# AUTOMATION IN DATA ANALYTICS USING ROBOTIC PROCESS AUTOMATION AND ARTIFICIAL INTELLIGENCE

By

GAN YU NYUK

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfilment of the requirements

for the degree of

BACHELOR OF COMPUTER SCIENCE (HONOURS)
Faculty of Information and Communication Technology
(Kampar Campus)

JUNE 2024

**UNIVERSITI TUNKU ABDUL RAHMAN**

# REPORT STATUS DECLARATION FORM

**Title**: **Automation in Data Analytics using Robotic Process Automation and Artificial Intelligence**

**Academic Session**: June 2024

I GAN YU NYUK

**(CAPITAL LETTER)**

declare that I allow this Final Year Project Report to be kept in

Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.

Verified by,

(Author's signature)

(Supervisor's signature)

**Address**:

11-11-03 Awana Condo,

Jalan Bukit Kukus,

11060 Ayer Itam, Pulau Pinang

Ts Dr. Ooi Joo On

Supervisor's name

**Date**: 11/9/2024

**Date**: 13/9/2024

| Universiti Tunku Abdul Rahman | | | |
|---|---|---|---|
| Form Title : **Sample of Submission Sheet for FYP/Dissertation/Thesis** | | | |
| Form Number: **FM-IAD-004** | Rev No.: **0** | Effective Date: **21 JUNE 2011** | Page No.: **1 of 1** |

**FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY**

**UNIVERSITI TUNKU ABDUL RAHMAN**

Date: ___11/9/2024________________

**SUBMISSION OF FINAL YEAR PROJECT**

It is hereby certified that ______***Gan Yu Nyuk***____________________ (ID No: __***21ACB06785***__ ) has completed this final year project entitled "_AUTOMATION IN DATA ANALYTICS USING ROBOTIC PROCESS AUTOMATION AND ARTIFICIAL INTELLIGENCE" under the supervision of _____Ts Dr. Ooi Joo On______ (Supervisor) from the Department of ____Digital Economy Technology_____, Faculty of ___Information and Communication Technology___.

I understand that University will upload softcopy of my final year project in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_______________

(*Gan Yu Nyuk*)

*Delete whichever not applicable

## DECLARATION OF ORIGINALITY

I declare that this report entitled "**AUTOMATION IN DATA ANALYTICS USING ROBOTIC PROCESS AUTOMATION AND ARTIFICIAL INTELLIGENCE**" is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature : ______________________________

Name : Gan Yu Nyuk________________

Date : 13 September 2024________________

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my supervisor, Ts Dr Ooi Joo On, who has given me the opportunity to engage myself in this project. His valuable guidance and instruction have served as major contributor towards the completion of the project. This is a good start to establish a career in Artificial Intelligence and Robotic Process Automation field.

Finally, I would also like to extend my appreciation to my family for always supporting me in my project development. Their support has encouraged me to be determined in completing this project. Without their support, I would not have come this far in working on this project.

## ABSTRACT

This project provides an integrated solution for beginner traders who do not have prior knowledge in analysing the market trends and have no idea where to start their trading journey. Hence, this project focuses on developing a web application by leveraging Artificial Intelligence (AI) and Robotic Process Automation (RPA), to automate data analytics and market predictions, providing valuable insights to users for their investment decisions. This project utilises React, an open-source JavaScript library, for building the frontend interface of the application, where it offers users a user-friendly platform to access real-time stock using RPA. Another tool being used in this project is a lightweight web framework written in Python, called Flask, where it contains features that make building web applications in Python easier. Flask is used for building the backend of the application. Moreover, the integrated development environment (IDE) chosen to carry out the project is Visual Studio Code.

This project currently comprises two functionalities, which are, viewing market information and predicting market trends. The first feature allows users to access real-time market data fetched from Yahoo Finance using Robotic Process Automation (RPA) implemented with UiPath. This is to provide comprehensive data analytics and visualization, where users able to get insights about the market trends and historical price movements. Furthermore, this project incorporates Long Short-Term Memory (LSTM) model for predictive analytics, where it can provide action recommendations (buy, hold, or sell) to users regarding the upcoming market trend. In a nutshell, the project aims to provide an automated trading analytics by leveraging an LSTM model for stock prediction and a sentiment analysis to analyse market news

## TABLE OF CONTENT

## LIST OF FIGURES

# LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| *RPA* | Robotic Process Automation |
| *AI* | Artificial Intelligence |
| *LSTM* | Long Short-Term Memory |
| *KNN* | K-Nearest Neighbour |
| *ANN* | Artificial Neural Network |
| *RF* | Random Forest |
| *RMSE* | Root Mean Square Error |
| *MAPE* | Mean Absolute Percentage Error |
| *MBE* | Mean Bias Error |
| *MSE* | Mean Square Error |
| *MAE* | Mean Absolute Error |
| *SMA* | Simple Moving Average |
| *EMA* | Exponential Moving Average |
| *RSI* | Relative Strength Index |
| *MACD* | Moving Average Convergence Divergence |
| *BB* | Bollinger Bands |

# CHAPTER 1 INTRODUCTION

In this chapter, the motivation, research objectives, project scope, contribution and report organization of this project will be presented.

## 1.1 Problem Statement and Motivation

One of the reasons many people are reluctant to invest in the financial market is because of the fear of losing money. This results from the lack of knowledge about investing or trading. However, to be proficient in analysing market trends, it requires years of experience in researching. This poses a challenge for the beginners because they do not know where to start. It is also nonetheless a tedious task for traders to navigate from websites to websites to monitor the market and it requires a lot of manual processes to extract recent news and stock prices for the intended tickers. Hence, providing access to real-time market data, comprehensive visualisations through automated process would ease beginners in well-informed decisions.

Moreover, the market trend is uncertain and volatile as it is influenced by various factors such as, the global economy, unexpected events, politics, company’s financial performance and so on. For example, unexpected event like COVID-19 pandemic has introduced unprecedented levels of uncertainty, causing drastic market fluctuations. Hence, predicting market trends is a difficult task due to the market’s volatility and uncertainty. Traders often require prior skills in predicting the market trend by using varieties of analysis such as, technical analysis, sentimental analysis and so on. Therefore, by leveraging Artificial Intelligence such as Deep Learning, it can automate the process of predict the future market trend and provide more valuable insights that can be considered for traders’ decisions.

## 1.2 Research Objectives

The aim of the project is to build a web application to ease the work of traders in analysing the future trend of the market as well as making decisions in investments. The web application will fetch historical market data to predict the future trend of the market and provide useful insights to users. Deep Learning model will used in this project to train on the historical data and predict the future market price. Moreover, the application will also have a dashboard to show users the analytics of the real-time market data. This project seeks to streamline traders' task by providing an automated solution and empower users to make more informed decisions with confidence.

- To implement an LSTM-based deep learning model to forecast future market prices. This model will use historical stock data to make accurate predictions and provide insights into potential investment strategies.
- To utilise sentiment analysis to evaluate the impact of news and social media on stock prices. Integrate RPA technology to extract real-time data from various sources, enhancing the sentiment analysis process and providing up-to-date insights for better decision-making.
- To create an automated trading bot that leverages both the LSTM model and sentiment analysis to make informed trading decisions. The bot will execute trades based on predicted stock prices and sentiment insights, streamlining the trading process for users.

## 1.3 Project Scope and Direction

This project aims to assist targeted users in making confident decisions for investments. It automates repetitive tasks that traders need to do to predict future price of tickers using historical data, such as technical analysis and fundamentals analysis, where it is repetitive when users must manually do the analysis and predict the future marker prices. Therefore, this application utilises Long Short-Term Model, a deep learning algorithm, to predict the future trend of the ticker and recommend actions, by analysing on the historical data fetched from yahoo finance. Moreover, RPA sequence is also created in UiPath to extract real-time data from open-source websites such as yahoo finance. This extracted data will be used to generate charts to visualise

the historical data of the ticker, so that users can better understand the trend of the ticker. The integration of AI and RPA into the web-based application will be implemented using React and Flask as the frontend and backend of the application respectively.

## 1.4 Contributions

The project develops and implements a predictive analytics model for stock market forecasting. By leveraging historical data and advanced deep learning technique, this model aims to automate predictions of stock prices, improving the efficiency of traders, and investors in their decision-making processes. For example, it can forecast whether a specific stock is likely to rise or fall soon, allowing traders to take appropriate positions. Hence, traders need not analyse the market data manually using investments analysis methods to predict the market price, Other than that, the project integrates Robotic Process Automation (RPA) into the data analytics workflow. RPA automates repetitive tasks, such as data extraction, saving time and reducing errors. This integration enhances the efficiency of data analytics processes. By integrating RPA, the project streamlines and accelerates data-related processes.

## 1.5 Report Organization

In Chapter 2, some past research related to this project are reviewed. Then, the project methodology, system design, model design, implementation issues and project timeline are presented in Chapter 3. Furthermore, Chapter 4 describes the preliminary work for FYP1. Furthermore, Chapter 5 summarises the conclusion of the project.

# CHAPTER 2 LITERATURE REVIEWS

## 2.1 Previous Works on Automated Trading Bot

### 2.1.1 3Commas

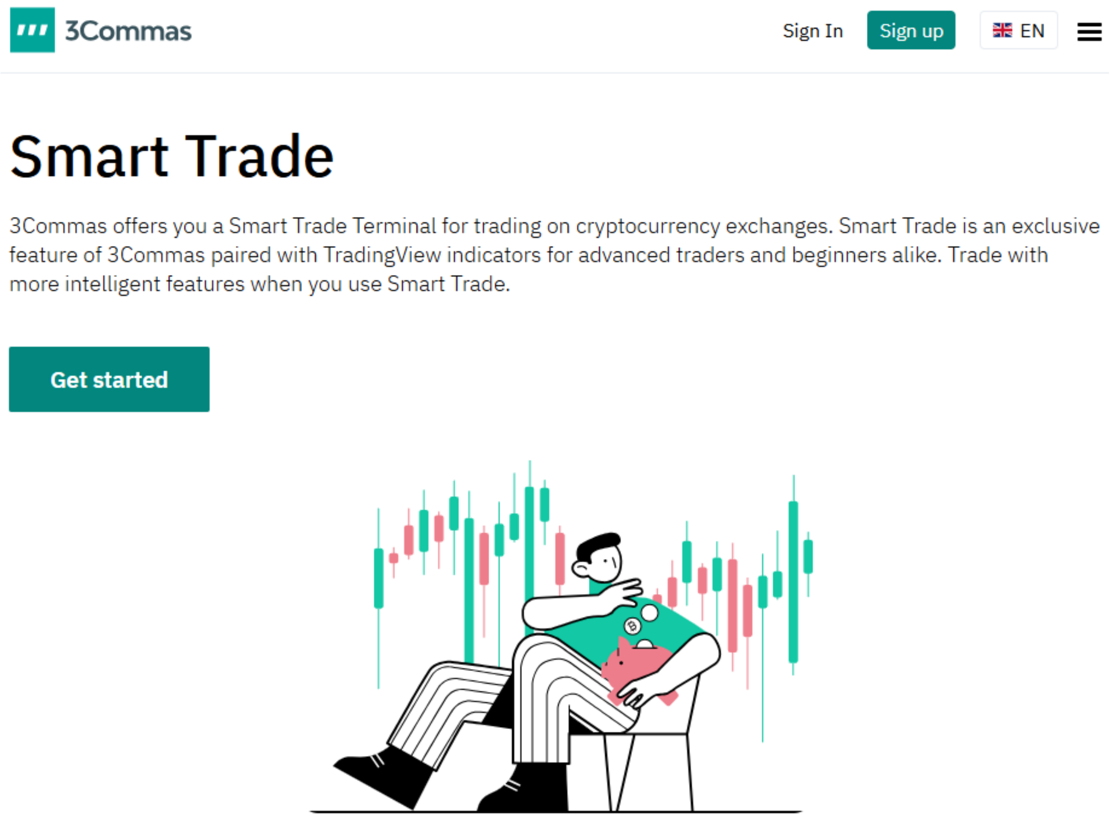



*Figure 2. 1 Web Application of 3Commas*

3Commas is a comprehensive trading platform known for its advanced tools for automated trading. This platform offers a variety of trading bots, such as Grid Bots, DCA Bots and Options Bots, which is highly customised to fit specific trading strategies. he platform also provides Smart Trading features, allowing users to set advanced trading rules with trailing stop-loss and take-profit functions. While this is a great platform for automated trading, a drawback of this platform is that its subscription costs could be expensive to certain users. In order to use some advanced features, it requires users to pay for subscription. 3Commas is ideal for traders seeking a versatile and powerful suite of tools for managing and automating their trades [1].

### 2.1.2 Pionex

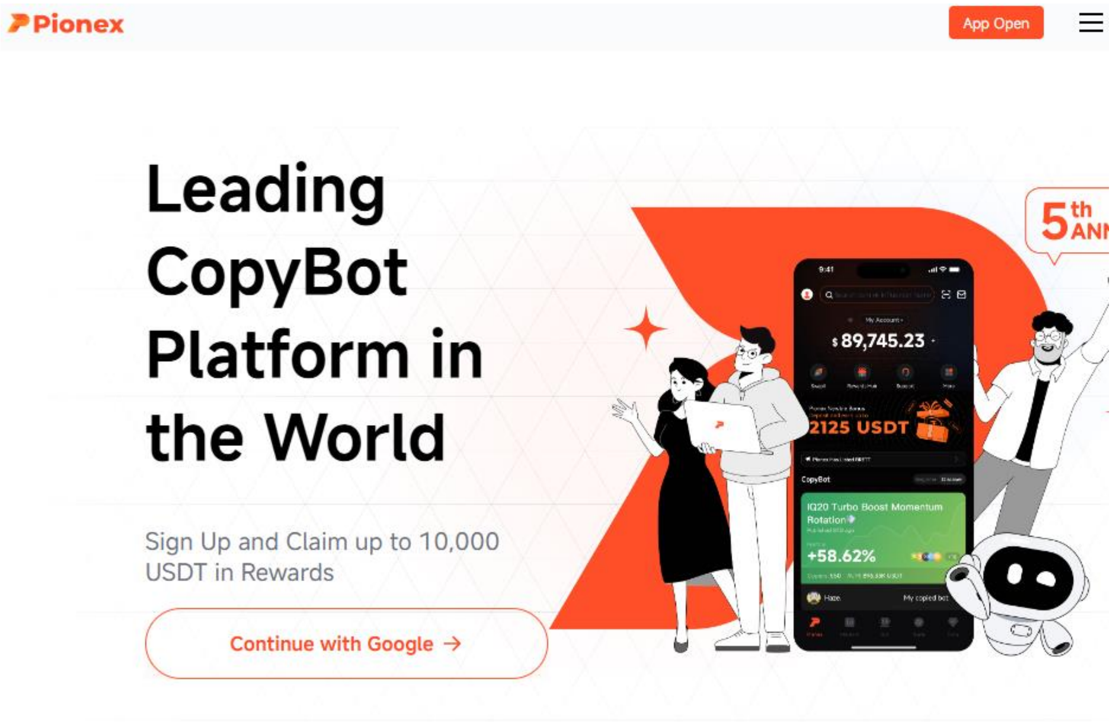



*Figure 2. 2 Web Application of Pionex*

Pionex is a cryptocurrency exchange which includes a built-in trade bot directly in its platform. This integration simplifies the trading process by providing ready-to-use bots like Grid and DCA Bots without additional fees for the bots themselves. With competitive trading fees and a user-friendly interface, Pionex is particularly attractive to beginners. However, its main limitations are the level of customization for its bots are reduced compared to other trade bot and it has a restriction to only trade within Pionex's own exchange. Pionex is best suited for those who prefer a straightforward, cost-effective trading solution with built-in automation [2].

### 2.1.3 TradeSanta

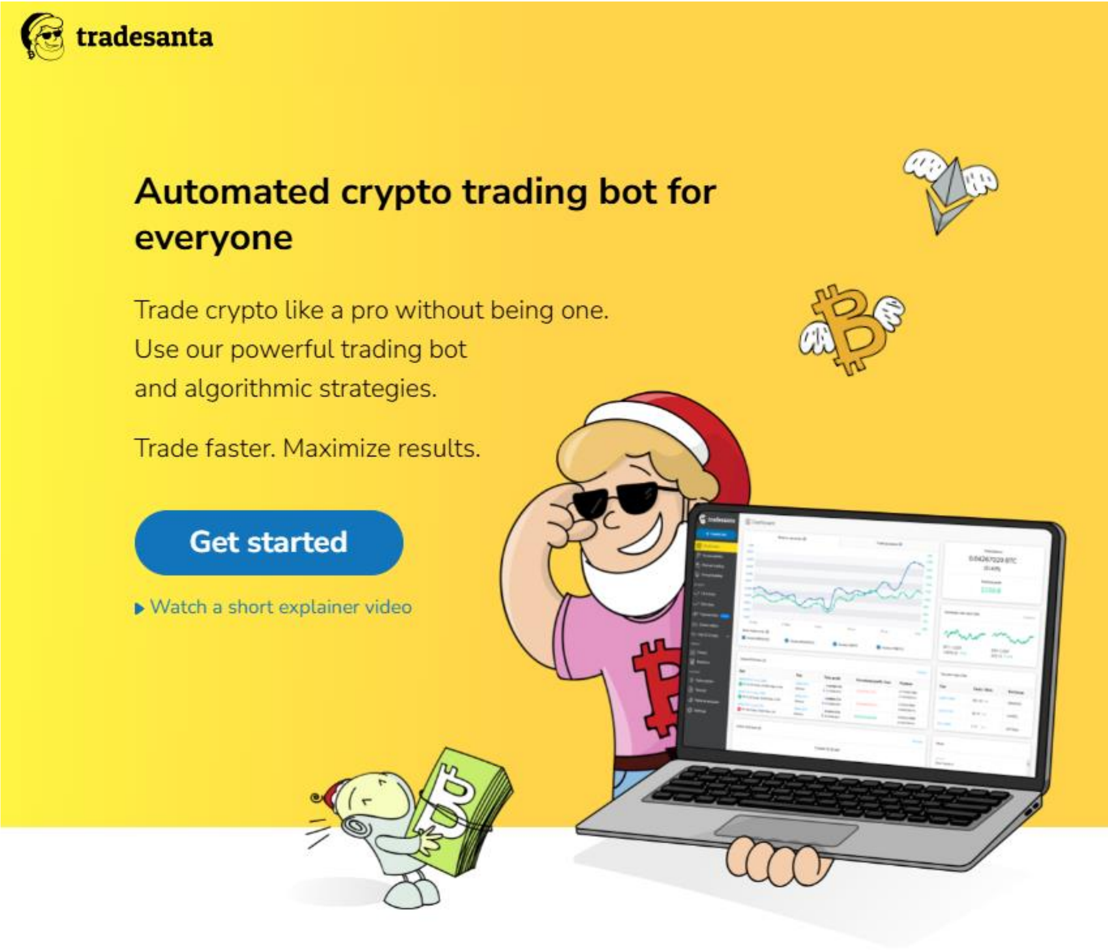



*Figure 2. 3 Web Application of TradeSanta*

TradeSanta offers a cloud-based platform for cryptocurrency trading. This website focuses on simplicity and ease of use. It provides various trading bots, including Long and Short bots, and allows for some degree of customization. The platform's cloud-based nature means users do not need to keep their devices online, and its user-friendly interface is accessible to traders of all levels. While TradeSanta's flexible subscription plans and multi-exchange support are advantages, its advanced features may be limited compared to other platforms. This makes TradeSanta a good choice for traders looking for a convenient and affordable solution with basic to intermediate automation capabilities [3].

