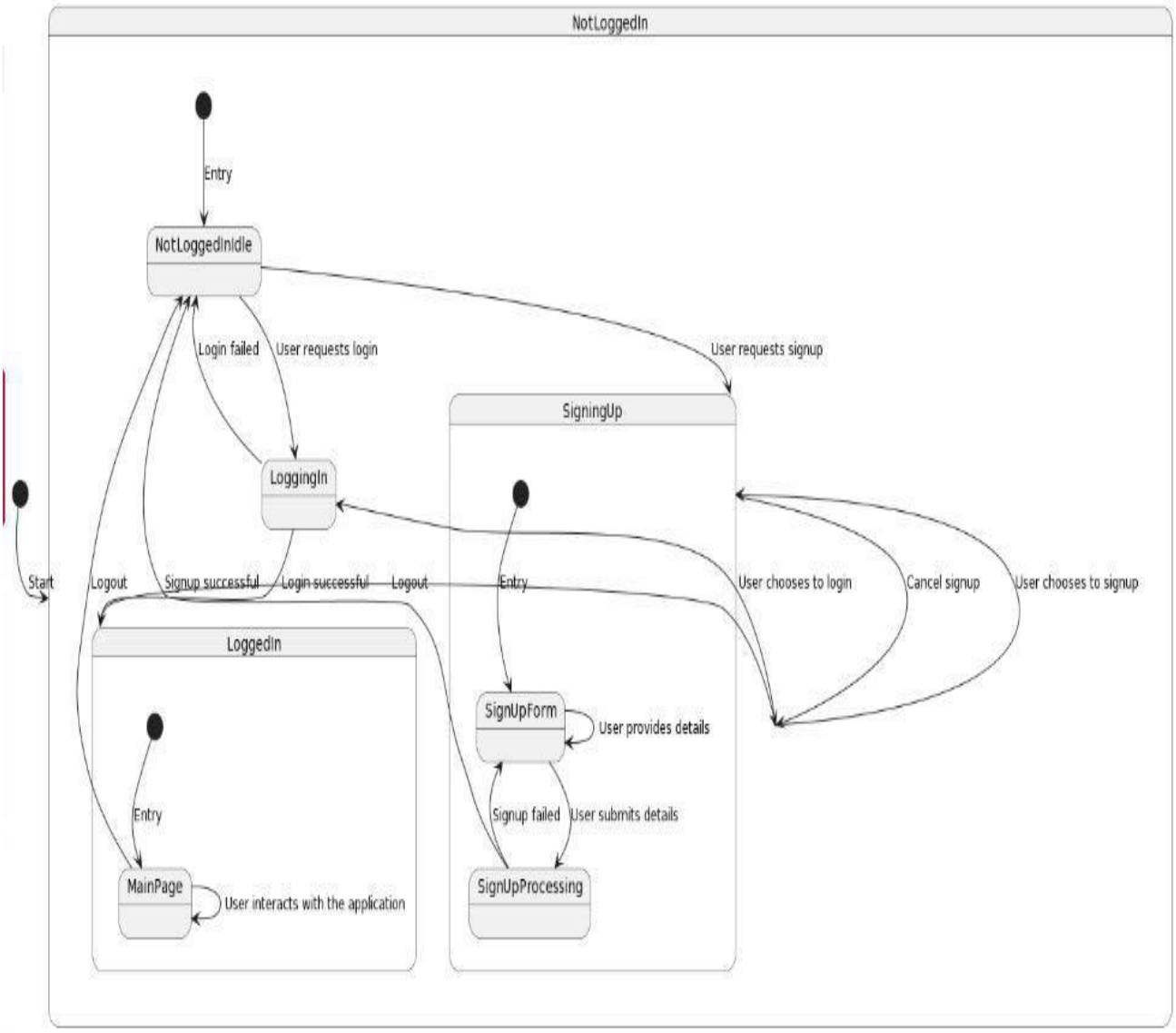


Figure. 4.8. State Transition Diagram

4.9. Component Diagram

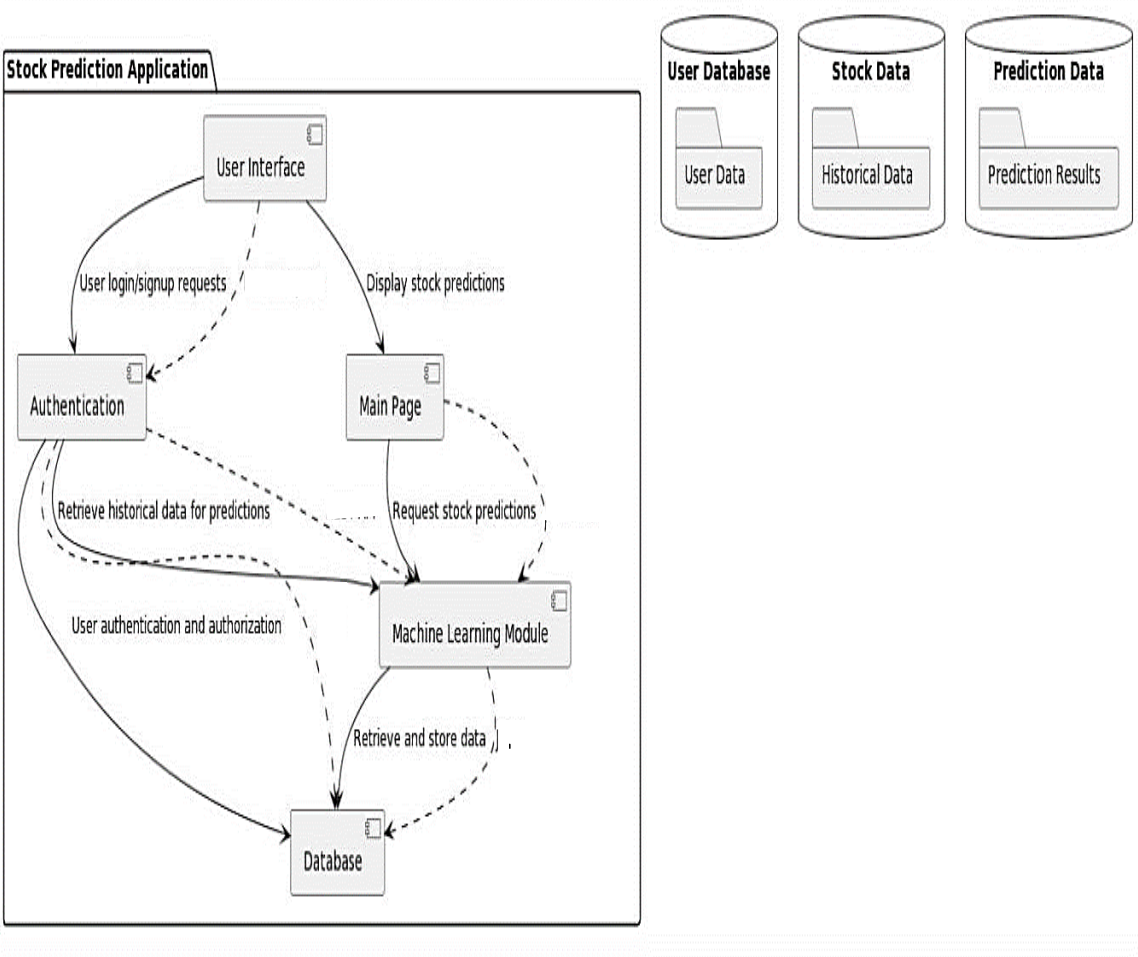


Figure. 4.9. Component Diagram

4.10. Deployment Diagram

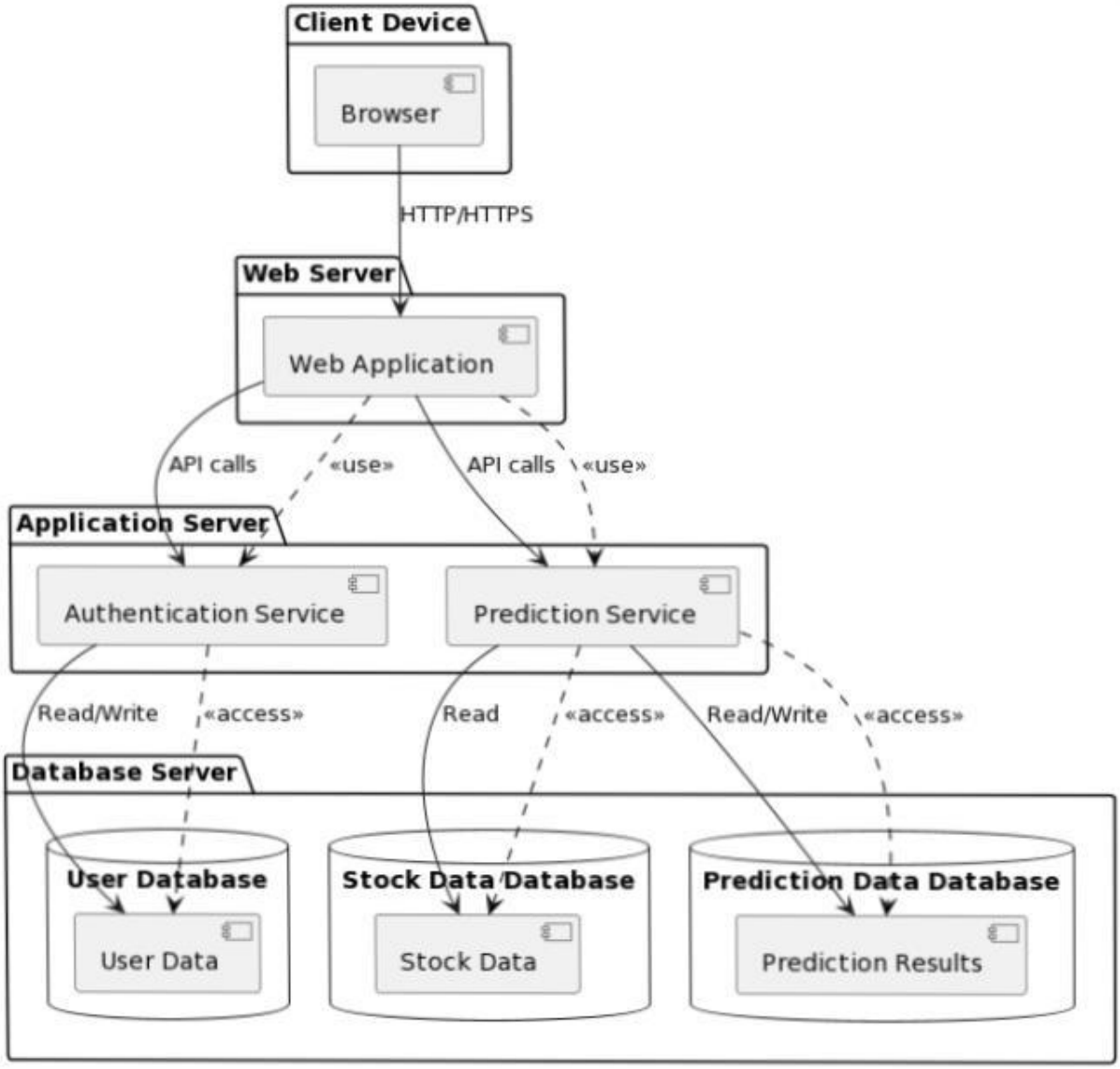


Figure. 4.10. Deployment Diagram

Chapter 5

Implementation

Chapter 5: Implementation

In this chapter we will go through all the techniques and other related components we used for the implementation of our product/application. For instance, we will look at what libraries we used, what coding standards we had to stick with and what type of environment we used for the deployment. This chapter mainly focus these kinds of things so that a user or reader could get more deeper understanding of the product/application we are developing.

5.1. Important Flow Control/Pseudo codes

Login/Signup:

function loginUser(username, password):

user = findUserInDatabase(username)

If user and password match(password, user.password):

return loginSuccess (user)

however: return loginFailure("Incorrect username or password");

function signupUser(username, password, email):

If userExistsInDatabase(username):

return signupFailure("username already exists");

however: saveUserToDatabase(username, password, email);

return signupSuccess()

Go to the main page:

function accessMainPage(user):

if isAuthenticated (user): .

Showroom () .

however: redirectToLoginPage () .

Stock forecast:

function requestStockPrediction(user, stocksymbol):

if isAuthenticated (user):

stockData = getStockData(stockIcon)

forecast = weForecast (stockdata)

Prophecy of the return however:

return error("User not authenticated");

General flow:

user = attemptLogin (username, password)

For example: User: showMainPage (user)

however: user = attemptSignup (username, password, email);

For example: User: stockSymbol = getInput("Enter stock symbol:");

prediction = requestStockPrediction(user, stockmark)

displayPrediction(prediction)

5.2. Components, Libraries, Web Services and stubs

Libraries: Below are some libraries that will help us development of our application;

- Streamlit/flask: we will use this for creating our web app's UI with Python.
- TensorFlow/Keras: For building and running our LSTM model to attain prediction results.
- Pandas: For data manipulation and analysis of our project's datasets.
- NumPy: For numerical computations of our project's datasets.
- Scikit-learn: For data preprocessing and additional machine learning tasks of our project's datasets and LSTM training and testing.

Components:

- User Interface (UI): This component of our application which is built with Stream lit/flask, this is the front-end where users interact with your application.
- Authentication System: This part of our project app, manages user login and signup processes.
- Prediction Engine: The core component of our project application that uses the LSTM model for stock price predictions.
- Database: This component our application stores user data, historical stock data, and prediction results.

5.3. Deployment Environment

Our Web Server: For the web server deployment Stream lit/flask will be used for running its own server.

Our Project's Database Server: In this, storing user data and possibly stock data for our project app, we might use a database server. We can also use Cloud-based solutions like Amazon RDS, Google Cloud SQL, or Azure SQL Database for our project's app as they provide scalability and ease of integration for the developers in their applications.

5.4. Tools and Techniques

Development Tools and Languages

- Python: We will use python programming language for both backend coding and machine learning model development.
- Stream lit: It's a python library will be used for creating the web application interface.
- TensorFlow/Kera's: Will be used for designing, training, and deploying the LSTM model.
- Pandas/NumPy: For data manipulation & numerical computation in the model training and testing during of our project.

Using some good machine Learning Techniques

- Long Short-Term Memory (LSTM) Networks: Core technique which we will be using for predicting the stock prices based on historical data.
- Data Preprocessing: Preprocessing techniques like normalization, handling missing values, and data transformation to prepare the dataset for the LSTM model.

Using Web Development Techniques

- Responsive Design: We will ensure the interface is user-friendly and is adaptable to different screen sizes across different devices.

5.5. Best Practices / Coding Standards

Security: We will implement secure user authentication using pre-defined practices for our project's app. Also, we will encrypt the passwords before storing them in the database.

Best Practices of stream lit: Using the Stream lit/flask or WordPress for our project app's user friendly interaction (e.g., st.button, st.text_input).

Machine Learning Model: Our project will save and load the trained LSTM model in an efficient way. We will also implement error handling for model prediction failures.

Code Structure and Organization: Organizing our project's code into clear, modular functions and classes so that it will help us in understanding the code more clearly and can also help us when we need to do any further kind of modification.

User Experience: Ensuring a responsive and user-friendly design for our project application will greatly help us in the development of our application.

5.6. Version Control

We will use git version control for our application.

Chapter 6

Testing and Evaluation

Chapter 6: Testing and Evaluation

This chapter revolves around one of the main stages of developing a software which is testing using different techniques and methods. In this chapter, we have used different white box and black box techniques to test the various aspects of our project such as testing individual components or units to thoroughly test them. Similarly, we also checked the load a software can handle when it is loaded with increased number of users simultaneously. Also, we tested the different use cases we developed earlier in the chapter 3 of this report.

6.1. Use Case Testing

Table 6.1. Use case testing

	Steps	**Description**
Login Success Scenario A: Actor S: System/Backend	**1** **2** **3** **4**	**S**: Enter Your login Details **A**: Enter his/her details **S**: Validate Login Details **S**: Allows access to main page of the website
Extensions	 **2a** **2b**	**Login Failed** **S**: Show error message and asks user to again enter login details **Login Failed** **S**: User not registered. Asks user to register first and then login

	Steps	**Description**
Prediction Success Scenario A: Actor S: System/Backend	**1** **2** **3** **4**	**S**: Enter a company stock's symbol **A**: Enter stock symbol **S**: Checks whether it exists or not **S**: Shows the Prediction results in a graph form
Extensions	 **2a** **2b**	**Invalid stock symbol** **S**: Display error message and asks user to enter a valid stock symbol for prediction **Invalid stock symbol** **S**: Symbol does not exist. Asks user to try again

6.2. Equivalence partitioning

Table 6.2. Equivalence Partitioning

Input a Stock Symbol		Submit

Test Scenario	**Test Scenario Description**	**Expected Outcome**
1	Enter “AAPL” symbol in input field	System should accept and show Predictions about AAPL
2	Enter “123ABC “in input field	System should not accept and display error message
3	Enter “xyzqdgoiossdsd “in input field	System should not accept and display error message
4	Enter “ “in input field	System should not accept, display error message and asks the user to enter something in the input field
5	Enter “@GOOGL “or “AMZN#” in input field	System should not accept, display error message and asks the user to enter correct symbol

6.3. Boundary value analysis

Table 6.3. Boundary Value Analysis

	Test scenario Description	Expected Outcome
1	Companies having length of ticker symbol between 3 – 5 characters	System should accept it
2	Boundary Value: Empty String “”	System should not accept it and asks the user to enter valid stock symbol
3	Boundary Value: Input having more than 6 characters	System should not accept it and asks the user to enter valid stock symbol

6.4. Data flow testing

First, we'll identify variables: -

1. __login__obj
2. LOGGED_IN
3. username
4. selected_page

Then Analyse the variables: Now we will determine for each variable where and how it is defined (DEF) and where it is used (USE). The DEF contains the statements where a variable receives a value. The USE contains the statements where the values are being referenced.