## 2.2 Previous Works on Stock Market Prediction (using Artificial Intelligence)

### 2.2.1 Deep Learning

Prediction of stock market remains a difficult task due to its stochastic nature. Various models were proposed to predict stock price, with Long Short-Term Memory (LSTM) being one of them. One example can be seen in research done by Pushpendra et al. [4], where they created a stock price prediction model using LSTM algorithm. In their research, they have used the 10-year historical data for the NIFTY 50 index from the National Stock Exchange (NSE). The data is used for training the prediction model. They obtained promising results from their proposed model, with up to 83.88% accuracy.

In [5], Akshat et al. focused on comparing performance between conventional LSTM and attention-based LSTM deep neural network in predicting the stock market future price. [5] trained the prediction model on dataset including stock history, finance tweets sentiment, and technical indicators. From their studies, it is found that the finance tweets that are posted from market closure to market open in the next day has more predictive power on next day stock movement. The study showed that the aggregated dataset has better performance on the attention-based LSTM model.

In the research of [6], A. Bansal et al. proposed to predict the stock closing prices using Bi-directional LSTM (BLSTM) based Seq2Seq Model. The historical data of Google is used in training the model and Seq2Seq Model is used as the proposed model because it helps to map the input sequences to the output sequences of the data, where it makes use of the encoder-decoder structure. BLSTM links LSTM in bi-directional manner and information will be extracted from both layers concurrently. Their proposed method is then compared with other models, such as KNN, Linear Regression, and Decision Tree, the results showed that BLSTM has the highest accuracy (97.3%) among the other models. Another research also worked on LSTM and BLSTM, where both models were compared in assessing the accuracy of both models in predicting stock market [7]. It is inferred that proper adjustment of different parameters is crucial because the accuracy in prediction has significant dependency on the parameters. Therefore, LSTM and BLSTM require proper parameters tuning to achieve higher accuracy. In the study of [7] BLSTM model generates lower Root Mean Square Error (RMSE) than LSTM when using the same parameters for both models.

In [8], Xuan Ji et. al aimed to fully utilise social media data in building a stock price prediction model. In their research, they have proposed a prediction model that consists of traditional financial features and social media text features to predict stock prices. In this research, Doc2Vec and SAE are used for training social media documents and eliminate noisy data in the stock price time series that is caused by stock market fluctuation. Finally, LSTM is used for prediction of stock price. Results proved that the model proposed by [8] is able to perform better than other models in prediction stock prices.

Other than LSTM, Convolutional Neural Network (CNN) is also used as stock price forecasting method in research done by Sayavong Lounnapha et .al [9]. The researchers used Thai stock market as the dataset to train on the CNN model and three companies (BBL, CAPLL&PTT) listed on the SET50 index are evaluated and compared with the actual stock price. Due to the reason that CNN model does not depend on any previous information for prediction, CNN is shown to be very accurate in forecasting the stock prices.

### 2.2.2 Machine Learning

In a study done by Durga P et al. [10], comparisons between various approaches, such as Linear Regression, XGBoost Regressor, and random forest are done by the researchers. They claimed that it is challenging in predicting the stock market accurately due to the complexity of stock market data. Data cleaning is an important step, to accurately predict stock market. In the study, researchers obtained stock price data with a period five years from Yahoo Finance as the basis for dataset for training three models, Linear Regression, XGBoost Regressor, and random forest. After comparison is done among the three models, the results showed that all of the models were able to achieve accuracy score of 95% or more, with the highest score of 98.65% being achieved by the XGBoost Regressor. Researchers further fine tune the XGBoost Regressor Model to maximise its performance. Moreover, researchers also incorporated Chatbot into their study, in which the Chatbot is trained using natural language processing to engage in conversations with users and offer responses to their enquiries.

In paper [11], R.S.Latha et al. proposed to use K Nearest Neighbour (KNN) in developing a machine learning model to predict the change of stock value in the next day. The reason R.S.Latha et al. used the KNN algorithm in their study is because the dataset comprises of only numerical data and hence the data has higher possibility of grouping together. In this study, the output for the prediction would be the stock movement for the date, which is represented in high or low. From their results, KNN used in this project could only achieve maximum accuracy of 59% and they claimed that the results obtained is because the dataset only consists of stock prices.

Researchers in [12] proposed to employ Artificial Neural Networks (ANN) and Random Forest (RF) techniques to predict the closing prices of the stock market. In this study, they aim to predict the close price for the following day for five companies that are of different sectors. They have also introduced new variables as the input to their model for prediction, by leveraging the financial data that consists of Open, High, Low and Close prices (OHLC) of stock. Mehar Vijh et al. compared the performance between the two models, by assessing through key indicators such as RMSE, MAPE and MBE. Through the comparative analysis, it indicates that ANN consistently outperforms RF in forecasting stock prices. The results highlight the effectiveness of ANN, with the best values obtained showing an MBE of 0.013, RMSE of 0.42, and MAPE of 0.77.

### 2.2.3 Strengths and Weaknesses

| Method/ Technique | Strength | Weakness |
|---|---|---|
| LSTM | • It is designed to handle sequential data.<br>• It is able to capture long-term dependencies in sequential data.<br>• It can learn features from the raw input | • It requires large amount of data to train effectively because insufficient training data can lead to overfitting or poor generalization.<br>• It is sensitive towards hyperparameters, |

| | | |
|---|---|---|
| | data, enhancing prediction performance. | hence discovering optimal set of hyperparameters can be time-consuming. |
| CNN | • It is well-suited for processing spatial information | • It has limited temporal context, which means it could be difficult for CNN to model complex relationships in time series data. |
| Random Forest | • It can capture the feature that has most significant effect on stock price predictions. | • It can be computationally expensive and training with large datasets. |

*Table 2. 1 Strengths and Weaknesses for each method*

**2.3 Previous Works using Robotic Process Automation**

In [13], researchers included automation technique into the process of stock analysis and stock price prediction. In this study, Yusuf A et al. employed Robotic Process Automation (RPA) to facilitate stock market analysis. Object cloning, keystrokes, and Excel operations are utilised to extract and process financial data. It launches a webpage, making up to five attempts to retrieve information. The bot, run through Automation Anywhere Control Room, logs into Google Finance, fetching current market data, and updates an Excel sheet with stock performance metrics. This automation offers significant advantages for traders, allowing scheduled execution during market hours and automated notifications based on predefined conditions, such as stock price targets.

[14], Vanayak Jadkar et al. considered that in a real-life scenario, user relies on a technical screener to identify stocks with recent breakouts and this process usually takes up to 15 minutes. Manually performing this task involves constant monitoring and manual input of stock names into the analysis model. However, in this study,

researchers utilised RPA to eliminate the need for manual intervention entirely. They made use of automation platforms like Automation Anywhere, in which it provides a control room where users can schedule the execution of a specific bot at regular 15-minute intervals. This automation allows users to receive analysis reports without any human interaction, even for multiple stock screeners running concurrently. This automation not only simplifies the process of running multiple stock screeners but also accelerates the generation of analysis reports. By eliminating manual steps, it enables users to make more informed decisions promptly.

# CHAPTER 3 SYSTEM METHODOLOGY/APPROACH

The development methodology, system requirements, system design and project timeline will be presented in this section.

## 3.1 Development Methodology

In this project, the development approach that will be used is the Agile Methodology. Agile methodology is a repetitive framework that prioritises continuous improvement. This allows teams to deliver working software incrementally, minimise risk, and deliver high-quality solutions that align closely with business goals. The reason this approach is chosen for this project is because it can facilitate rapid development and ensures that there are minimal errors in the project by iterative testing. Hence, the quality of the project is assured.

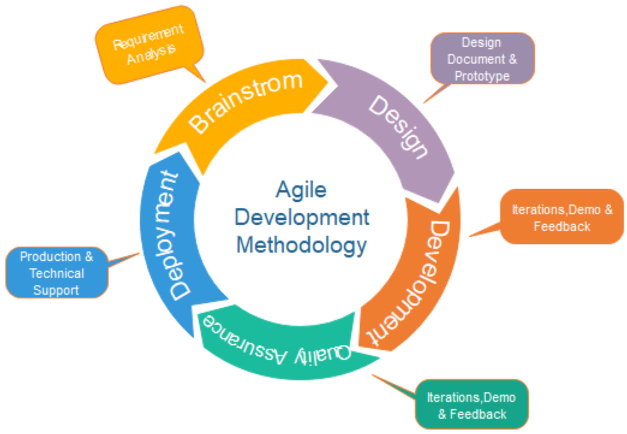



*Figure 3. 1 Agile Development Methodology*

**Planning**

This phase is initiated when discussion has been made with supervisor to decide on the title and scope of the project. Then, extensive research has been done on AI algorithms to predict market price and use case of RPA. Different AI algorithms has been reviewed and compared based on past research. Apart from that, some time is also allocated for brainstorming on the functionalities this project would bring to users, in alignment with the project title. Moreover, the timeline of the project is drafted to ensure that key milestones can be achieved on time. Other than that, project budget and resources are also taken in account when doing project planning.

**Design**

After having a rough idea for the requirement of the project, the overall structure of the system is defined by drafting system architecture diagram, use-case diagram, and sequence diagram. By drawing diagrams, it helps understand clearly on how the system will work and ensure that the logic and structure of the system is correct. During this phase, the system architecture design is drawn to visualise how the different component interacts with each other. In this phase, requirement is also defined. The functional requirement for this website is stock prediction, sentiment analysis, trading automation. Whereas, the non-functional requirements are performance, and usability. By drawing the use case diagrams, it visualises the interaction between actor and the functionalities. The use case description further specifies the flow of each use case.

**Development**

The LSTM prediction model was the first thing to develop in FYP1. Firstly, the coding was done on Jupyter Notebook to run and test the code. The code for building the model includes, data sourcing, data preprocessing, feature engineering, model training, model evaluation, prediction, and visualization. These steps are carried out repeatedly until the prediction model can obtain high accuracy for its prediction. Moreover, the basic code for development of web applications is done initially by connecting the frontend and the backend using API. After the connection has been established, the prediction model is integrated into the backend of the application and functions have been added into the codebase. Then, a simple user interface is coded for the web application. RPA sequences is also developed in UiPath but have not been integrated into the system.

During FYP2, the prediction model is further improved and fine-tuned to prevent overfitting of the model. Then, other main features such as the sentiment analysis, paper trading and trade bot are built subsequently.

**Testing**

In this phase, testing has been done iteratively to ensure that the web application works smoothly. Initially, when prediction model was built, testing and fine-tuning is carried out repeatedly to ensure that the model is able to predict accurately. Then, after developing the website, the functions that are integrated into the application are also checked, to ensure they work well in the web application. To test the system, unit testing, integration testing is done to ensure that all the important units on the website are working as expected. Test cases are written to test if the unit achieve the expected results, if no, debugging will need to be carried out to fix the issue. Other than that, the performance of the website will also be tested to ensure to bring a smooth browsing experience to users

## 3.2 System Requirement

### 3.2.1 Hardware

The hardware used for carrying out this project is computer. A computer is needed for building LSTM model, RPA sequence, and the web application.

| Description | Specifications |
|---|---|
| Model | MateBook D 15 |
| Processor | Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz |
| Operating System | Windows 11 Home Single Language |
| Graphic | Intel(R) UHD Graphics |
| Memory | 16 GB DDR4 |
| Storage | 512 GB SSD |

*Table 3. 1 Specifications of Computer*

### 3.2.2 Development Tools

*Figure 3. 2 Logo of VS Code*

The development tool that is used in developing the web application is Visual Studio Code. It serves as the integrated development environments (IDE) for the development of the LSTM prediction model as well as the frontend and backend of the web application. This IDE is lightweight and fast because it consumes fewer system resources, making it a good choice for developing this project. Moreover, it contains wide variety of extensions that can help enhance the code and make development process faster. The programming languages used in this IDE is Python and React.js, in which Visual Code provides and conducive environment for developing Flask backend APIs and React frontend applications.

## 3.3 System Design

### 3.3.1 System Architecture Diagram

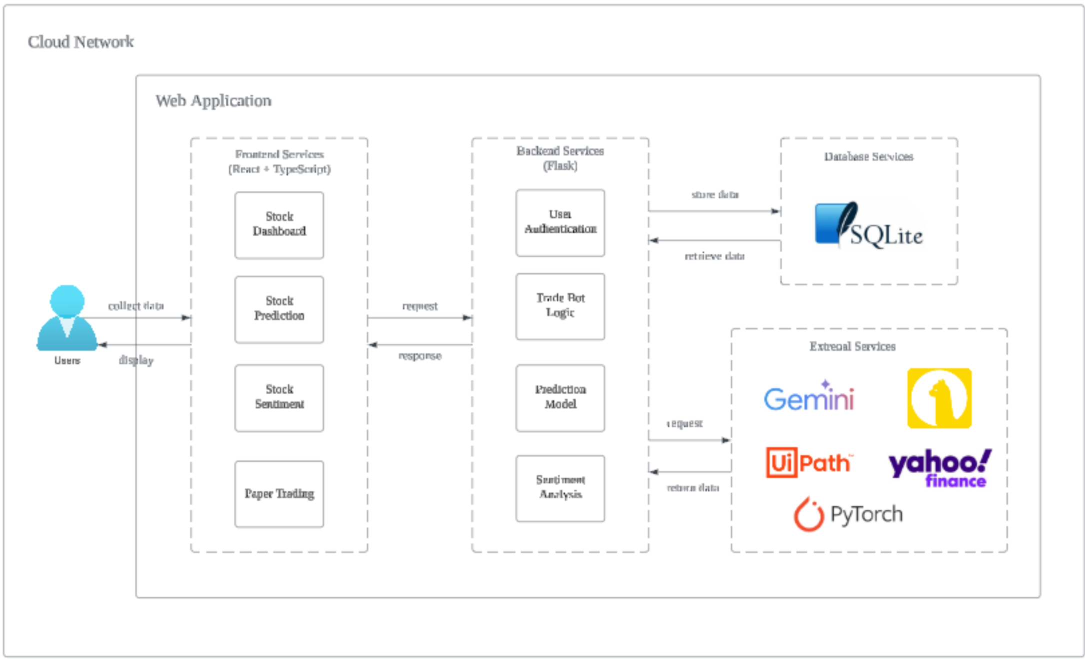



*Figure 3. 3 System Architectural Diagram*

Figure above illustrates the architectural design for the web application. First, users will interact with the client-side of the web application, where users will provide data to perform a certain function on the web application. The frontend services are comprising of stock dashboard, stock prediction, stock sentiment and paper trading. Then, the client will interact with the API Gateway Endpoint, which in simple word, is the web server. When user requests for a stock prediction, the web application processes the request, potentially using machine learning models built with PyTorch. On the other hand, when user requests for sentiment analysis, it will request the backend services and the backend services will utilise UiPath for data collection and VADER to do sentiment analysis. Whenever users intend to perform a function on the application, the client will send request to the server and wait for server to response. Upon receiving request to perform any functions on the web application, the server will then fetch or call the function to perform the operation and hence pass the results back to the client, in which client will pass display the results to users. When calling the functions, the server will also call the external services to aid in perform the functions. Moreover, user data will be store in the SQLite3 database.

### 3.3.2 Use Case Diagram

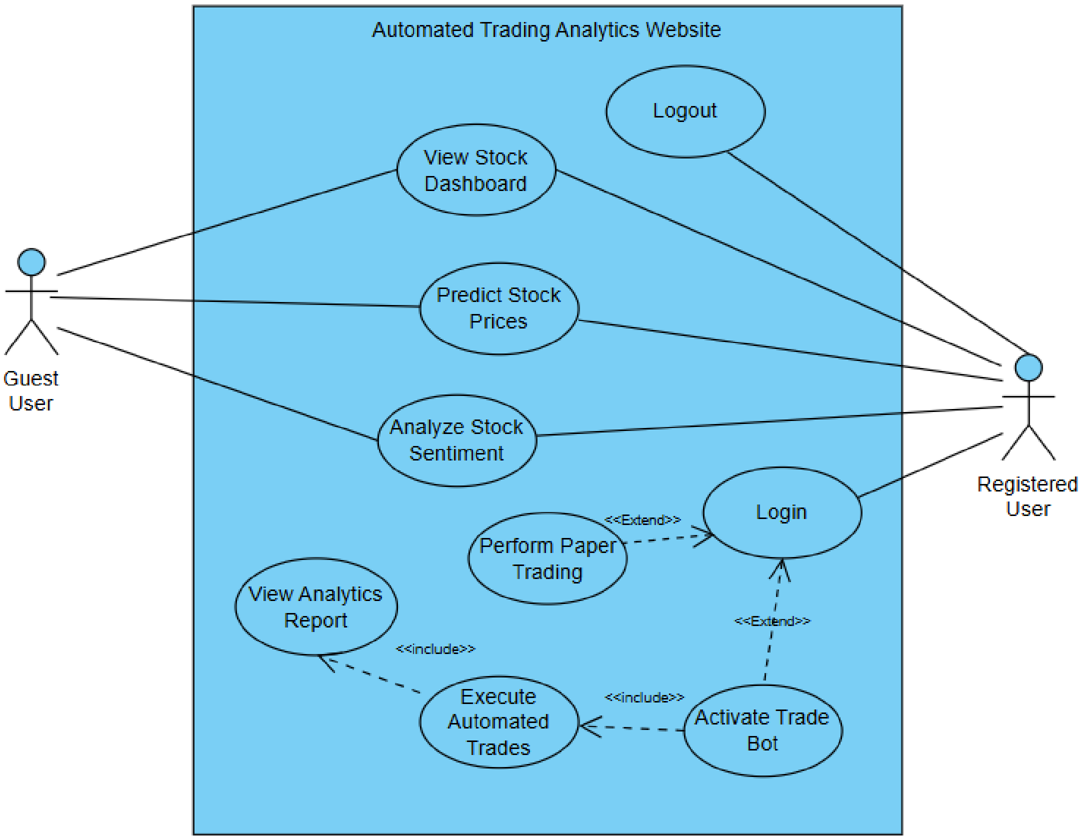



*Figure 3. 4 Use Case Diagram*

Based on the figure above, there are various use cases in the web application, which are view ticker information, receive prediction results, and get action recommendations. For both view ticker information and receive prediction results use cases require user to fill in the ticker symbol intended to view or predict. After searching for the ticker in view ticker information, the information such as the price, historical data and overview of the searched ticker will be displayed to users. Moreover, users are required to fill in the necessary information such as ticker symbol and days to predict for prediction, and hence users will then receive the prediction results with the forecasted future price. Users can also get the action recommendation for a ticker once the ticker has been predicted.

### 3.3.3 Use Case Description

**View Stock Dashboard**

<table>
<tr><td>Use Case: View Stock Dashboard</td><td>ID: 1</td><td>Importance Level: High</td></tr>
<tr><td>Primary Actor: Guest user, Registered user</td><td colspan="2">Use Case Type: Detail, Essential</td></tr>
<tr><td colspan="3">Stakeholders and Interests:<br>Guest User, Registered User - wants to view general stock market data and insights</td></tr>
<tr><td colspan="3">Brief Description: This use case allows users to search and view stock market data, including stock prices, historical trends and stock information.</td></tr>
<tr><td colspan="3">Trigger: The user navigates to the stock dashboard page via the sidebar menu.<br><br>Type: External</td></tr>
<tr><td colspan="3">Relationships:<br>Association: Guest user, Registered user<br>Include: -<br>Extend: -<br>Generalization: -</td></tr>
<tr><td colspan="3">Normal Flow of Events:<br>1. The user navigates to the stock dashboard page from the side navigation menu.<br>2. The system displays a search bar for ticker input and the latest stock data of ‘AAPL’ by default.<br>3. The user searches for the desired ticker to view.<br>4. The system validates ticker and displays the stock data, including prices, trends, and financial news.</td></tr>
</table>

| |
|---|
| **Subflows: -** |
| **Alternate/ Exception Flows:**<br>4a: The user searches a stock that is not available or invalid.<br>4b: The system displays nothing on the dashboard.<br>4c: The user fills in valid ticker symbol in the search bar. |

*Table 3. 2 Use Case Description for View Stock Dashboard*

**Predict Stock Prices**

| | | |
|---|---|---|
| **Use Case:** Predict Stock Prices | **ID:** 2 | **Importance Level:** High |
| **Primary Actor:** Guest user, Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Guest User, Registered User - wants to predict future stock prices based on historical data and trends to inform their investment decisions. | | |
| **Brief Description:** This use case allows a user to predict future stock prices using the platform's machine learning model (LSTM). The prediction feature helps users make informed trading decisions by providing insights into potential price movements for selected stocks. | | |
| **Trigger:** The user navigates to the stock prediction page via the sidebar menu.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Guest user, Registered user<br>**Include: -**<br>**Extend: -**<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user selects a ticker for which they want a prediction and chooses the number of future days to predict.<br>2. The user clicks the “Predict Stock” button.<br>3. The system retrieves historical data for the selected stock.<br>4. The LSTM prediction model is applied to analyse historical trends and forecast the next day’s stock price. | | |

| |
|---|
| 5. The system displays the predicted stock price to the user, along with recommendation to user. |
| **Subflows: -** |
| **Alternate/ Exception Flows:**<br>3a: The system experiences an internal server error.<br>3b: The user retries the prediction later. |

*Table 3. 3 Use Case Description for Predict Stock Prices*

**Analyse Stock Sentiment**

| | | |
|---|---|---|
| **Use Case:** Analyse Stock Sentiment | **ID:** 3 | **Importance Level:** High |
| **Primary Actor:** Guest user, Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Guest User, Registered User - Interested in analysing the sentiment of stock-related news to inform their decisions about stock investments. | | |
| **Brief Description:** This use case allows users to analyse news sentiment of stocks using the VADER sentiment analysis model. This result help users understand the recent market sentiment for specific ticker. | | |
| **Trigger:** The user navigates to the sentiment analysis page via the sidebar menu.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Guest user, Registered user<br>**Include: -**<br>**Extend: -**<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user selects a ticker to analyse.<br>2. The system retrieves relevant news data related to the selected ticker.<br>3. VADER analyses the sentiment of the news articles.<br>4. The system displays the sentiment analysis results (positive, neutral, negative) in a graphical or tabular format. | | |

| |
|---|
| **Subflows: -** |
| **Alternate/ Exception Flows:**<br>2a: No news is retrieved.<br>2b: The system displays empty table as a result. |

*Table 3. 4 Use Case Description for Analyse Stock Sentiment*

**Login**

| | | |
|---|---|---|
| **Use Case:** Login | **ID:** 4 | **Importance Level:** High |
| **Primary Actor:** Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Registered user - Interested in securely accessing their account to execute paper trading through the Alpaca trading platform. | | |
| **Brief Description:** This use case allows user to log into their Alpaca trading account, authenticate themselves and authorise the system to execute the paper trades. | | |
| **Trigger:** The user clicks the "Login" button on the application.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Registered user<br>**Include: -**<br>**Extend:** Perform Paper Trade<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user clicks the "Login" button.<br>2. The system displays a login form that requests for user's login credentials.<br>3. The user enters their username and password.<br>4. The system validates the credentials.<br>5. The system redirects user to the Alpaca authorization page.<br>6. The user authorises the application to access their Alpaca account for paper trading | | |

| |
|---|
| 7. The system confirms the authorization and redirects user to the Paper Trading page. |
| **Subflows: -** |
| **Alternate/ Exception Flows:**<br>2a: Invalid login credentials.<br>2b: Systems displays error message and prompt the user to try again. |

*Table 3. 5 Use Case Description for Login*

**Perform Paper Trade**

| | | |
|---|---|---|
| **Use Case:** Perform Paper Trade | **ID:** 5 | **Importance Level:** High |
| **Primary Actor:** Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Registered user - Wants to simulate stock trading without real financial risks to practice trading strategies. | | |
| **Brief Description:** This use case allows users to place buy/sell order in a simulated environment without real money involved. | | |
| **Trigger:** The user navigates to the Paper Trade section.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Registered user<br>**Include:** Login<br>**Extend: -**<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user logs in and navigates to the "Paper Trading" section.<br>2. The system displays user account information, order and position information<br>3. The user decides to either buy or sell the stock in the simulated environment.<br>4. The system processes the trade and updates the user's paper trading account with the relevant information. | | |

| |
|---|
| 5. The system displays the user's updated order and position information. |
| **Subflows: -** |
| **Alternate/ Exception Flows:**<br>2a: The user enters invalid ticker to buy.<br>2b: The system displays an error message.<br><br>3a: The user buys a stock with insufficient funds.<br>3b: The system displays an error message. |

*Table 3. 6 Use Case Description for Perform Paper Trade*

**Activate Trade Bot**

| | | |
|---|---|---|
| **Use Case:** Activate Trade Bot | **ID:** 6 | **Importance Level:** High |
| **Primary Actor:** Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Registered User - Wants to automate stock trading based on predictions and sentiment analysis, which removes the need for manual execution. | | |
| **Brief Description:** This use case allows registered users to activate a trade bot that automatically executes trades in user’s Alpaca paper trading account based on the system's stock price predictions and sentiment analysis results. | | |
| **Trigger:** The user presses the "Activate Trade Bot" button.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Registered user<br>**Include:** Login<br>**Extend:** -<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user clicks the "Activate Trade Bot" button.<br>2. The system enables the trade bot and starts monitoring stock predictions and sentiment data. | | |

| |
|---|
| 3. The bot makes decisions to buy/sell stocks in the user's Alpaca account based on the LSTM predictions and sentiment analysis.<br>4. The bot automatically places trades according to the predefined strategy.<br>5. The system logs the trades and generates a summary report for the user. |
| **Subflows: -** |
| **Alternate/ Exception Flows: -** |

*Table 3. 7 Use Case Description for Activate Trade Bot*

**Logout**

| | | |
|---|---|---|
| **Use Case:** Logout | **ID:** 7 | **Importance Level:** High |
| **Primary Actor:** Registered user | **Use Case Type:** Detail, Essential | |
| **Stakeholders and Interests:**<br>Registered User - Wants to securely log out of their Alpaca account. | | |
| **Brief Description:** This use case allows registered users to log out from their account securely. | | |
| **Trigger:** User selects the "Logout" button in the navigation menu.<br><br>**Type:** External | | |
| **Relationships:**<br>**Association:** Registered user<br>**Include: -**<br>**Extend:** -<br>**Generalization: -** | | |
| **Normal Flow of Events:**<br>1. The user clicks the "Logout" button.<br>2. The system requests the server to terminate the user's session.<br>3. The system clears any session-related data and redirects the user to the public view of the website. | | |
| **Subflows: -** | | |
| **Alternate/ Exception Flows: -** | | |