Finally perform Data flow testing: -

1. **_login__obj:**
- DEF: The variable is defined during the instantiation of the __login__ object.
- USE: It is used in the subsequent lines to build the login UI and retrieve the username.
2. **LOGGED_IN:**
- DEF: The variable gets assigned a value which is based on the result of __login__obj.build_login_ui().
- USE: It is used in the conditional statement to determine whether the user is logged in.
3. **username:**
- DEF: The variable is assigned the value returned by __login__obj.get_username().
- USE: It is used to display the username.
4. **Selected page:**
- DEF: The variable is assigned that value which is selected from the sidebar.
- USE: This is used to determine which page content to display.

6.5. Unit testing

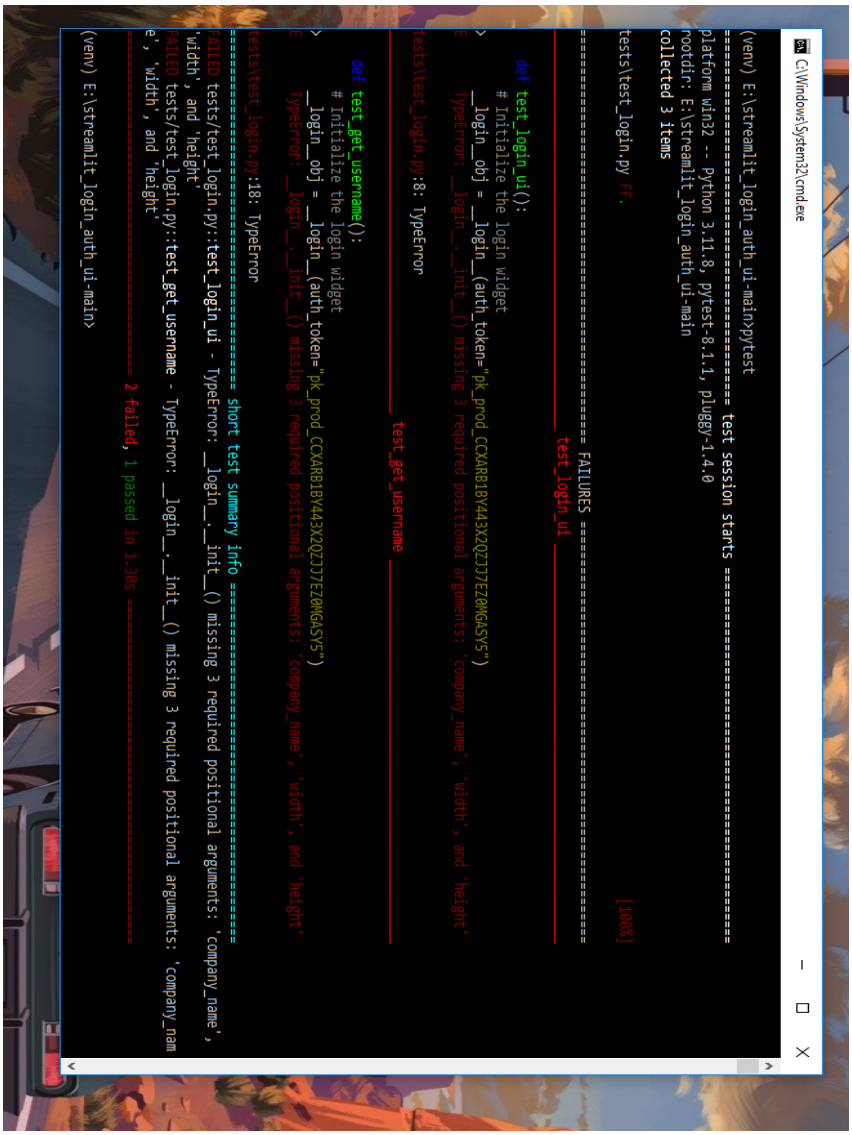


Figure. 6.5. Unit Testing

6.6. Integration testing

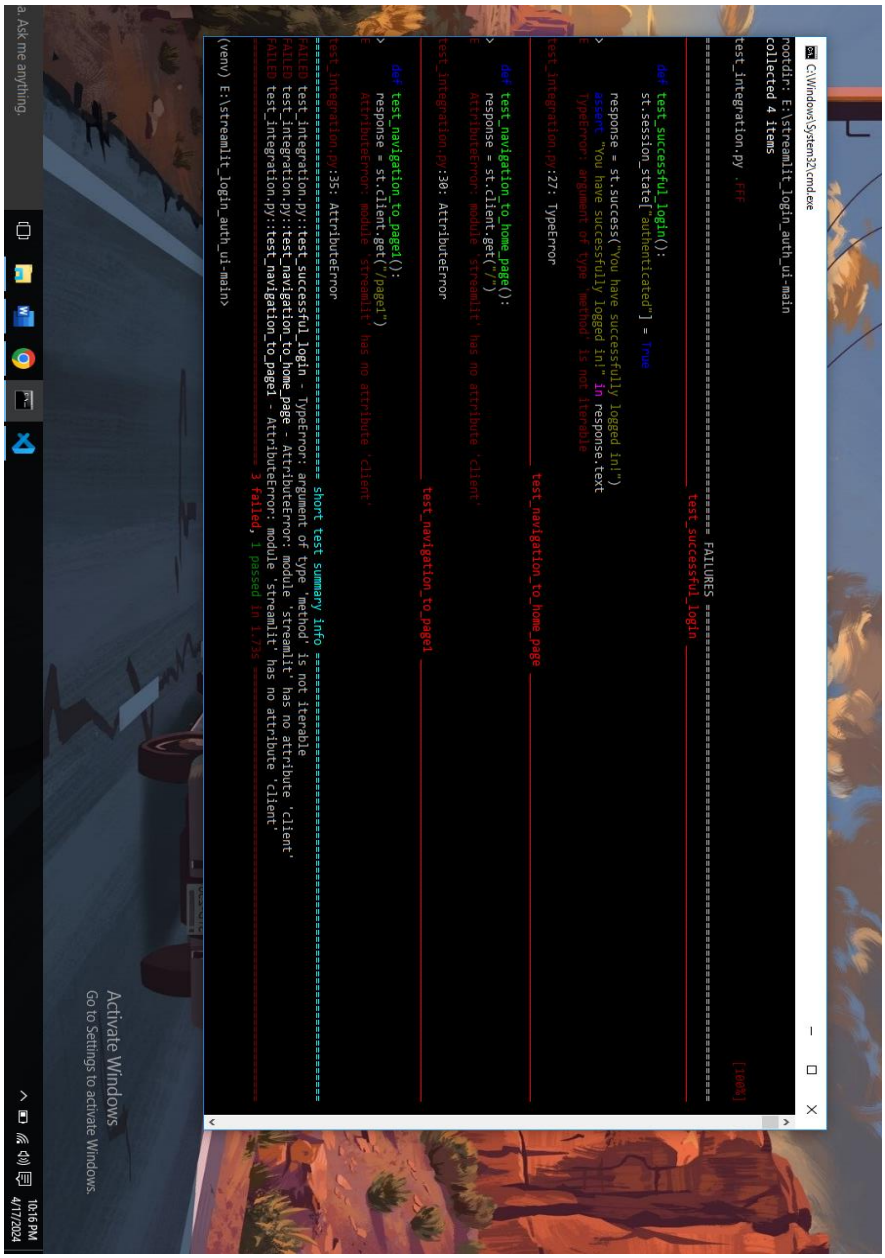


Figure. 6.6. Integration Testing

6.7. Performance testing

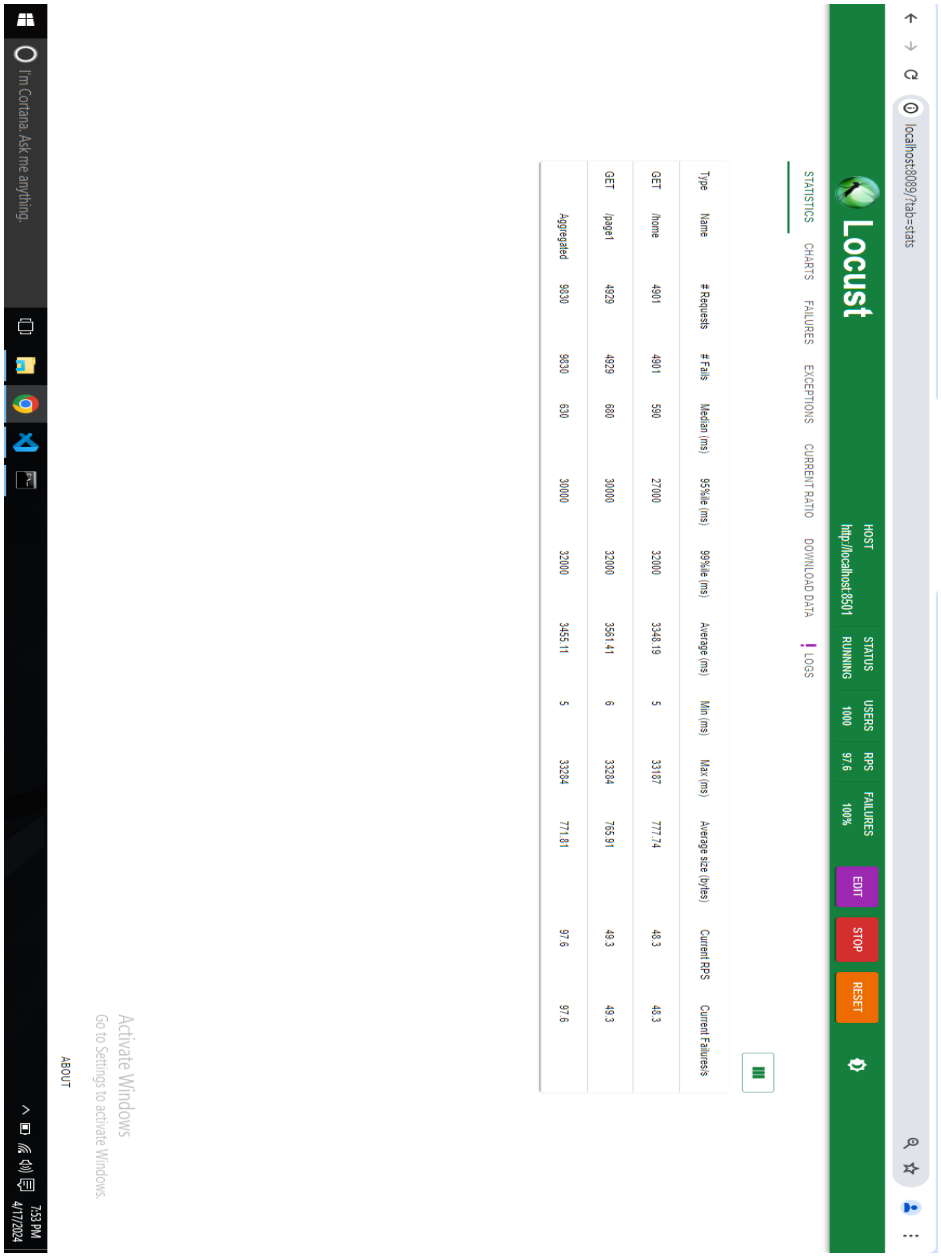


Figure. 6.7. Performance testing

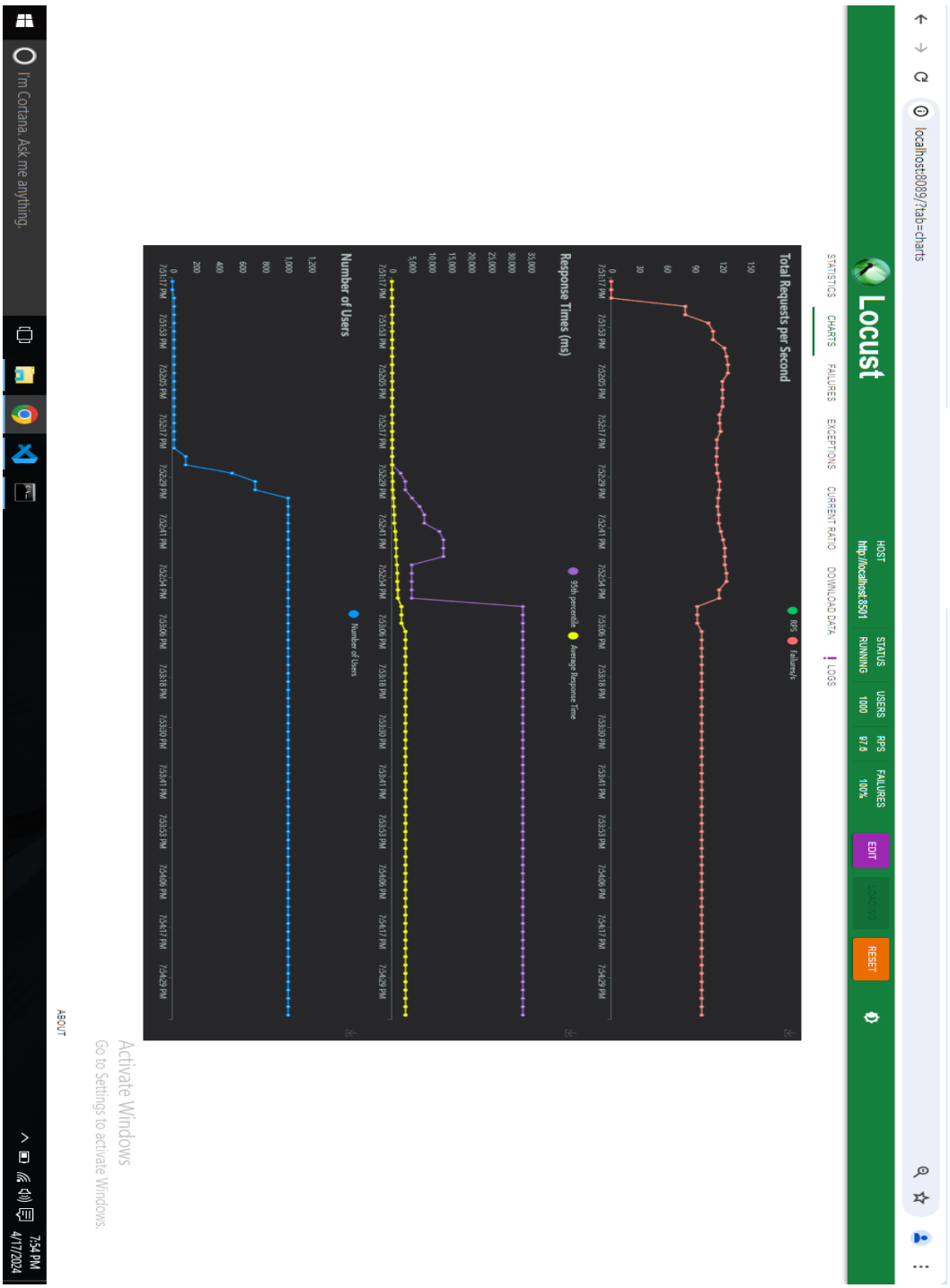


Figure. 6.7.1. Performance testing

6.8. Stress Testing

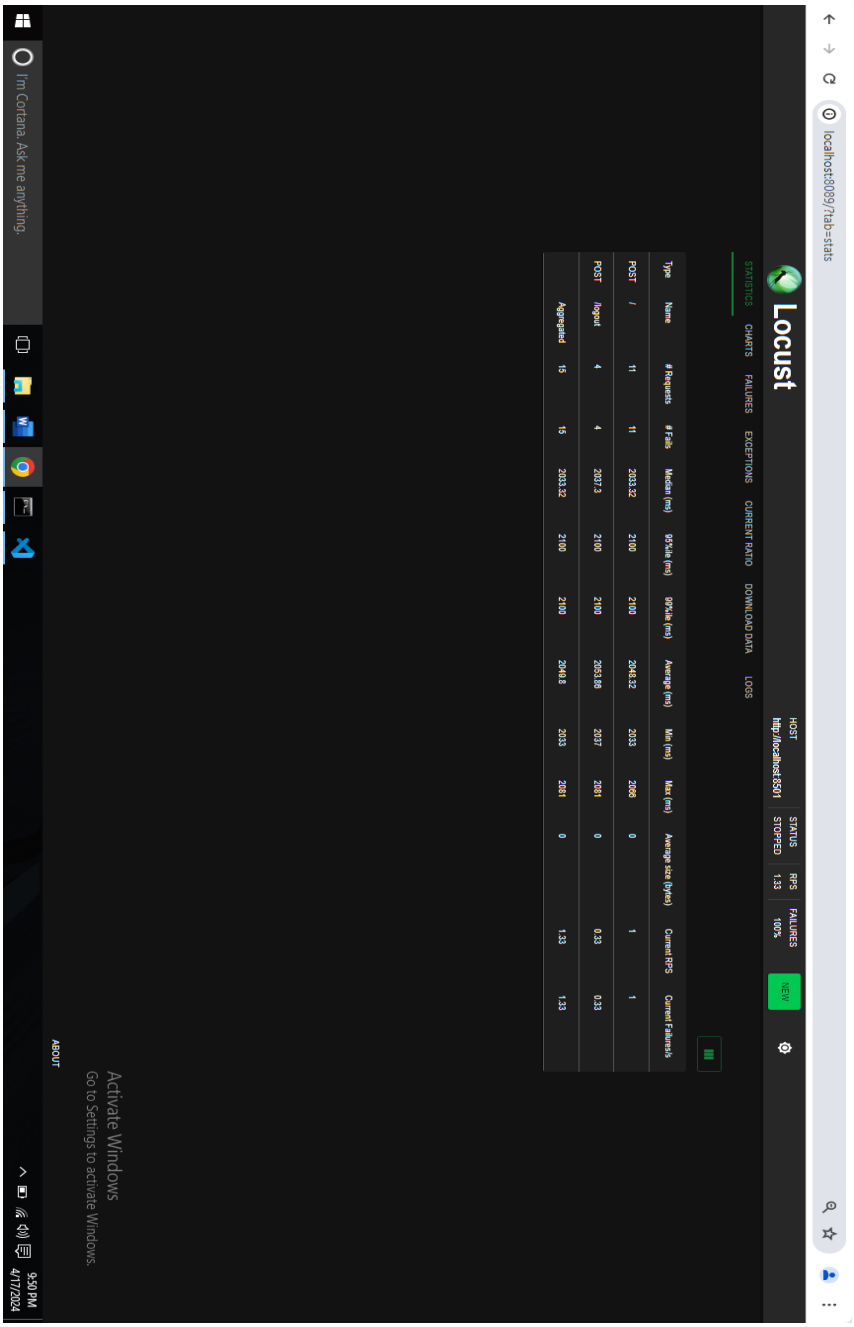


Figure. 6.8. Stress Testing

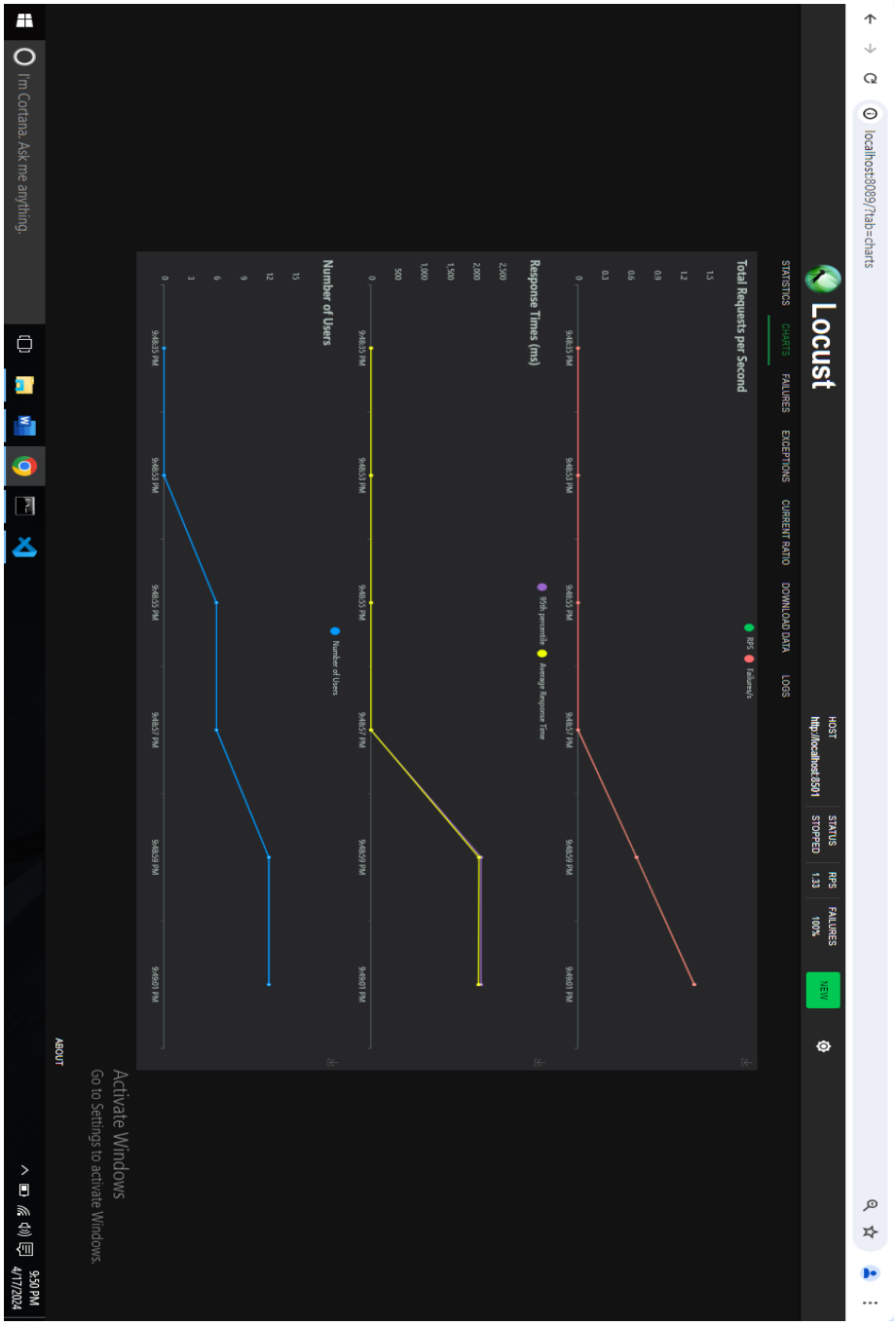


Figure. 6.8.1. Stress Testing

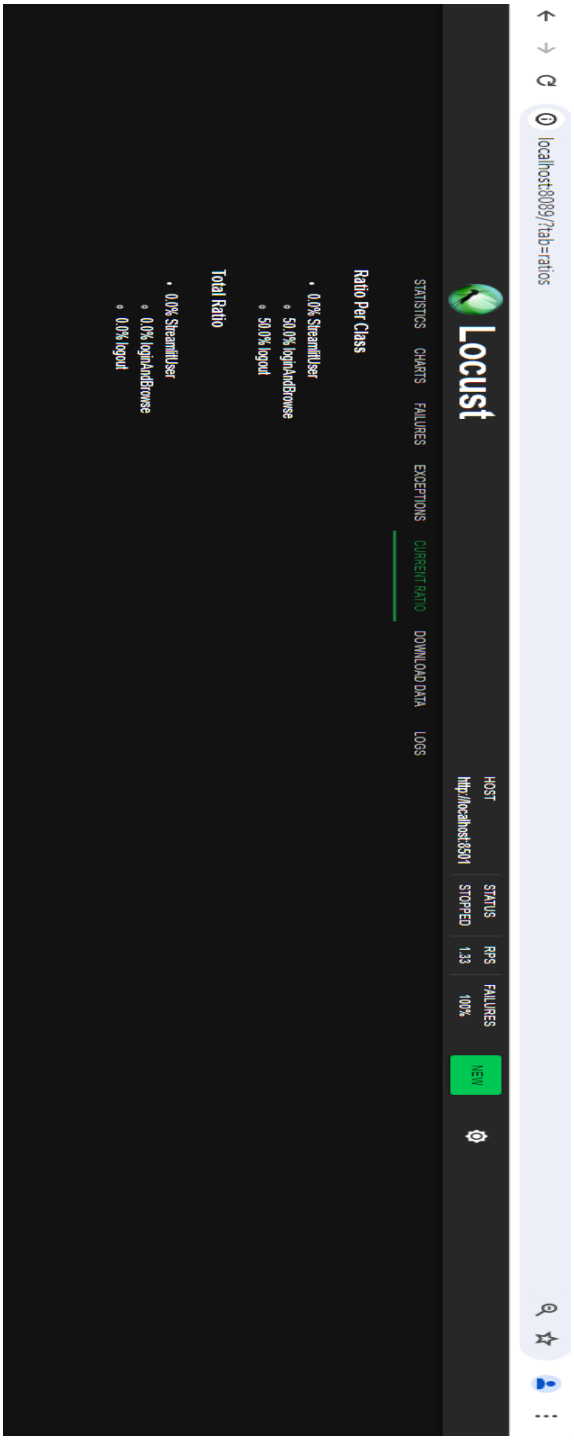


Figure. 6.8.2. Stress Testing

Chapter 7

Summary, Conclusion and Future Enhancements

Chapter 7: Summary, Conclusion & Future Enhancements

7.1. Project Summary