*Table 3. 8 Use Case Description for Logout*

### 3.3.4 Activity Diagram

#### View Stock Dashboard

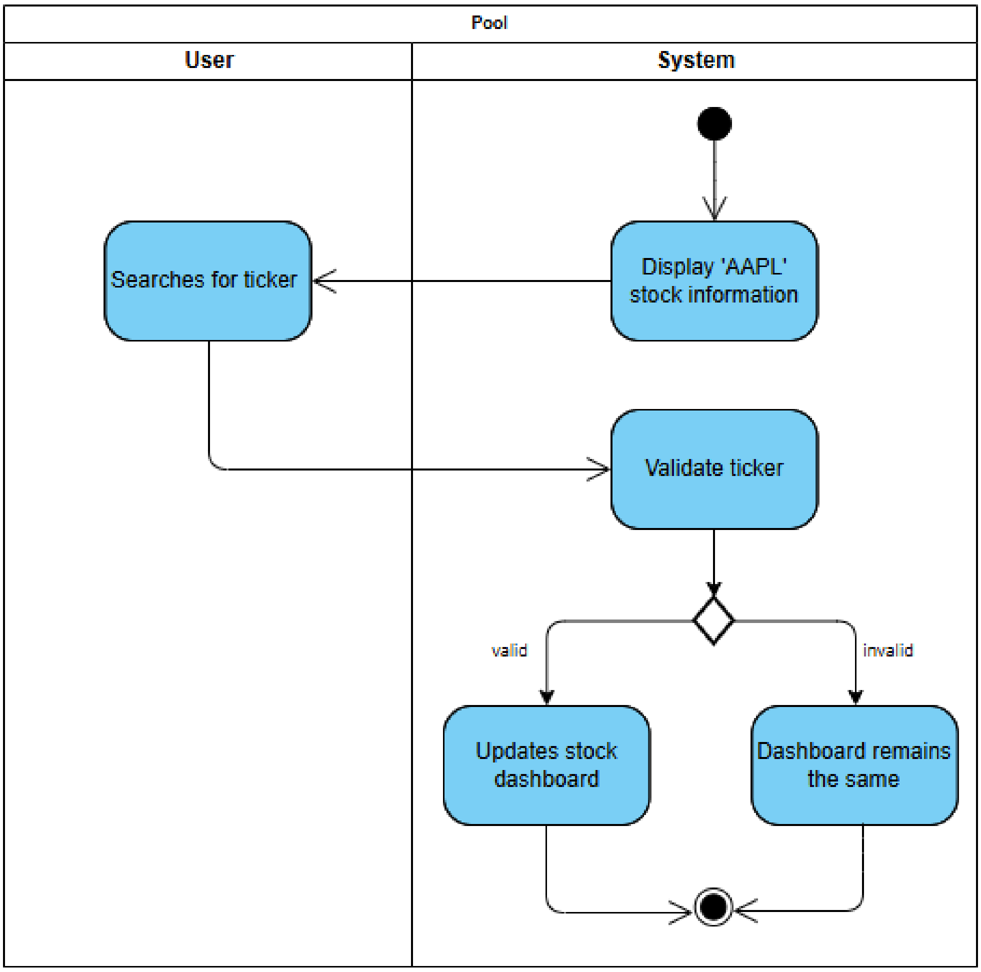



*Figure 3. 5 Activity Diagram of View Stock Dashboard*

**Predict Stock Prices**

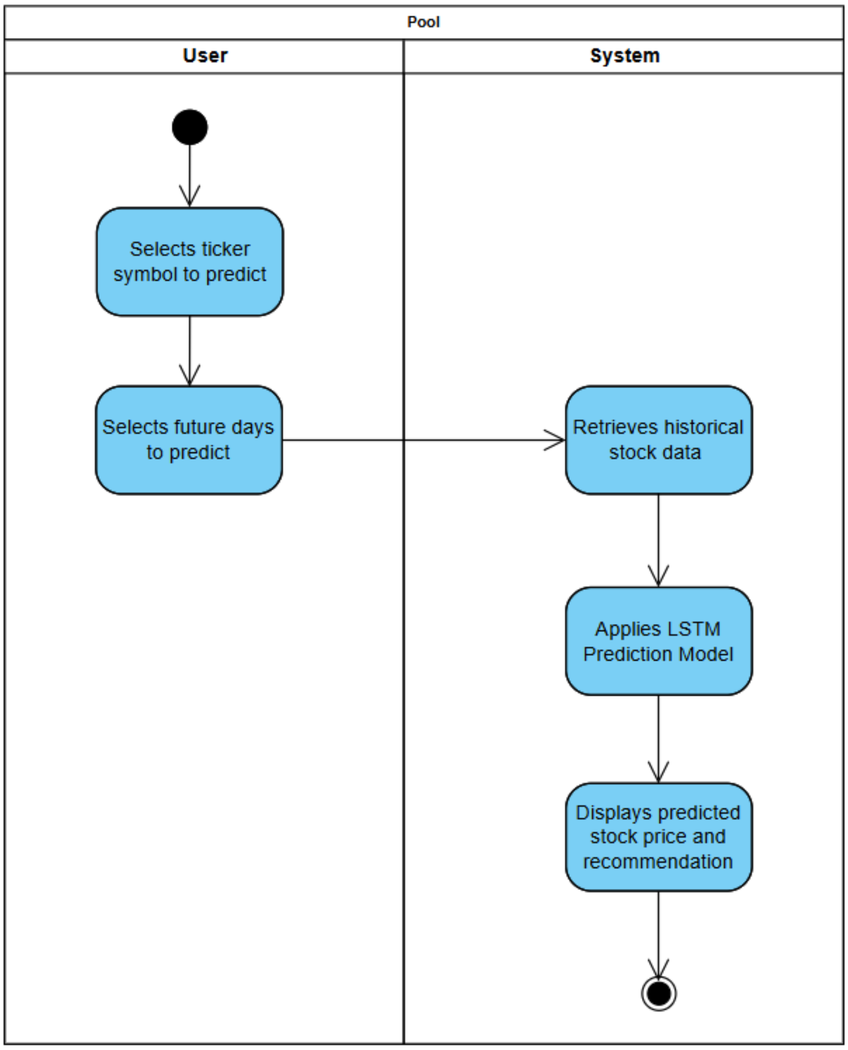



*Figure 3. 6 Activity Diagram of Predict Stock*

### Analyse Stock Sentiment

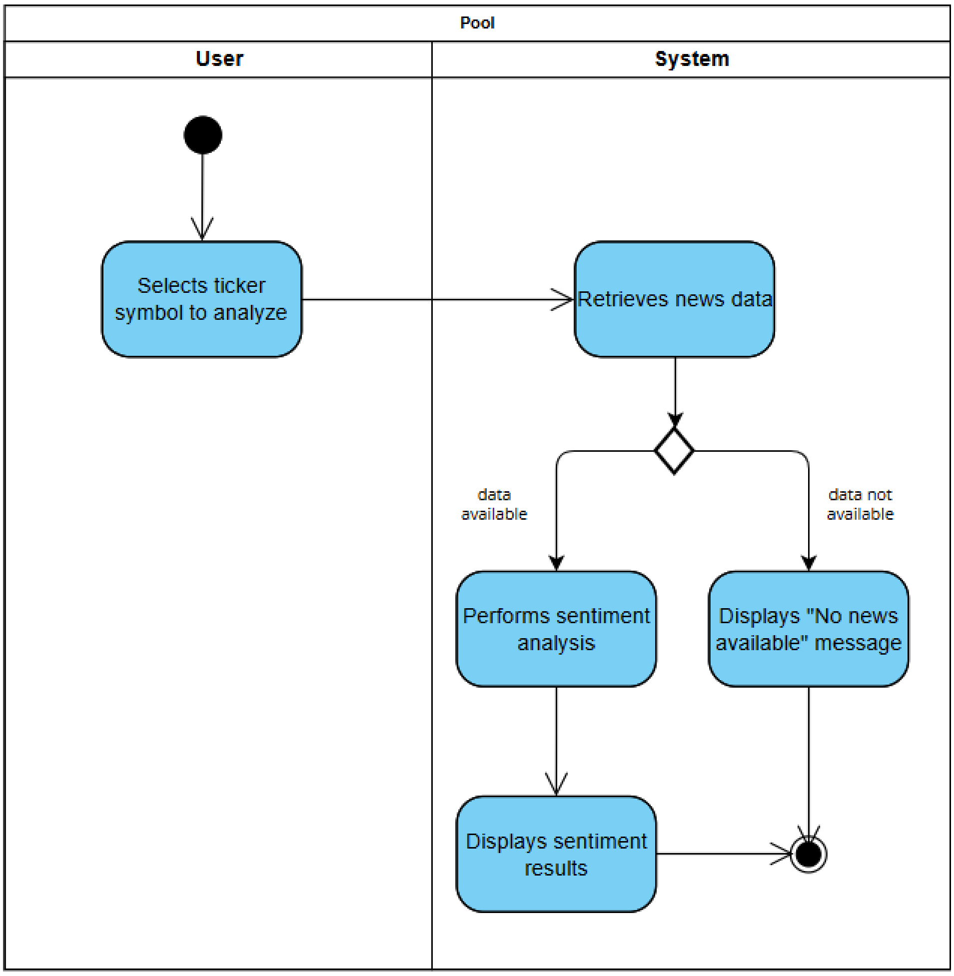



*Figure 3. 7 Activity Diagram of Analyse Stock Sentiment*

## Login

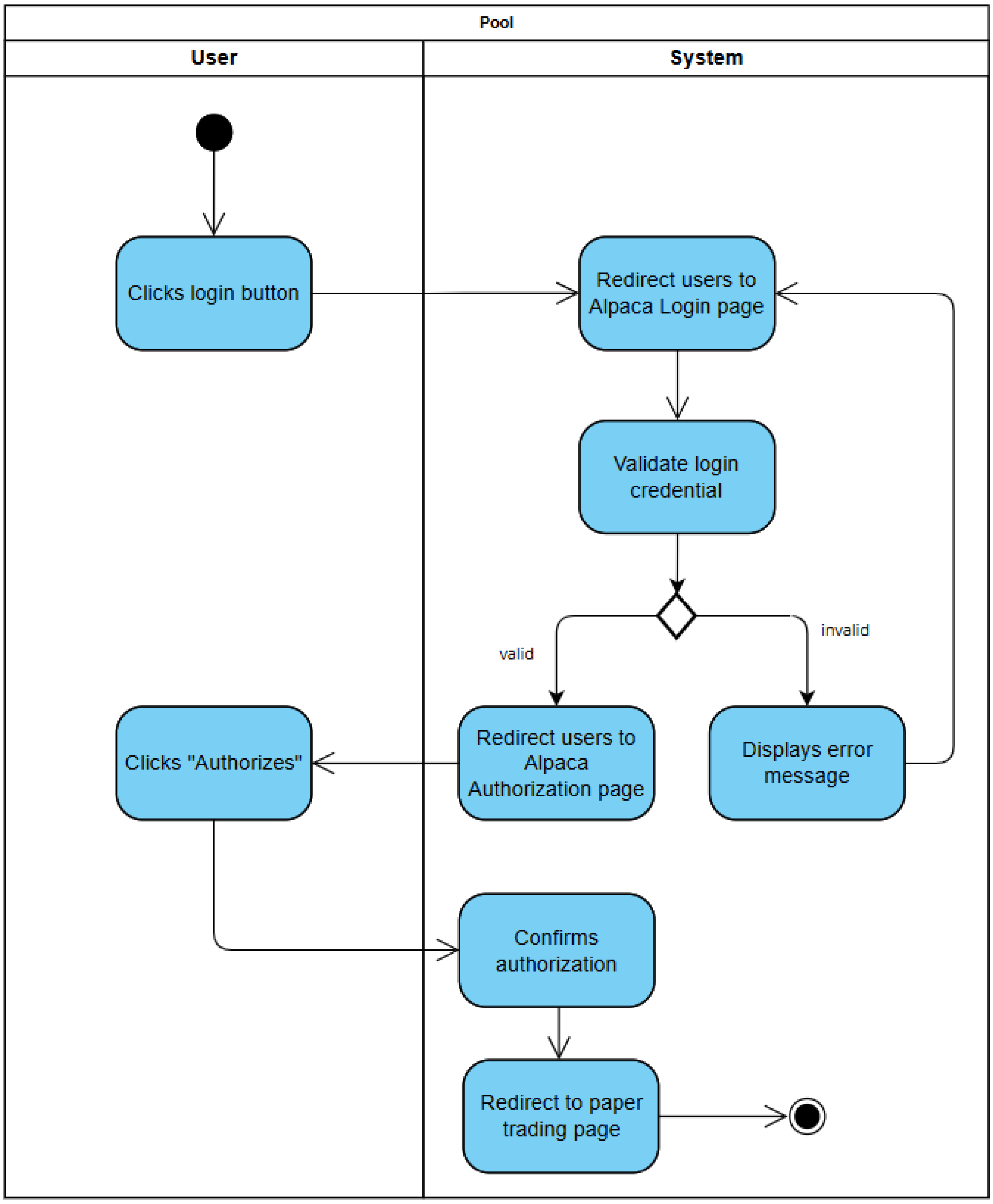



*Figure 3. 8 Activity Diagram of Login*

**Perform Paper Trade**

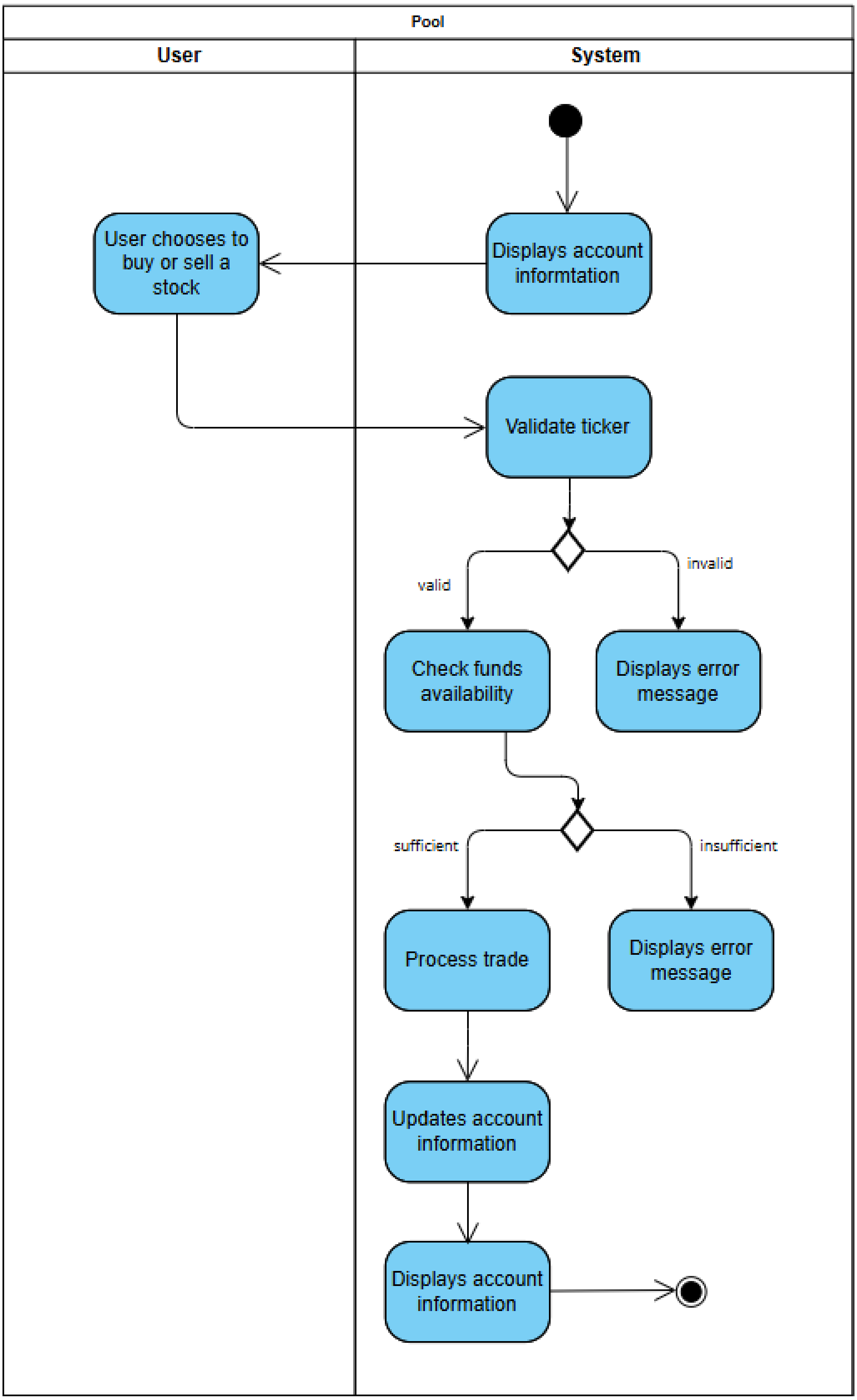



*Figure 3. 9 Activity Diagram of Perform Paper Trade*

**Activate Trade Bot**

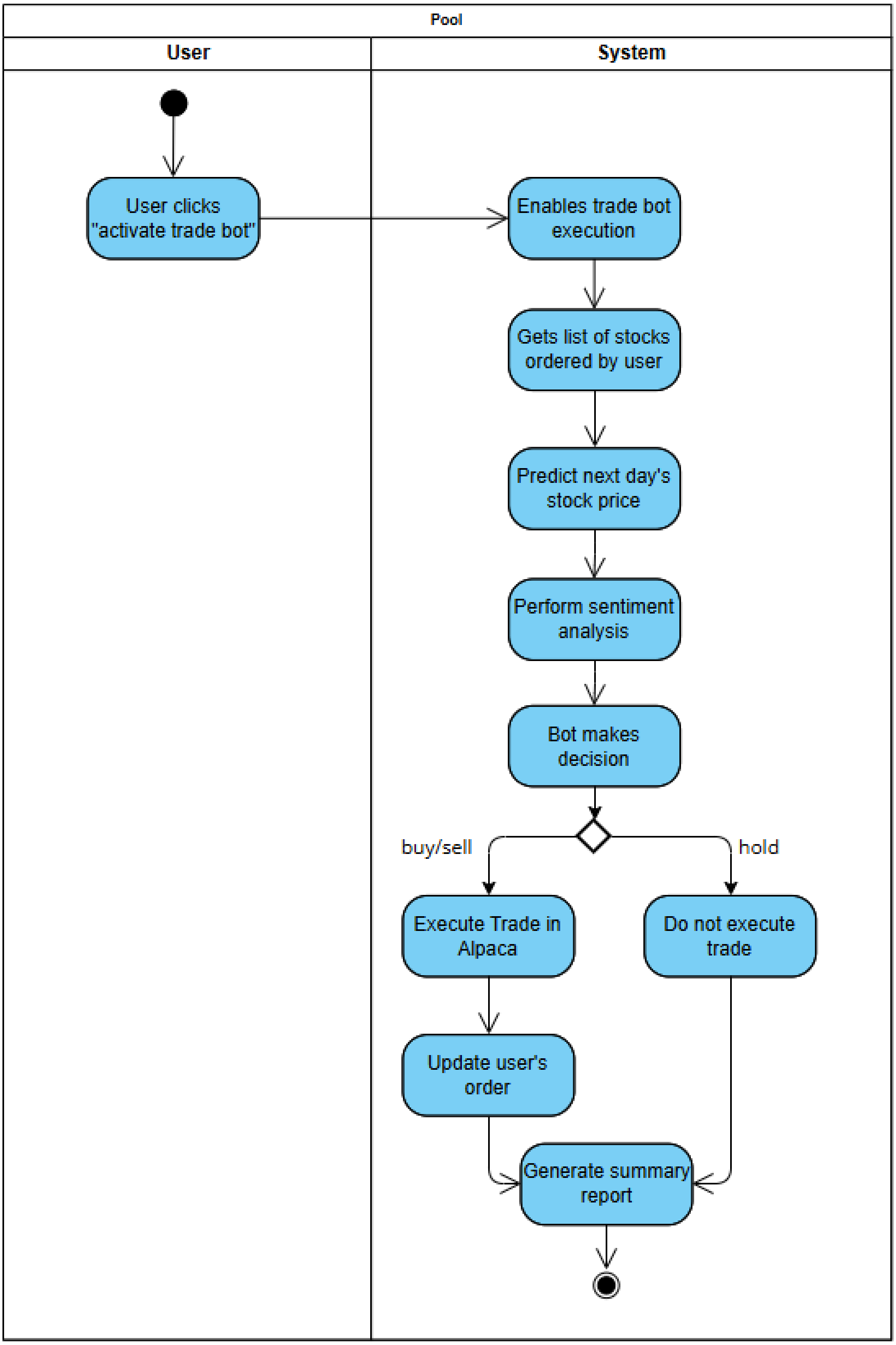



*Figure 3. 10 Activity Diagram of Activate Trade Bot*

**Logout**

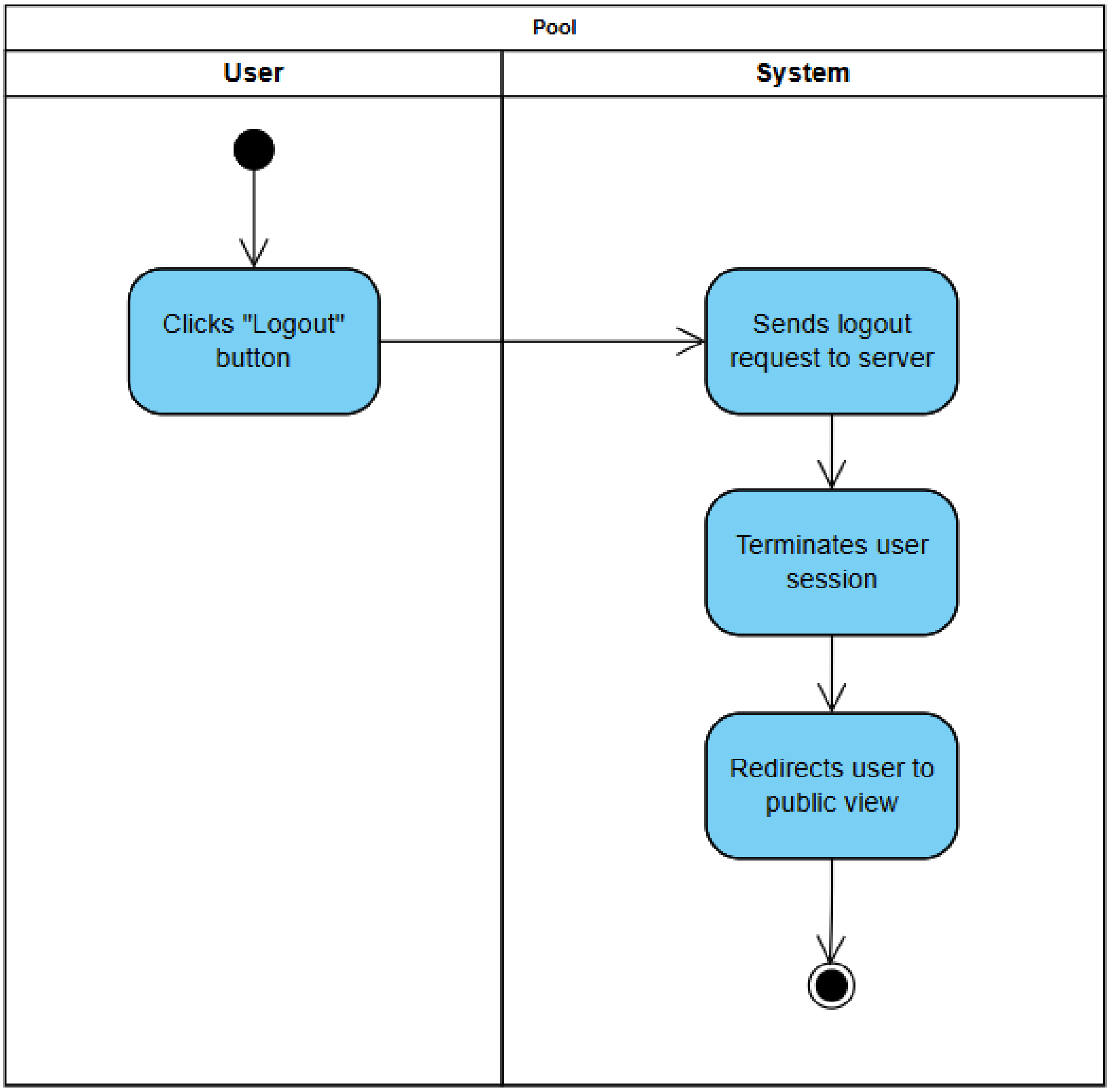



*Figure 3. 11 Activity Diagram of Logout*

## 3.4 Project Timeline

### FYP 1

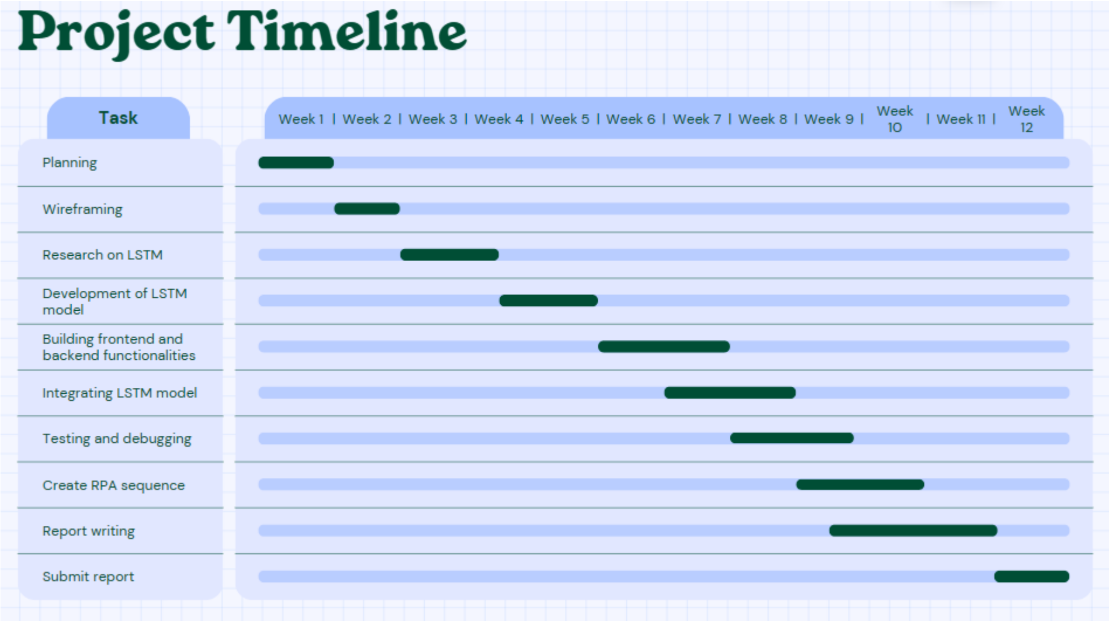



*Figure 3. 12 Project Timeline for FYP1*

### FYP 2

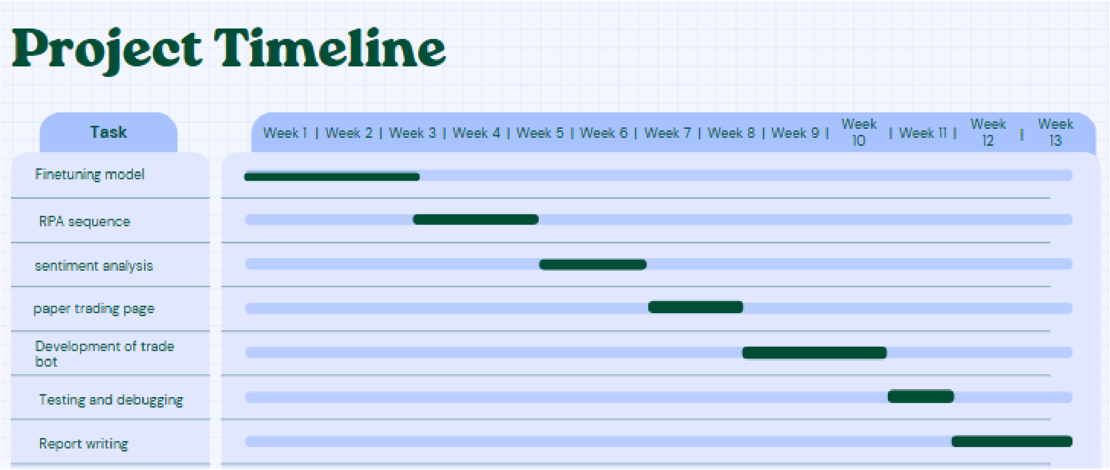



*Figure 3. 13 Project Timeline for FYP2*

# CHAPTER 4 SYSTEM DESIGN

## 4.1 Setting up

### 4.1.1 Software

Various software are required to be installed in the laptop to assist in the development of this project:

1. Visual Studio Code: 1.88.0
2. UiPath Studio: Community Edition

## 4.2 UI Development

In this part, the development of user interface (UI) for the automated trading analytics website will be discussed comprehensively. The goal of the user interface design is to provide user an interactive and user-friendly platform for trading analytics purposes. Hence, the user interface design is designed to be simple and minimal as the project prioritises on the functionality over the user interface.

### 4.2.1 Framework and Tools

The user interface of the website is built using React, an open-source JavaScript library. React is used in this project development is because the library is component-based, in which it is designed in a way that allows user to build web applications using components. This means that the components can be reused across different parts of the web application, and this promotes code reusability. For example, if user needs to customise a button and include it on different web pages, user need not rewrite the code in every page but build the button as a component and simply integrate this reusable component in the web pages. This indicates that the component-based library helps save time and resources. Moreover, Tailwind CSS is used to streamline the styling of the website. By using Tailwind CSS, it eliminates the need for writing custom CSS for every component, which allows user to build web applications faster. This is because it provides a variety of utility classes for user to build custom web design. Apart from that, Yarn is used in this project to manage dependencies, in which libraries can be easily installed, updated, and managed for frontend development. By using Yarn, the conflicts between installed libraries can be prevented because Yarn helps to lock package versions. For a smoother and more efficient development experience, Vite is used in this project development. This is because Vite provides a faster development build by utilizing ES modules in the browser, which means that Vite will only processes and serves the current code, rather than the entire project. For example, if user is editing one file or component, Vite will only processes that specific file or component to browser, therefore leading to faster build time. Vite also provides instant feedback to the changes made to the code in browser, without reloading the page. This further enhances the efficiency of developing the frontend. In developing the user interface of this project, frameworks or tools that are chosen shares the same characteristics, which can provide a development environment that streamline the building process by enhancing the efficiency and speed.

### 4.2.2 Main Components

There are several main components in this website, which includes View Stock Dashboard, Stock Prediction, Sentiment Analysis, Paper Trade. These are the web pages that are available on the website. Each page represents different functions of the project, providing users with different tools and features to interact with, which enhance the overall trading and analytics experience.

Firstly, the initial page that user will see is the stock dashboard page. This page displays the stock information, such as stock overview, current stock price, and the stock trend. The page is separated into different sections and each section is encapsulated in a grid. The grid is then arranged in a way that is intuitive and user-friendly. The sections included in this page are namely, the search bar, stock overview, stock chart, and stock information. The dashboard does not include advanced charts or information because this simplicity is believed to be able to make beginner traders to navigate and understand trades more easily, which aligns with the aim of this project. In this page, user can search for the ticker symbol that they would like to view. The searching action will prompt a list of selections based on user's input in the search bar. Upon selecting the symbol, the frontend will then pass the request to backend of the web application to retrieve stock information form Yahoo Finance API and the dashboard will then be updated with the stock information retrieved.

Furthermore, the stock prediction page and sentiment analysis page are designed in this project to allow user to perform analysis on stock. These pages allow users to interact with the stock prediction model and sentiment analysis tool. These tools are located in the backend of this website, which will be explained in detailed in the following section. The purpose of these pages is to let user view the analysis results done by the prediction model and sentiment analysis tool, which can aid them in making decisions when doing trading manually. In these respective pages, the user interface is separated into two sections, which are the input section and the results section. Users are required to input the ticker symbol for both pages in order to perform the analysis. Upon submitting the ticker symbol, the stock prediction model or the sentiment tool will be executed.

Next, one of the main components in this website is the paper trade page. This page is s significant page for this website. The page allows users to place buy and sell order within the simulation environment while enabling trade bot to showcase its performance on this page. Instead of live trading, paper trading is preferred to be used in this project as it aligned better with the project's objective. It provides a risk-free environment where the trade bot can perform its tasks, and beginner traders can practice hands-on trading without the risk of losing real money. This page is associated with Alpaca trade account, and hence user must be registered to an Alpaca account before being able to perform paper trading on this website. Therefore, users must log in to their Alpaca account and authorise their account to this website before performing paper trading. Once authorization is done, users are now able to perform trade operation in a simulated environment in this web page. Any actions done in the paper trading page will be updated instantly in user's Alpaca account. All the trading operations are implemented using Alpaca API. The backend of the website will request the Alpaca endpoint and hence any operations done in this web page will be updated simultaneously in user's Alpaca trade account. Other than that, this page also includes another section for activating the trade bot. There will be a toggle switch that allows user to activate or deactivate the trade bot. Upon activation of the trade bot, it triggers the trade bot on the backend on the website to execute automated trading. Each time the bot executes, it will generate and display a summary report to user, to ensure that users understand the decision made by the bot.

Moreover, another main component in this website is the side navigation menu. This is crucial for the entire website because it provides the navigation to different pages. This component utilised React’s useState to manage state between the pages. The default page that user will view is the stock dashboard page. When user clicks another page in the side navigation menu, the state will update the current page to the page that is clicked by user, and hence rendering the selected page to user.

### 4.3 Stock Prediction Development

The stock prediction feature is one of the core components in this project, in which it is designed to forecast future stock prices based on historical stock data. The primary goal of this prediction is to enable users, particularly beginner traders, to make more informed decisions by leveraging the deep learning model for stock price predictions. This functionality helps bridge the gap between data analysis and practical trading decisions, allowing users to visualise potential market movements without relying solely on intuition. Various libraries are necessary for the implementation of the LSTM prediction model. The libraries needed are:

- Numpy
- Pandas
- Torch
- Sklearn
- Matplotlib
- Yfinance

#### 4.3.1 Model Selection and Architecture

For this project, LSTM (Long Short-Term Memory) was chosen to perform stock prediction. This is because of its ability to affectively model sequential data and capture long-term dependencies in time-series forecasting.

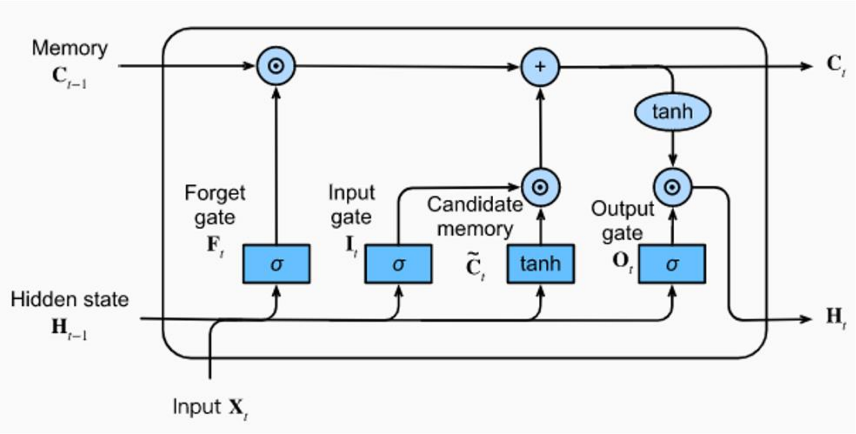



*Figure 4. 1 LSTM Architecture*

Figure 4.1 shows the LSTM layer, which is one of the components in the LSTM prediction model implemented in this project. The architecture of the prediction model consists of input data, LSTM layer, Fully Connected Layer, and then the output. The first component in the LSTM model is the input data, where it consists of lag features representing the past values of the target variable and the lag features corresponds to a time step in the sequence. Next is the LSTM layer, where it serves as the core component in the prediction model because it is responsible for processing sequential input data. There are several LSTM cells found in the prediction model, in which each of them processes the input data sequentially. The three main components found in LSTM cell are forget gate, input gate and output gate. The determination of information that should be dropped from the previous hidden state is done by the forget gate, while the input gate finds out the new information from the current input that should be stored in the cell state. Moreover, the output gate plays its role in deciding the information that should be output as the hidden state. The hidden state in LSTM means an internal representation of the network's memory, whereby it will be updated at each time step as the cell processes input data. The following layer is the fully connected layer. This layer maps the output of the LSTM to the final prediction. This layer is important because it will combine the features extracted by the LSTM layer to make accurate predictions. Finally, the output of the model will be obtained by applying activation functions to the output of the fully connected layer and the output generated will be the predicted results.

### 4.3.2 Data Preprocessing

Firstly, before training the LSTM model, data acquisition is performed. The yfinance library provides access to daily stock prices, which includes the open, close, high, low and the volume. Hence, yfinance library is used to fetch the 10 years of historical data of a specific ticker to be used as the input of the prediction model. For each selected ticker, this historical data forms the foundation of the input for the prediction model. Then, several technical indicators, such as SMA, EMA, RSI, MACD, and Bollinger Bands, are calculated and included in the data frame as well. This is because technical indicators are indicator used to analyse entry and exit points for trades. By including this into the data frame, it is able to provide additional features that enhances the model's capability to capture the market patterns better. Below are the explanations for each technical indicator:

1. Simple Moving Average (SMA)

   This indicator is used to smooths out the price data by averaging prices over a specified number of periods.

   Its formula is as follows:

$$SMA_{window} = \frac{1}{n}\sum_{i=0}^{n-1} Close_{t-i}$$

2. Exponential Moving Average (EMA)

   This indicator gives more weight to recent prices, and hence it is more responsive to most recent information.
   Its formula is as follows:

$$EMA_{window} = Close_t \times \frac{2}{n+1} + EMA_{t-i} \times (1 - \frac{2}{n+1})$$

3. Relative Strength Index (RSI)

   This indicator measures the speed and change of price movements to identify overbought or oversold conditions.

   It is calculated as follows:

$$RSI = 100 - (\frac{100}{1 + RS})$$

4. Moving Average Convergence Divergence (MACD)

This indicator indicates changes in the direction, momentum, strength, and duration of a trend in a stock's price.
Its formula is as follows:

$$MACD = EMA_{fast} - EMA_{slow}$$

$$MACD\_Signal = EMA_{MACD}$$

5. Bollinger Bands

This indicator consists of a middle band (SMA) and two outer bands (standard deviations) that indicate volatility.

Its formula is as follows:

$$BB_{Middle} = SMA_{window}$$

$$BB_{Upper} = BB_{Middle} + 2 \times SD$$

$$BB_{Lower} = BB_{Middle} - 2 \times SD$$

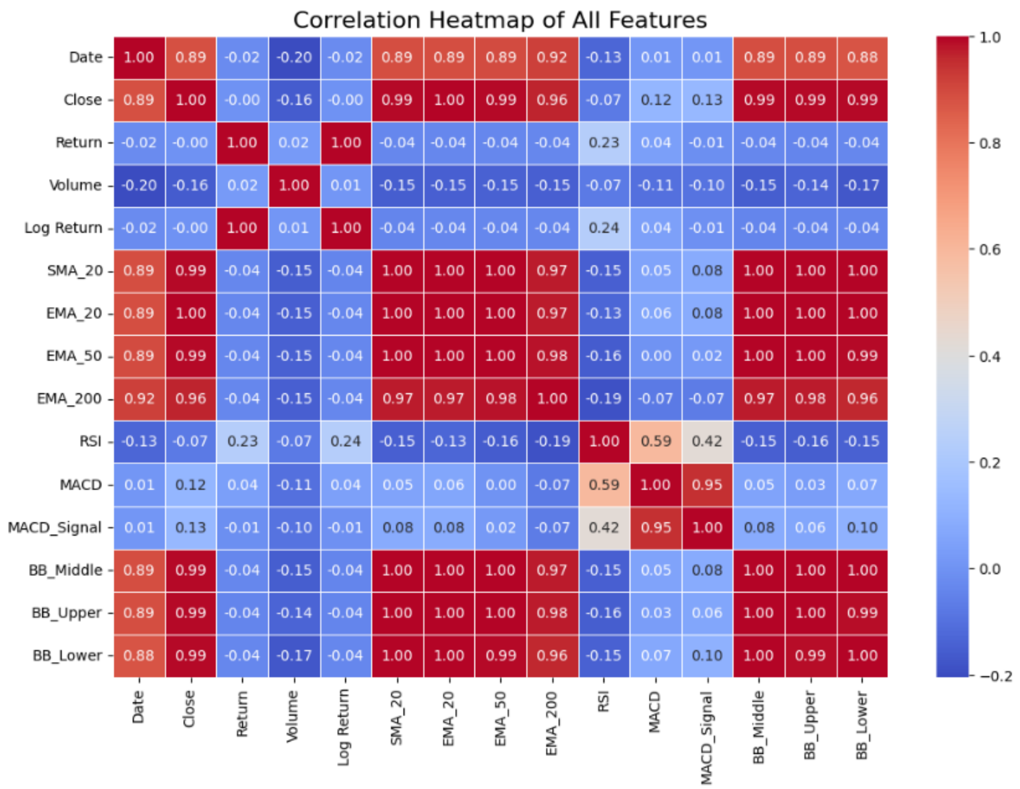



*Figure 4. 2 Correlation Graph of Features*

Then, a correlation graph is plotted is to assess the relationship between different features. By doing so, the most features that have the strongest association with the close prices (target variable) are selected as the input to the LSTM model. To enhance the model's efficiency. From the correlation graph, Bollinger Band, SMA, EMA have strong association with the closing price. Hence, SMA. EMA and BB_Upper and BB_lower are included in the input together with closing price. Now the data frame contains Close, SMA, EMA, BB_Upper and BB_Lower as the attributes. This project will focus on short term trading, and hence SMA_20 and EMA_20 will be chosen over window of 50 and 200. This is because the SMA_20 and EMA_20 indicators are more responsive to recent price changes, which allows the model to quickly adapt to the latest market conditions. After that, data cleaning is also done to drop rows with null values that were created during feature calculations.

After that, the dataset is split into training and testing sets using 70/30 ratio. The preprocessing steps also include scaling the data using MinMax Scaling to scales features and make the values to range between 0 and 1 to improve model's performance. Moreover, the data is partitioned into sequences of 20 days for input and the corresponding next day's data for output.

### 4.3.3 Model Training and Tuning

Model: "sequential_2"

| Layer (type) | Output Shape | Param # |
|---|---|---|
| lstm_4 (LSTM) | (None, 20, 128) | 68,608 |
| dropout_4 (Dropout) | (None, 20, 128) | 0 |
| lstm_5 (LSTM) | (None, 64) | 49,408 |
| dropout_5 (Dropout) | (None, 64) | 0 |
| dense_2 (Dense) | (None, 5) | 325 |

Total params: 118,341 (462.27 KB)
Trainable params: 118,341 (462.27 KB)
Non-trainable params: 0 (0.00 B)

*Figure 4. 3 Model Architecture*

Before model training, the architecture of the LSTM model is built. For building the model, PyTorch are used. This model comprises of two LSTM layers, the first with 128 units and return sequences enabled, while the second layer with 64 units and with the return sequences being disabled. Both LSTM layers use the ReLU activation function with 20% dropout applied after each to mitigate overfitting. Finally, the final output is provided by the dense layer, with the number of units corresponding to the features being predicted. Then, the model is compiled using the Adam optimiser. The learning rate of the model is set to be 0.0001 and Mean Squared Error (MSE) set as the loss function. During training, early stopping is implemented to terminate the process if the validation loss indicates that the model is not learning for 5 consecutive epochs. The model is trained for a maximum of 10 epochs with a batch size of 32. The training epoch is set to 10.

### 4.3.4 Model Evaluation

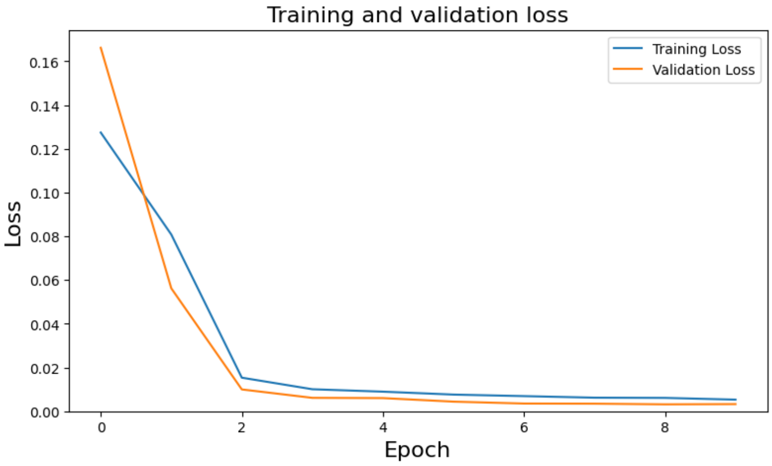



*Figure 4. 4 Loss visualization for AAPL*

Model evaluation begins with a crucial step, which is to visualise the training and validation loss during model training. By plotting the losses across epochs, the model's performance can be assessed. Ideally the training and the validation loss curves should converge and stabilise, which indicates that the model is not just memorising the training data. However, a significant divergence between the two curves might suggest that the model is overfitting, which indicates that adjustments to model's hyperparameter are necessary.