The stock market is a collection of various companies listed publicly to raise funds for their company. People buy and sell their interest stocks in order to make huge profits out of it. But doing it so precisely is not easy as it seems because stock market changes every second and no one knows whether the values are going to rise or fall. For this purpose, we are building a web-based application where a user can see what particular stock will rise or fall. We will be using machine learning's LSTM model approach for this as it has been proved to show excellent results related to time series data. Along with this the user can also compare stocks of two different companies and the user can also download dataset of a stock in a csv format so that they can analyse it according to their interest.

7.2. Achievements and Improvements

There are number of achievements we have got while the development of this project. Like for instance we not only used LSTM for predicting the prices of the stocks but we also made the users to take more deep insights about a particular stock by giving them an option where they can compare different stocks performances and also giving them real time stock prices at one place. Also, our model, when subjected to testing, gave almost accurate results which in turn means that the model is accurate and can predict or forecast the prices of the stocks up to some accuracy.

7.3. Critical Review

Overview: This project of ours called Stocks Price Prediction uses machine learning's LSTM approach to predict or forecast the prices of stocks for the upcoming days. The application has been designed keeping in mind the interface a user would love to interact with.

Methodology:

For this project, we have used historical datasets for the training the LSTM model and then testing the model to find how accurate the predictions or results are. The datasets will be fetched by yahoo finance API.

Conclusion: So, the critical review defines how the system is going to work and what weaknesses or strengths our project application is going to have.

7.4. Lessons Learnt

One of the main lessons we learnt while developing this project was that we should have to use clean and concise historical dataset for the training and testing of the LSTM model. Similarly, we learnt that while using machine learning we should have to consider various factors for this software because it will help in developing the accuracy of the model further.

7.5. Future Enhancements/Recommendations

- Just like any other software, our project application will also need an updating regularly so that we can ensure that it remains up to date with the current technologies of the world and we can further improve it.
- For instance, besides using historical dataset of the stocks we can also use various data like data from news, data from social media platforms and any other alternative data that we can use with LSTM to further improve its accuracy.
- Also, we can use different algorithms like ARIMA, Linear regression, random forest or any other algorithm which is best for time series data prediction with the LSTM to enhance the prediction accuracy results.
- We can use almost near to real time datasets for training the LSTM model so that we remain up to date with the latest changes of market trends.

Appendices

Appendix A: User Manual