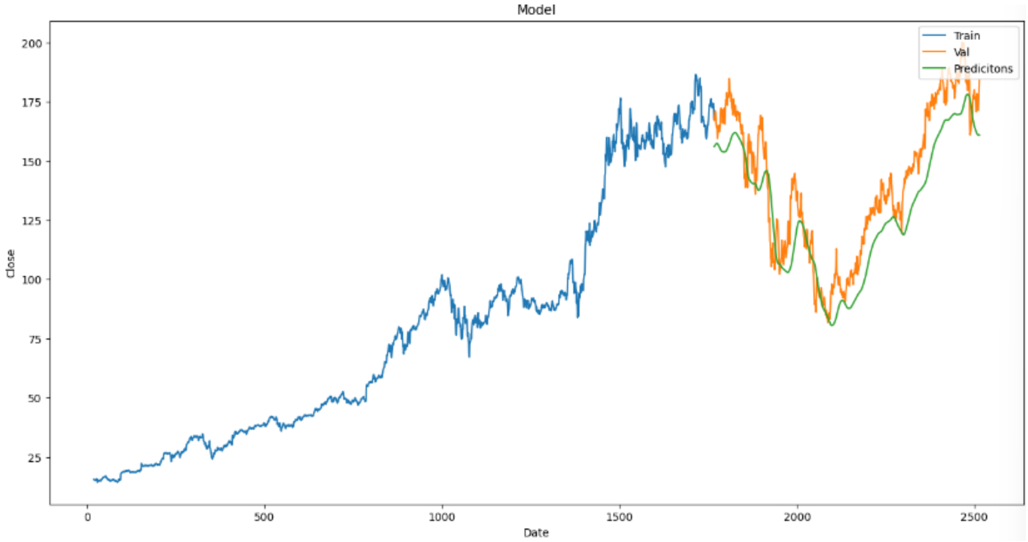



*Figure 4. 5 Predicted vs Actual Value on Testing Set*

Moreover, to assess the model's accuracy, error metrics such as RMSE, MAE are calculated. For this experiment, we have trained the model using AAPL stock, the RMSE and MAE is 7.65 and 5.83 respectively. The RMSE is used for gives insights into how much the predicted values deviate from the actual values, and hence the lower the RMSE indicates the more accurate the prediction. On the other hand, the MAE offers indication of the typical size if error in the model's performance. The figure above also shows the predictions of the model on test data. It is shown that the predictions are accurate in term of the trend. The model is trained spontaneously on the website using the historical data of the selected ticker. This approach is used, rather than pre-training the model on a large array of different is because it is computationally expensive to train model on almost all tickers, which is not practical.

## 4.4 Sentiment Analysis Development

Stock market is extremely volatile and influenced by multiple factors. Hence, to make the stock price prediction more accurate for executing trades, sentiment analysis is done to further aid the decision for trading strategies. Sentiment Analysis plays an important role in predicting stock market trends by extract insights from news articles and social media. This approach used to interpret the market news to determine the sentiment behind it. In the context of stock market prediction, the sentiment analysis uses NLP and machine learning techniques to class the text as positive, negative or neutral.

### 4.4.1 Data Collection

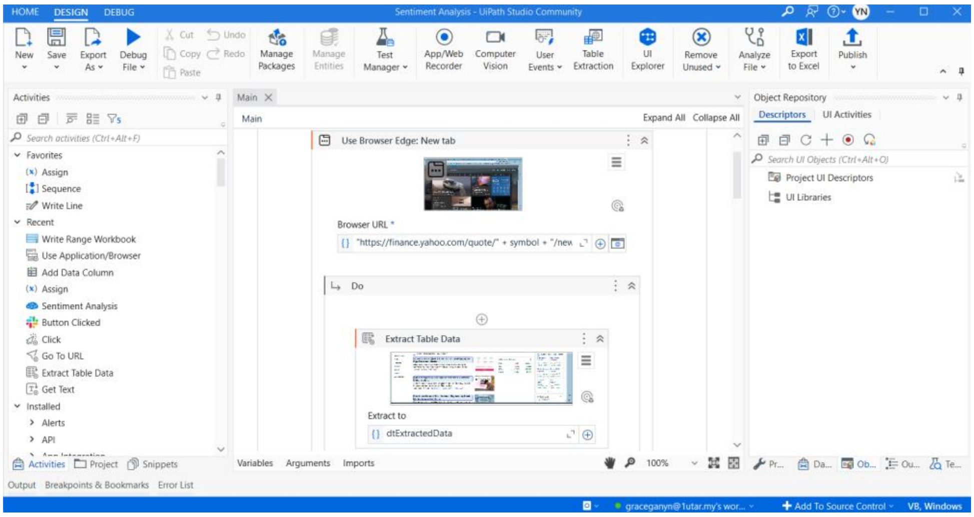



*Figure 4. 6 RPA Sequence Built in UiPath*

In developing the sentiment analysis tool for this project, RPA is leveraged to collect data from different sources. The RPA is constructed using UiPath. The process begins by taking in input for ticker's symbol and a date. Then, the web scraping function in UiPath is used to perform the task to source news data from different sources. The RPA is designed to source data from Yahoo Finance and Twitter. The reason why RPA is used to collect news data, rather than calling API that can return news data is due to most of the API has limited usage due to the API pricing. By taking in the ticker symbol and date as the input, RPA will only fetch news data for the specific stock and the specific date. The collected data is then cleaned and organised into a data frame. The data frame should contain different column for news headline, description, and date. This data frame will then be return as the output to the backend to be processed.

### 4.4.2 Sentiment Analysis Model

In this project, VADER (Valence Aware Dictionary and sEntiment Reasoner) is used for the performing the sentiment analysis. VADER is chosen due to its effectiveness in handling social media text and its capability of assigning sentiment scores (positive, negative, neutral) to each piece of text. Moreover, VADER also assign texts with a compound score, which is normalised value ranging from -1 (most negative) to 1 (most positive). After receiving news data from the RPA, the sentiment analysis model starts to analyse the data based on the headline and the description. Both headline and description will be given a score (positive, negative, neutral). Based on this generated score, the final score is then generated by setting a logic to check if either headline or description is positive, then the final score will be positive; if either headline or description is negative, then the final score will be negative. Finally, the sentiment with the highest occurrence across all rows will then be the overall sentiment for the given stock on the specific day.

### 4.5 Trade Bot Development

The trade bot in this project is designed to execute trades based on the combined results of stock price prediction and sentiment analysis. The purpose of this bot is to handle the entire decision-making process, eliminating the need for manual intervention by the user. This provides an automated solution for the users in trading, where trade analysis will be done by the bot and hence trigger the buy/sell actions based on the results.

Firstly, the user's order list will be passed to the trade bot, which means the bot will only perform trade execution on the stocks that have been ordered by the user. Then, the two key components (stock prediction model and sentiment analysis model) are integrated into the trade, in which at the beginning of the trade bot execution, the trade bot will call the predict stock and sentiment analysis functions to run and generate results. Then, the trade bot receives the results from stock prediction and sentiment analysis. Based on the results, the trade decision logic is constructed. If the results generated by both functions indicate a positive result, the bot will place a buy order in the anticipation of price rise, by calling the Alpaca API. On the other hand, if both results indicate a negative trend, then trade bot will issue a sell order to mitigate loss. However, discrepancy between prediction and sentiment results is inevitable. Hence, logic to handle conflict between both results are also implemented. For example, in the scenario where the prediction forecasts an increase, but the sentiment analysis is negative, or vice versa, the bot will neither buy nor sell.

Moreover, the trade bot will be scheduled to execute once a day at a specific time. In this case, the bot is set to execute at 8.00 a.m. Eastern Time, which is before market open. This is because the trade bot forecasts price for next day, by trading before the market open ensures the bot reacts before significant price moves occur. This timing aligns well with the LSTM's forecast. The scheduling of the trade bot is done by using cron. By doing this, the trade bot is able to run even if the user is not logged in to the website. Hence, the trade execution is fully automated.

To provide a comprehensive insight of the bot's performance and trading outcome, a simple summary report will be generated and displayed to user. The Gemini API is leveraged to generate report to users. At the end of a trading session, the trade bot will call the Gemini API to compile a report that includes the prediction made by the LSTM model, sentiment analysis results, trades executed by the bot, and additional information. In the development of this use case, a prompt message is passed to the Gemini API, and it will generate summary report based on the prompt and the input. The generated report is designed to be beginner-friendly, in which it should be easy-to-understand and is able to keep users well-informed with the bot's action.

### 4.6 Backend of Web Application

The backend of the Web Application is developed using Flask, in which it utilises Python to build the server-side of the web application. The backend contains function to retrieve list of tickers for each market and pass them to the client-side of the web application to allow users to select ticker in the Prediction Interface. Moreover, the backend is responsible for predicting the future price of the ticker and then pass the results to the frontend to display it to users. Hence, the LSTM prediction model is built in the backend of the web application, in which it exists in a function called predict stock where the server will call this function when user initiates the prediction process. Apart from that, the sentiment analysis model, trade execution logic and trade bot are also developed in the backend of the web application. All the functionalities of the website are integrated in the backend so that it streamlines the process and does not expose sensitive operations or data to the frontend, ensuring a secure and efficient user experience while maintaining a clear separation between client-side interactions and server-side processing. Moreover, the backend uses SQLAlchemy to interact with the database, in which the database stores the user information. The database used in this project is SQLite because it is lightweight and require minimal configurations. Its serverless architecture makes it ideal for applications where simplicity and ease of use are paramount, without the overhead of a full-fledged database server.

## 4.7 Integration

To ensure a cohesive user experience for the web application, the integration of the frontend and backend of the website are integrated. By integrating both the frontend and backend, frontend can request for data from the backend, whereas backend is able to pass the response back to the frontend. This is done by expose backend’s API endpoints to the frontend. For example, to request for stock data, frontend can call the GET /api/stockdata route. Moreover, the authentication of user is also a result of the integration, where the backend will receive the access token from Alpaca’s authorization, and the frontend will then include the token in the headers of requests to access protected routes, ensuring secure interactions. In this project, frontend used the ‘fetch’ method to route to the backend’s endpoint. By using fetch, the frontend can asynchronously request and receive data, ensuring a responsive and interactive user experience while keeping the data exchange process efficient and straightforward.

# CHAPTER 5 SYSTEM IMPLEMENTATION

## 5.1 Software Setup

### 5.1.1 Development Environment

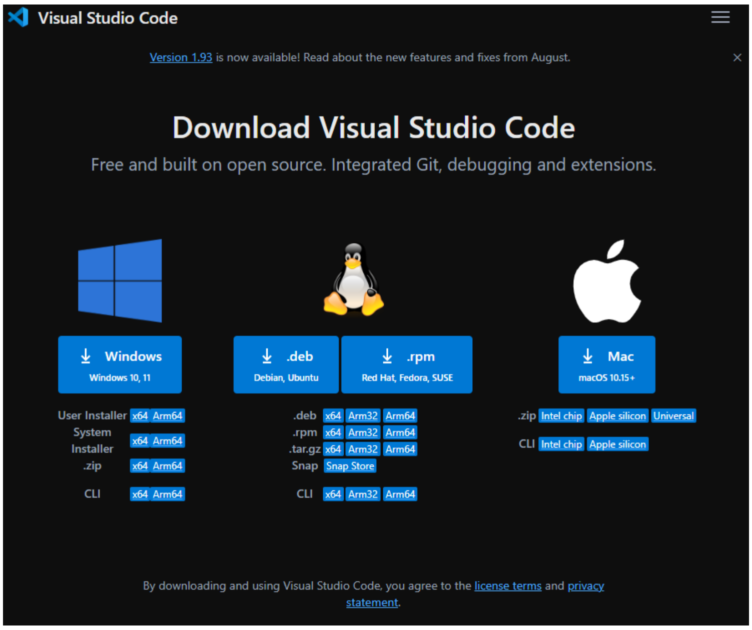



*Figure 5. 1 VS Code Installation*

The development of the website is carried out using Visual Studio Code, which provides features such as code syntax highlighting, debugging tools, various extensions for enhanced productivity. Visual Studio Code can be installed from its official website.

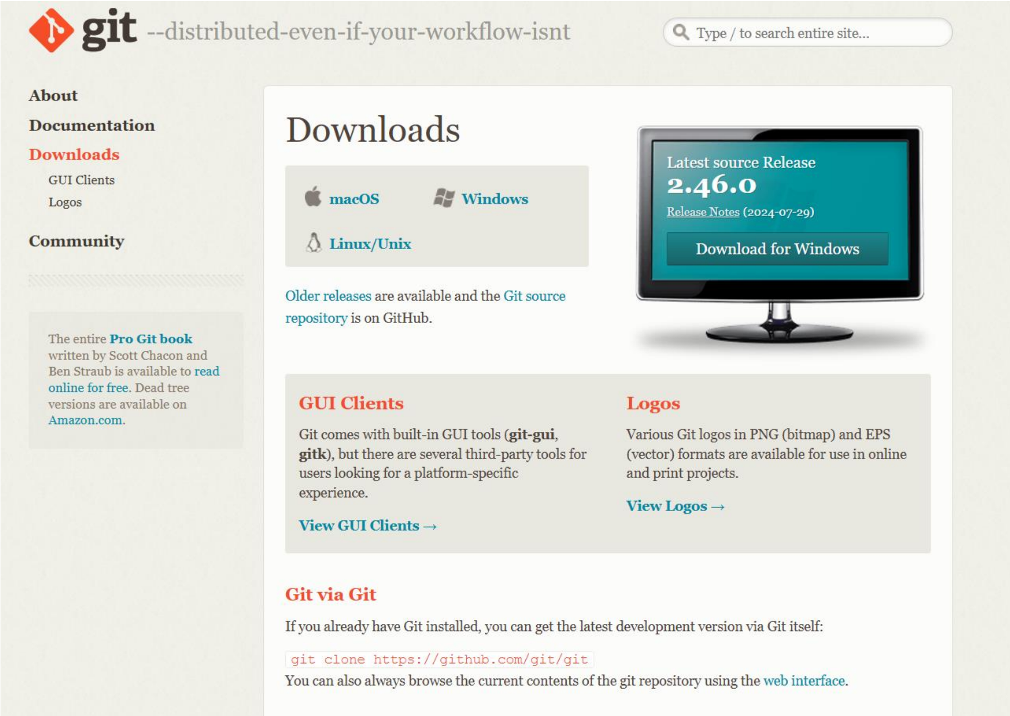



*Figure 5. 2 Git Installation*

Git is used in this project development for version control. Git enables tracking of codes changes, and this is useful when changes made are required to revert to the previous version. Git can be downloaded from the official site. Next, user can install Git based on their operating system.

After Git and VS Code have been installed, it is required to configure Git in VS Code in order to use it for version control. User must type Git:Clone in the command palette. Then, set up Git user information.

### 5.1.2 Frontend Setup

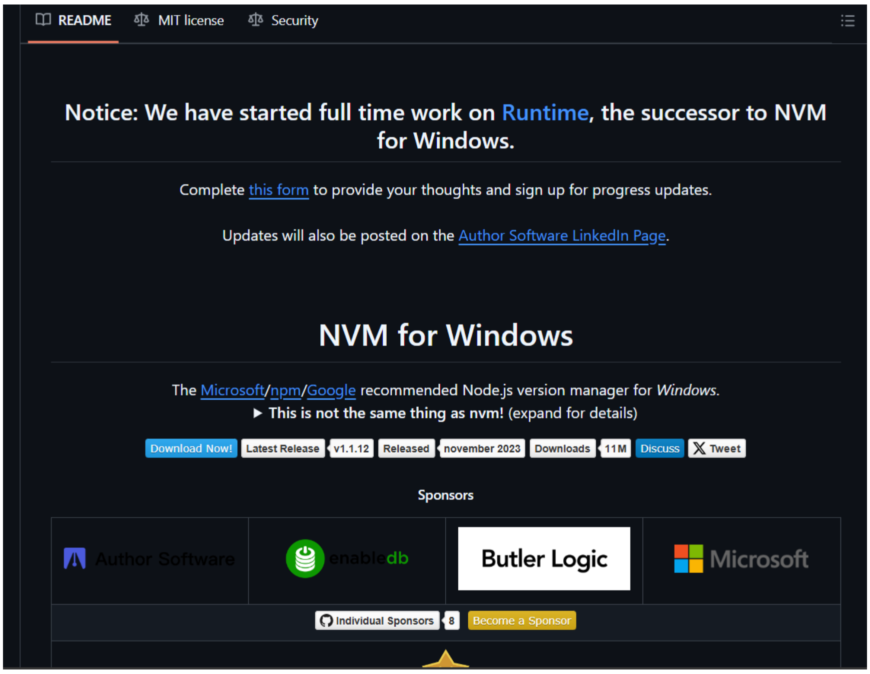



*Figure 5. 3 nvm Installation*

For the frontend development of this project, nvm is needed. Firstly, the nvm can be installed from the Windows NVM package manager for Windows user. Next, run "nvm –version" in the terminal to verify nvm installation. Then, install the latest version of Node.js.

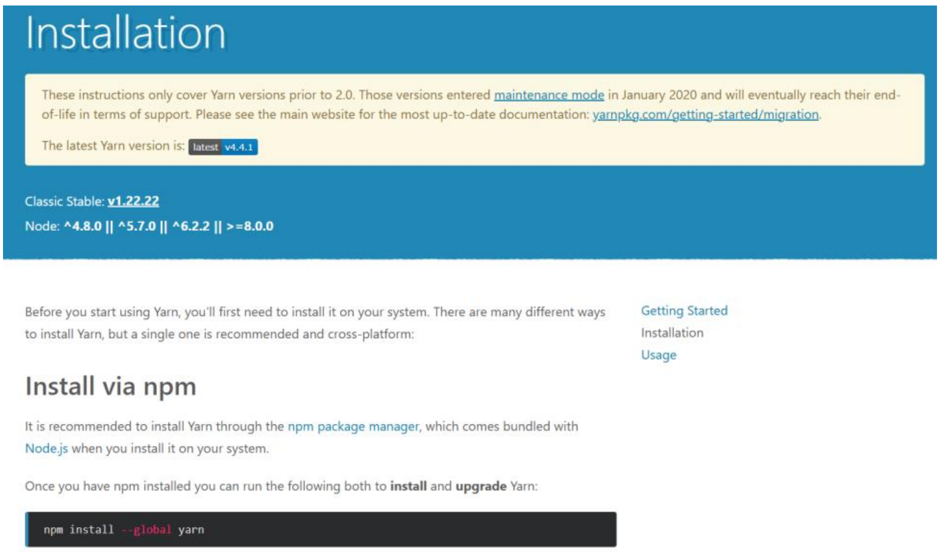



*Figure 5. 4 Yarn Installation*

Yarn is also used in the development of this project. Yarn is a package manager that replaces npm for faster and more reliable dependency management. Firstly, with Node.js installed, run the command “npm install –global yarn” to install Yarn. Then, similar to nvm, run “yarn –version” to verify yarn installation.

Next, setup for React project using Yarn is required. To create a react app uing yarn, run the command “yarn create react-app my-app”. Next, navigate into the project folder and start the development server by typing “yarn start” in the terminal.

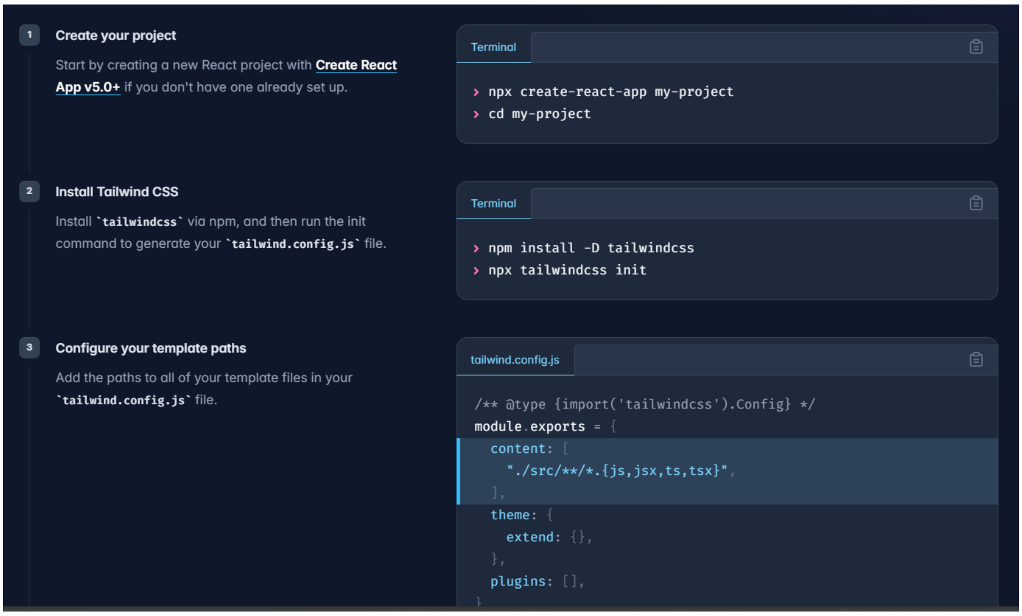



*Figure 5. 5 Tailwind Configuration*

Tailwind CSS is used in this project to style the React components. Firstly, install Tailwind CSS and its dependencies using Yarn: yarn add -D tailwindcss postcss autoprefixer. After that, initialise the Tailwind CSS configuration files by running “npx tailwindcss init -p” in the terminal. Then, configure the content paths in the tailwind.config.js file and add tailwind directive to the css file. Tailwind CSS is now ready to be used.

### 5.1.3 Backend Setup

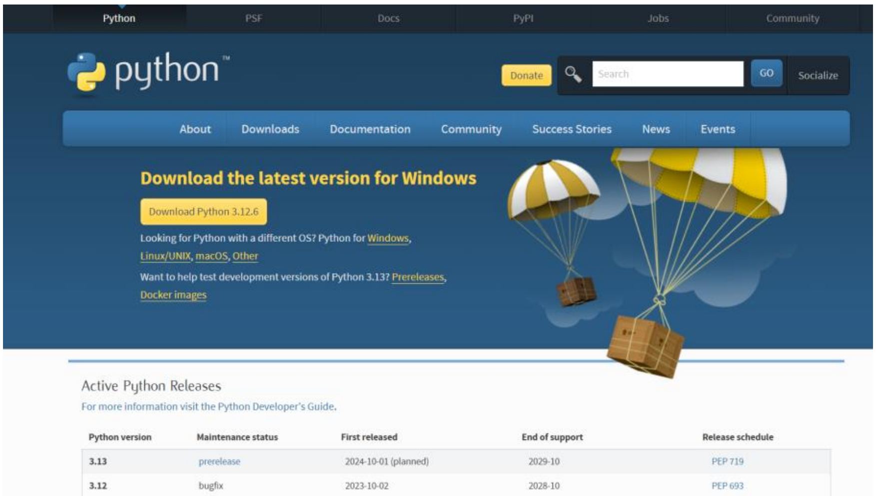



*Figure 5. 6 Python Installation*

For windows user, Python can be installed from its official website. During installation, the check box "Add Python to PATH" must be checked.

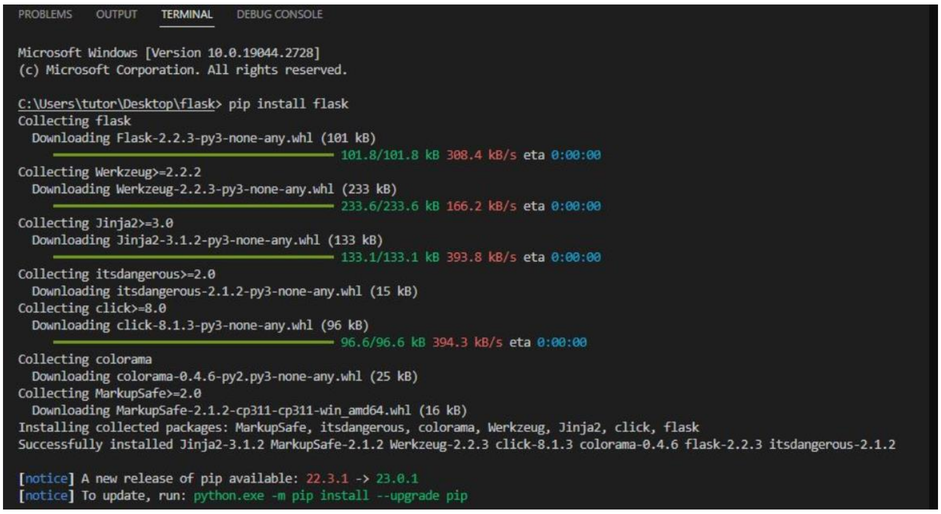



*Figure 5. 7 Flask Installation*

For the backend development, Flask is used. Firstly, install venv to create a virtual environment. Then, install Flask by running the command "pip install Flask". Then, verify Flask installation by checking its version. After this, set up Flask application with the following command: mkdir flask-app, cd flask-app, touch app.py

### 5.1.4 RPA Setup

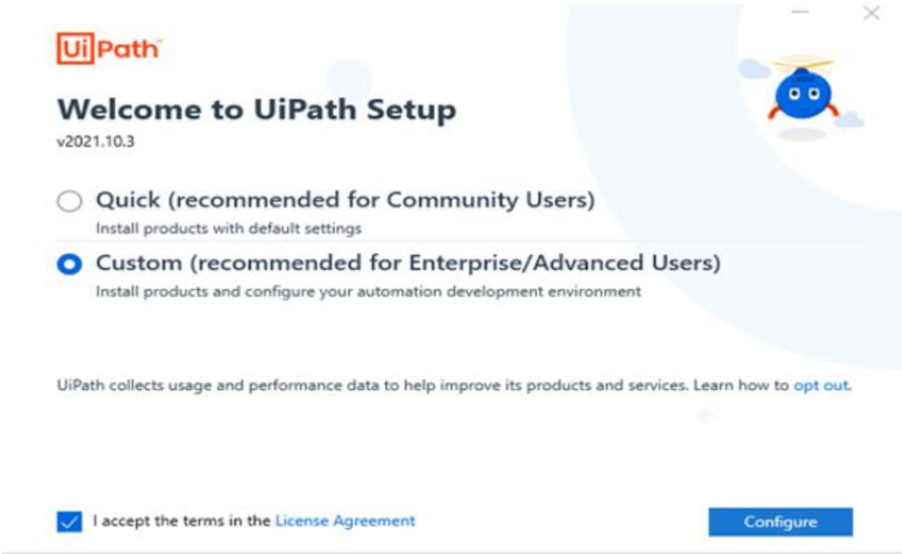



*Figure 5. 8 UiPath Installation*

Before installing UiPath, go to UiPath website and sign up for a free account. After signing up, log in and navigate to the Automation Cloud. Then click on the Resource Centre and download UiPath Studio. Open downloaded installer and follow the instructions. After installation, UiPath will automatically launch

## 5.2 System Operation (With Screenshot)

### 5.2.1 Stock Dashboard


*Figure 5. 9 Stock Dashboard Page*

Based on the figure shown above, the 'View Market Information' Interface comprises of a few sections, which are namely, the header which contains search bar to search for ticker, ticker price overview, ticker info and ticker chart. This page serves as a dashboard to summarise and display the important information of a ticker. This dashboard currently only shows the historical data of the ticker in candlestick pattern. This interface is expected to show more types of charts to show ticker information in an understandable way. In this part, finnhub API is used to fetch ticker symbol, and ticker information, however, the historical data needed to display the chart requires a premium package to be fetched. This way, RPA will be used to fetch historical data from yahoo finance in order to exhibit the chart. As shown in the figure, the theme of the web application is also developed where users get to switch between light mode and dark mode based on their preferences. The style of the web application is mostly implemented using Tailwind CSS and React Apex Chart is used for displaying the chart.

### 5.2.2 Predict Stock Prices

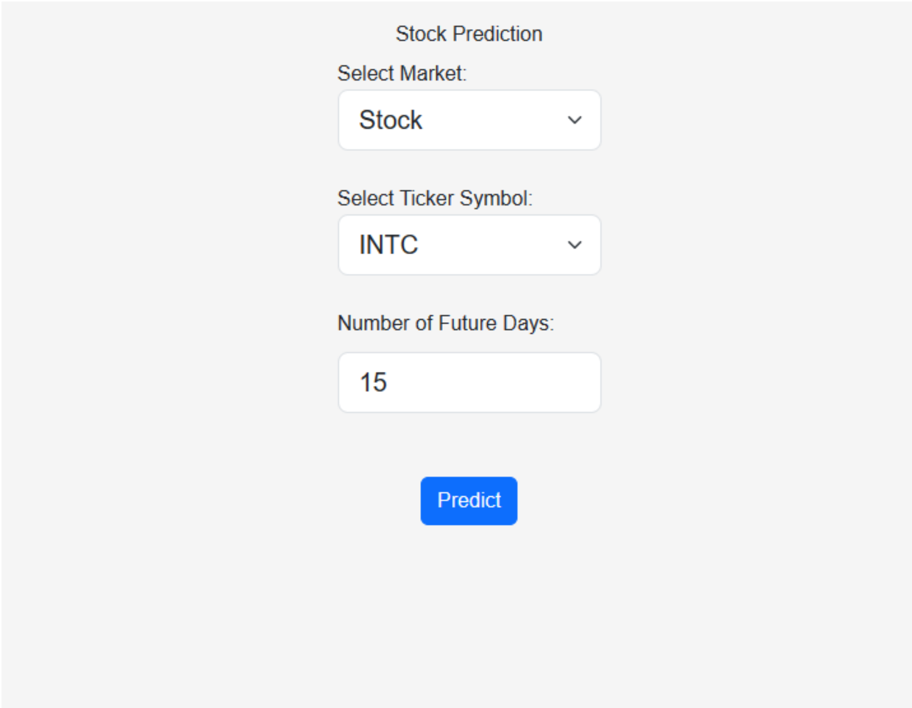



*Figure 5. 10 Stock Prediction Page (Before Prediction)*

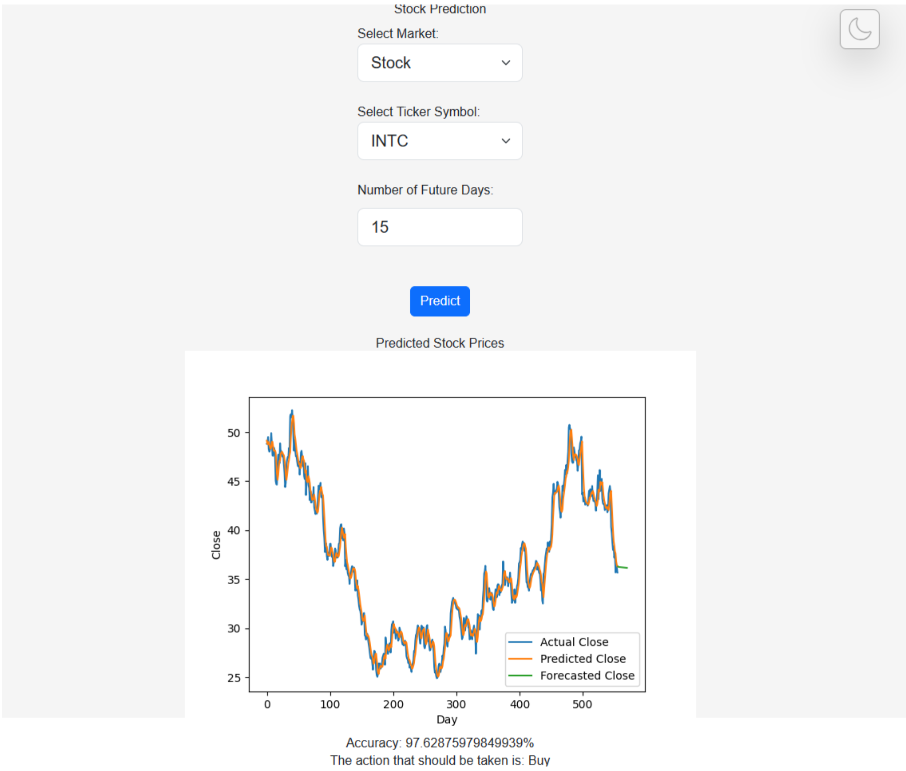



*Figure 5. 11 Stock Prediction Page (After Prediction)*

Figure above depicts the interface for the prediction of future ticker trend. When users first navigate to this page, it requires user to fill in information for prediction such as market, ticker symbol and the future days to predict. Choices of market selection provided in this web application are stock, world indices, currencies and ETFs. Upon filling in all the necessary information, by pressing the "predict" button will start the prediction. This process may require some time to load as the prediction model need some time to train and generate results. After results have been generated, a graph of the predicted value vs actual value will be shown, and the forecasted trend will also be displayed in the chart. Below the chart is the section for showing the accuracy, calculated using $R^2$ and the action recommendation for user.

### 5.2.3 Sentimental Analysis

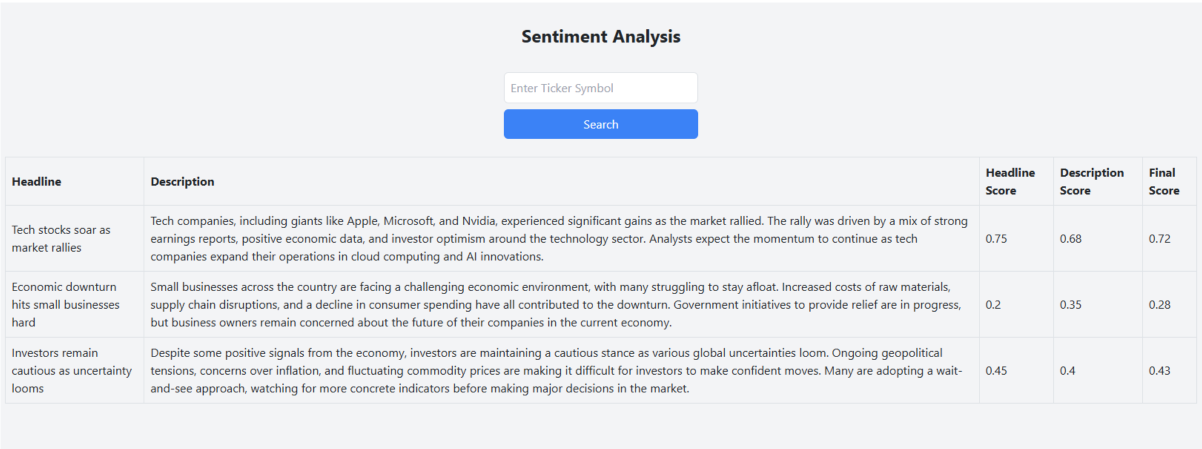



*Figure 5. 12 Sentimental Analysis Page*

This is the sentimental analysis page, where users can search for specific stock to analyse. After user has searched for a stock, the system will take in the symbol as input and call the function to retrieve relevant news data for today (the current date). Then, the VADER analyser will be responsible for analysing the news data by categorise them as positive, negative or neutral. After analysing the text, scores will be generated and displayed in a tabular form to users.

### 5.2.4 Paper Trading

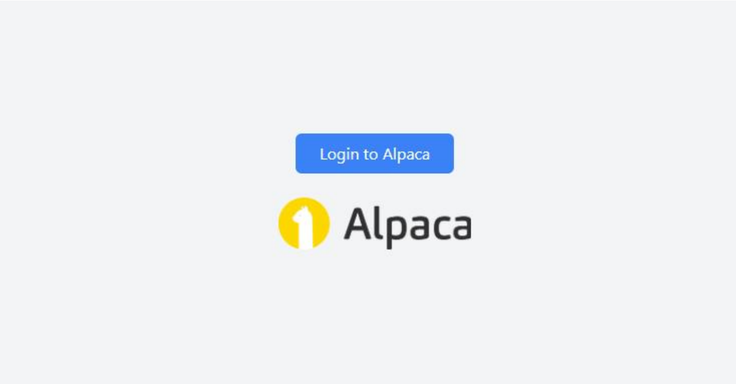



*Figure 5. 13 Login to Alpaca*

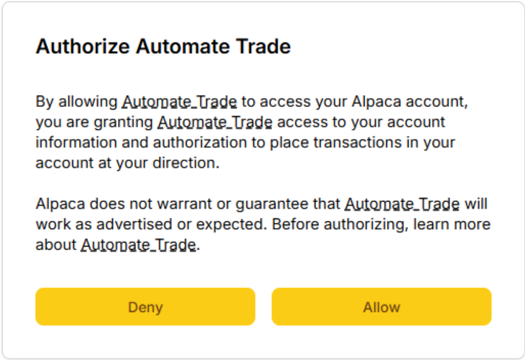



*Figure 5. 14 Alpaca Account Authorisation*

This is the interface that user will see before logging in into their Alpaca's Trading account. Login to Alpaca is required before user perform any paper trade in this website. This is because users need to authorise their Alpaca accounts to this website so that any actions performed in this website is instantly reflected in their Alpaca account.