In this part of our project application documentation, users of app are guided thoroughly about the project that has been developed, and more specifically on how to use the system more clearly, concisely and easily.

Appendix A: Project Overview

This is a web-based application using machine learning's LSTM approach which helps the user to get the predictions about their interest stock symbol.

A.1. How to Interact/use our application

You can simply login in to application and select the prediction feature from the navigation bar and from there you can enter a valid stock symbol which will get you the predicted results.

A.1.1. Different aspects of our project's application

You can enter any stock symbol as long as it is valid and does exist in the stock market. If you do not enter the correct ticker symbol, then our project application will again ask to enter the input correctly. You are shown the results in a form of graph for the easiness and to understand better.

A.1.1.1. Where the data for our project's app comes from?

The data that LSTM uses for its training and testing the model comes from yahoo finance which is integrated in our project application through yahoo finance API.

Appendix B: Administrator Manual

This section of the appendix is clearly defining the roles the admins have in the project and how they can use these functionalities apart from the other users of the applications.

B.1. User Management

B.1.1. User Roles and Permissions

Our application supports multiple user roles, including admins and viewers. Administrators have full access to all functionalities of the project. Viewers have only access to prediction results and historical dataset of company stocks.

B.1.1.1. User Account Management

In our application, an admin can have the power to create and delete user accounts as deemed appropriate or necessary.

Reference and Bibliography

Reference and Bibliography

1. Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.
2. Moritz, B., & Zimmermann, T. (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. Available at SSRN 2740751.
3. J. Bollen and H. Mao. Twitter mood as a stock market predictor. Computer, 44(10):91–94, 2011.
4. a. Mittal and a. Goel. Stock Prediction Using Twitter Sentiment Analysis. Tomx.Inf. Elte.Hu, (June), 2012.
5. Veliota Drakopoulou. “A review of fundamental and technical stock analysis techniques”. In: Journal of Stock & Forex Trading 5 (2016).
6. Pengfei Yu and Xuesong Yan. “Stock price prediction based on deep neural networks”. In: Neural Computing and Applications 32 (2020), pp. 1609–1628.
7. Gaoshan Wang et al. “The Effect of Online Investor Sentiment on Stock Movements: An LSTM Approach”. In: (2020).

ORIGINALITY REPORT

10%	7%	1%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	2%
2	Submitted to Higher Education Commission Pakistan Student Paper	2%
3	Submitted to Bournemouth University Student Paper	1%
4	www.ijsrd.com Internet Source	1%
5	www.kdnuggets.com Internet Source	1%
6	core.ac.uk Internet Source	1%
7	Submitted to Study Group Australia Student Paper	<1%
8	pdfcoffee.com Internet Source	<1%
9	ruor.uottawa.ca Internet Source	<1%

10	www.ijraset.com Internet Source	<1%
11	"Big Data Analytics for Cyber-Physical System in Smart City", Springer Science and Business Media LLC, 2020 Publication	<1%

Exclude quotes On Ex Exclude matches Off