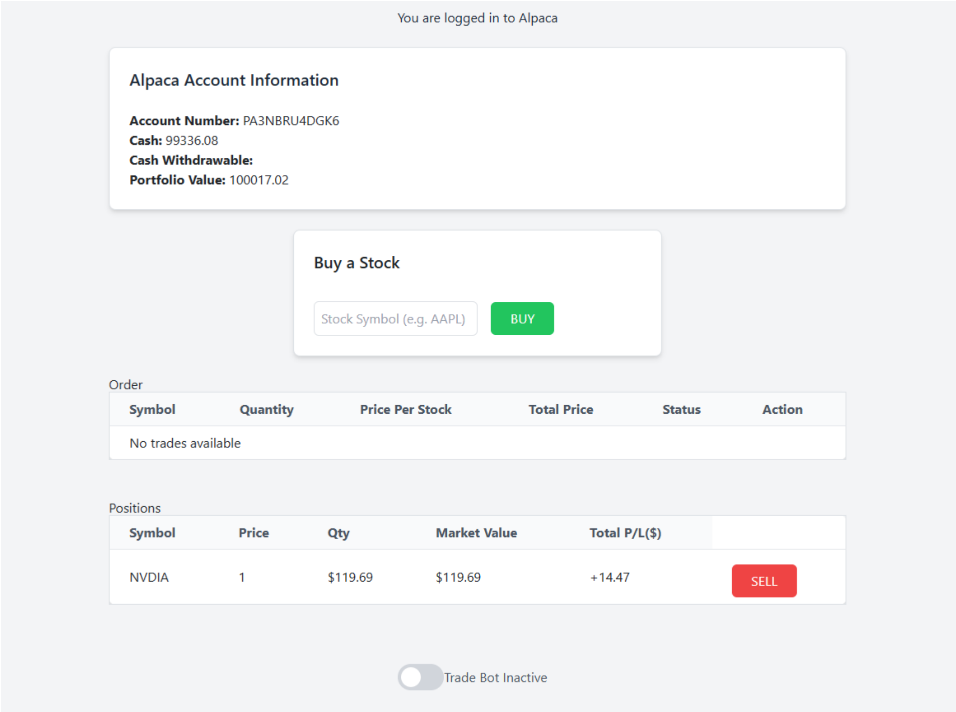



*Figure 5. 15 Paper Trading Page*

After users confirm the authorization, users will be redirected back to the paper trading page. Now data being displayed in this page are being fetched using the Alpaca API. In this web page, users can make simple simulated trade. However, for more trading tools, it is recommended that users can directly browse into their Alpaca account. This page is simple and containing minimal functions because the focus of this project is on the algorithm of automated trading. Hence, more focus will be given to the trade bot and how it works on user's Alpaca account. From the Figure above, there is a toggle button to activate trade bot. Once the button is activated, it schedules for the trade bot to run at a specific time once a day. The trade bot will make trade decision based on the prediction model and the sentimental analysis.

## 5.3 Implementation Issues and Challenges

### Challenges faced during FYP 1

One of the issues faced during the implementation of the system is that the LSTM model when first integrated into the backend of the web application, it cannot run the prediction smoothly as there are errors that kept arising. Initially, the model was built separately in Jupyter Notebook to ensure the model is able to perform well in prediction. However, when migrating the code into the backend of the web application, it faces errors. Hence, debugging was done step-by-step all the way from data preparation to model evaluation. Moreover, another issue faced is the prediction results passed from backend to frontend is unable to display on the web application. To show the prediction results, a simple chart will be displayed using matplotlib. The solution to this is that the visualisation is plotted in the backend, and it will just pass the encoded plot to the frontend to be displayed. After searching for solution online, this issue is resolved by debugging a line of code. There is also one significant issue faced, which is a run time error will occur after the plotted image for prediction is displayed. The error indicates that the main thread is not in main loop. This usually happens when the graphical user interface (GUI) elements are accessed from a thread other than the main thread. This needs to be solved by ensuring that GUI-related operations are always performed within the main thread.

**Challenges faced during FYP 2**

The stock prediction model developed during FYP1 exhibited overfitting, making it difficult to fine-tune and generalise to unseen data. Then, the training process of the model is revisited, by implementing regularization techniques, and adjusting the model architecture. Moreover, the market contains vast number of stocks, and training the model on all the tickers are practically impossible. To overcome this, the prediction process was modified to focus on one ticker at a time and the model is trained spontaneously each time user trigger a stock prediction operation rather than pre-training the model on all tickers. This ensures efficiency and relevance for each stock prediction.

Furthermore, the trade bot operates in the backend of the web application, there is no visible indication for the end-user to know that the bot has executed, making it difficult for users to verify its activity. To provide users with a clear acknowledgement about the execution of the bot, generating report or log after each execution can be a useful strategy. Hence, the trade bot will log the results and display them to the users and also generate report using Gemini API.

# CHAPTER 6 SYSTEM EVALUATION AND DISCUSSION

## 6.1 System Testing and Performance Metrics

### 6.1.1 User Authentication

| Test Case | Test Description | Test Data | Expected Result | Pass/Fail |
|---|---|---|---|---|
| User Login (Valid Credentials) | Test user login with valid credentials | Username: user1, Password: pass123 | User should be logged in and redirected | Pass |
| User Login (Invalid Credentials) | Test user login with incorrect password | Username: user1, Password: wrong | Error message: "Invalid credentials" and redirected to public view | Pass |
| User Logout | Test user logout | Logged-in user | User should be logged out and redirected | Pass |

*Table 6. 1 Test Cases for User Authentication*

### 6.1.2 Stock Prediction (LSTM model)

| Test Case | Test Description | Test Data | Expected Result | Pass/Fail |
|---|---|---|---|---|
| Stock Prediction (Valid Ticker) | Test stock prediction with a valid ticker | Ticker: AAPL | Prediction for AAPL should be displayed | Pass |
| Stock Prediction (Invalid Ticker) | Test stock prediction with an invalid ticker | Ticker: ABCD | Error message: "Invalid ticker" | Pass |

| | | | | |
|---|---|---|---|---|
| Stock Prediction (API Failure) | Test behaviour when stock prediction API fails | Ticker: AAPL | Error message: "Unable to fetch prediction" | Pass |

*Table 6. 2 Test Cases for Stock Prediction*

### 6.1.3 Sentiment Analysis

| **Test Case** | **Test Description** | **Test Data** | **Expected Result** | **Pass/Fail** |
|---|---|---|---|---|
| Sentiment Analysis (Valid Ticker) | Test sentiment analysis with a valid ticker | Ticker: AAPL | Sentiment score and news headlines should display | Pass |
| Sentiment Analysis (Invalid Ticker) | Test sentiment analysis with an invalid ticker | Ticker: ABCD | Error message: "Invalid ticker" | Pass |
| Sentiment Analysis (No Input) | Test sentiment analysis with no input | Empty | Error message: "Ticker is required" | Pass |

*Table 6. 3 Test Cases for Sentiment Analysis*

### 6.1.4 Trade Bot Activation

| **Test Case** | **Test Description** | **Test Data** | **Expected Result** | **Pass/Fail** |
|---|---|---|---|---|
| Trade Bot Activation (Button Click) | Test activating the trade Bot | Click 'Activate' | Trade Bot should activate and start trading | Pass |

| | | | | |
|---|---|---|---|---|
| Trade Bot Deactivation (Button Click) | Test deactivating the trade Bot | Click 'Deactivate' | Trade Bot should stop trading | Pass |

*Table 6. 4 Test Cases for trade Bot Activation*

### 6.1.5 Paper Trading

| **Test Case** | **Test Description** | **Test Data** | **Expected Result** | **Pass/Fail** |
|---|---|---|---|---|
| Stock Purchase (Valid Transaction) | Test stock purchase with valid input | Ticker: AAPL, Quantity: 10 | Purchase should be successful | Pass |
| Stock Purchase (Invalid Ticker) | Test stock purchase with an invalid ticker | Ticker: INVALID, Quantity: 10 | Error message: "Invalid ticker" | Pass |
| Stock Purchase (Insufficient Funds) | Test stock purchase with insufficient funds | Ticker: AAPL, Quantity: 1000 | Error message: "Insufficient funds" | Pass |

*Table 6. 5 Test Cases for Paper Trading*

## 6.2 Testing Result (Based on Performance Metrics)

| Performance Metrics | Results |
|---|---|
| Response Time | Data retrieval – 530ms<br>Page load – 208ms |

| | |
|---|---|
| | API requests – 2.14s |
| CPU Usage | Idle – 0%<br>Navigation – 4% |
| Memory Usage | 36.2MB |
| Time taken for Stock Predictions | 12s |
| Time taken for Sentiment Analysis | 10s |
| Time taken for trade Bot | 40s – 1m |

*Table 6. 6 Test Results Based on Performance Metrics*

**6.3 Project Challenges**

The main challenge in the developing this project is to build and maintain an accurate predictive (LSTM) model that does not overfit or underfit. Overfitting occurs when a machine learning model learns the noise of the training data, and in turn leads to the model performing poorly on unseen data. This happens easily when training model on stock historical prices, this is because stock market data is inherently noisy with a lot of random fluctuations, in which it causes the model to memorise these random patterns. In stock market prediction, feature selection is a challenging task. This is because if too many features are included could contribute to overfitting as well. Moreover, the market trend is influenced by numerous factors including economic indicator, political events and market sentiment, which makes it difficult for a time-series model to predict the future trend just based on the historical prices.

Moreover, another challenging part about this project is the integration of different units, such as integrating prediction model and sentiment analyser into the trade bot and integrating RPA sequence into the sentiment analysis. Each component requires careful consideration to ensure seamless interaction between one and another. Integration of the prediction model with the sentiment analysis involves aligning their outputs to inform trading decisions accurately. Hence, the outputs from the prediction model and sentiment analyser must be in the correct format so that it aids in streamlining the trade bot execution. The RPA sequence needs to be effectively integrated with the sentiment analysis tool to automate the process of collecting and processing news data. This integration demands careful design to ensure that the RPA sequences can handle the data extraction and processing tasks efficiently, and that the results are accurately fed into the sentiment analysis pipeline.

## 6.4 Objectives Evaluation

The first objective of this project is to develop an LSTM-based deep learning model to forecast future market prices. This objective has been fulfilled as the project has implemented a LSTM model to forecast market prices by utilising the historical stock data to generate predictions, providing insights that can guide investment strategies. By incorporating techniques to manage overfitting and evaluating the model's performance with appropriate metrics, the project ensures that the LSTM model delivers accurate and reliable forecasts.

Next, the second objective is also fulfilled, which is to utilise RPA for sentiment analysis. In this project, RPA is developed using UiPath to construct the logic to fetch news data from different sources. RPA is integrated with sentiment analysis to analyse the sentiment of the news data, which provides insights from another perspective. This integration allows for a more comprehensive understanding of market sentiment, improving the quality of the analysis and the relevance of the insights generated.

The third objective is to create an automated trading bot that leverages both the LSTM model and sentiment analysis to make informed trading decisions. This objective is also fulfilled as the bot is implemented in this project, in which it integrates predictions from the LSTM model and sentiment insights to execute trades automatically. This functionality streamlines the trading process for users, ensuring that trades are based on a combination of historical data and current market sentiment, thereby optimizing trading strategies.

# CHAPTER 7 CONCLUSION

## 7.1 Conclusion

In a nutshell, the market price is hard to predict because market price is volatile, where it will constantly drop and rise within a short period of time. This is because financial markets are affected by multiple factors, for example, economic indicators, political events, performance of the company and so on. It is common for experienced traders to use different investments analysis to predict the future trend of a certain ticker. However, this is not beginner-friendly, as beginners may not have enough experience to be able to evaluate the market prices based on investment analysis and it requires years of experience to excel in that. In facing these challenges, many beginners refuse to invest in the financial markets. Hence, by providing an automated solution to predict the stock market, this project benefits those who need to learn the future trend of the market price in order to take correct actions in actual investment.

Moreover, it also eases traders in analysing and predicting the market price manually. It is a tedious and time-consuming process for traders to manually calculate the P/E ratio or even analyse based on different technical indicators, where traders have to constantly track and monitor the trades. With the implementation of automated data analytics, traders no longer have to hassle to predict trades manually. This is because this project will predict the future trend or prices of trades by using time-series analysis, in which the prediction will be done by Deep Learning model and action recommendations will be generated. Hence, this project aims to help ease the trading process by providing automated web application, to predict the future market trend and allow users to view real-time stock information.

## 7.2 Recommendation

Firstly, future work can be done to enhance the model accuracy. Continuous refinement and improvement can be made to improve the LSTM model, by incorporating additional features, experimenting with different hyperparameters, and using more advanced neural network architectures. This could involve integrating other types of deep learning models, such as Transformer-based architectures, and employing techniques such as cross-validation to ensure the model's robustness and generalizability. Regularly updating the model with new data will also help maintain its accuracy over time.

Then, research in enhancing the sentiment analysis can also be done by using more sophisticated Natural Language Processing (NLP) techniques and tools. This could include using pre-trained language models like BERT or GPT to capture nuanced sentiments and context from news articles and social media posts. Additionally, consider integrating sentiment analysis from multiple languages and sources to provide a more comprehensive view of market sentiment.

Last but not least, the trade bot performance can be enhanced by leveraging advanced algorithm for decision making in trade. Example of algorithms that can be used for further improvement is reinforcement learning or genetic algorithms. By using these techniques, trade bot can learn its mistake form the past experience and improve its decision strategy. In the scenario, where prediction model and sentiment analysis generate results that can contradicting, instead of hold, trade can learn and evaluate the decision based on the past experience to execute buy or sell.

## REFERENCES

[1] 3Commas, "3Commas: Smart Trading Terminal," *3Commas,* Accessed: Sep. 13, 2024. [Online]. Available: https://3commas.io

[2] Pionex, "Pionex: Crypto Trading Bot," *Pionex,* Accessed: Sep. 13, 2024. [Online]. Available: https://www.pionex.com

[3] TradeSanta, "TradeSanta: Crypto Trading Bot," TradeSanta, Accessed: Sep. 13, 2024. [Online]. Available: https://tradesanta.com

[4] P. S. Sisodia, A. Gupta, Y. Kumar, and G. K. Ameta, "Stock Market Analysis and Prediction for Nifty50 using LSTM Deep Learning Approach," in Proceedings of 2nd International Conference on Innovative Practices in Technology and Management, ICIPTM 2022, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 156–161. doi: 10.1109/ICIPTM54933.2022.9754148.

[5] Yichuan Xu and Vlado Keselj, Stock Prediction using Deep Learning and Sentiment Analysis. 2019.

[6] A. Bansal, D. N. Patel, K. Rishabh, and M. Sneha, "Stock Prediction Model using Seq2Seq and Bi-directional LSTM," in 4th International Conference on Circuits, Control, Communication and Computing, I4C 2022, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 275–278. doi: 10.1109/I4C57141.2022.10057690.

[7] Nile University and Institute of Electrical and Electronics Engineers, Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model.

[8] X. Ji, J. Wang, and Z. Yan, "A stock price prediction method based on deep learning technology," International Journal of Crowd Science, vol. 5, no. 1, pp. 55–72, Apr. 2021, doi: 10.1108/IJCS-05-2020-0012.

[9] L. Sayavong, Z. Wu, and S. Chalita, "Research on stock price prediction method based on convolutional neural network," in Proceedings - 2019 International Conference on Virtual Reality and Intelligent Systems, ICVRIS 2019, Institute of Electrical and Electronics Engineers Inc., Sep. 2019, pp. 173–176. doi: 10.1109/ICVRIS.2019.00050.

[10] P. Durga and T. Sudhakar, "Machine Learning Classification Methods for Stock Market Forecasting," in International Conference on Innovative Data

Communication Technologies and Application, ICIDCA 2023 - Proceedings, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 14–21. doi: 10.1109/ICIDCA56705.2023.10100106.

[11] R. S. Latha et al., "Stock Movement Prediction using KNN Machine Learning Algorithm," in 2022 International Conference on Computer Communication and Informatics, ICCCI 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICCCI54379.2022.9740781.

[12] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques," in Procedia Computer Science, Elsevier B.V., 2020, pp. 599–606. doi: 10.1016/j.procs.2020.03.326.

[13] A. Yusuf, L. Kumar, S. Third, and N. Krishnan, "RPA Bot Working With Stock Market Share Prices | Stock Analysis & Stock Price Prediction," 2022. [Online]. Available: www.jetir.org

[14] Vanayak Jadkar, Mayur Khandate, Vedant Gampawar, Pritesh Bhutada, and Leena Deshpande, "RPA for Stock Selection Process and Price Prediction Model using Machine Learning," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 7, pp. 2321–8169, Jul. 2022.

# APPENDIX

## FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| | |
|---|---|
| **Trimester, Year:** Y3S3 | **Study week no.:** 2 |
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

**1. WORK DONE**

[Please write the details of the work done in the last fortnight.]

During the past week, I was doing research on how to fine tune an LSTM model to prevent overfitting.

**2. WORK TO BE DONE**

Continue my work on fine-tuning the prediction to prevent overfitting.

**3. PROBLEMS ENCOUNTERED**

No significant problem encountered.

**4. SELF EVALUATION OF THE PROGRESS**

I had no idea on how to fine-tune the model by tuning the hyperparameter and the using other approaches

Supervisor's signature

Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| | |
|---|---|
| **Trimester, Year:** Y3S3 | **Study week no.:** 4 |
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

**1. WORK DONE**

[Please write the details of the work done in the last fortnight.]

I tried doing feature engineering by including some technical indicators as the input for the model. I have also tried using different learning rate and included early stopping. I have also included dropout layer in the model.

**2. WORK TO BE DONE**

RPA Sequence to fetch new data for sentiment analysis
A function to do sentiment analysis needs to be done

**3. PROBLEMS ENCOUNTERED**

Now the model does not overfitting, but the prediction accuracy is also not as high as before.

**4. SELF EVALUATION OF THE PROGRESS**

I have learned that cross-validation, regularization, early stopping, use more training data and so on, are approaches to prevent overfitting in ML model.

Supervisor's signature

Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| | |
|---|---|
| **Trimester, Year:** Y3S3 | **Study week no.:** 6 |
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

**1. WORK DONE**
[Please write the details of the work done in the last fortnight.]

Constructed RPA Sequence to fetch new from Yahoo Finance and Twitter. Sentiment analysis is implemented in the web site.

**2. WORK TO BE DONE**

Paper Trading page is also done, which includes integrating it with alpaca API.

**3. PROBLEMS ENCOUNTERED**

Problem faced when trying to create RPA sequence as there are limited tutorial found on the internet and I was not familiar with the UiPath tools.

**4. SELF EVALUATION OF THE PROGRESS**

Now I have learned how to use UiPath and will utilise it to automate any daily tasks that can be automated.

Supervisor's signature

Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| **Trimester, Year:** Y3S3 | **Study week no.:** 8 |
|---|---|
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

**1. WORK DONE**

[Please write the details of the work done in the last fortnight.]

Paper Trading Page is done.
20% of the trade bot logic is developed.

**2. WORK TO BE DONE**

Continue to work on the trade bot

**3. PROBLEMS ENCOUNTERED**

Sometimes calling the Alpaca API encounter issues.

**4. SELF EVALUATION OF THE PROGRESS**

Extensive debugging has been done to solve the problem, with the aid of ChatGPT.

Supervisor's signature

Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| | |
|---|---|
| **Trimester, Year:** Y3S3 | **Study week no.:** 10 |
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

| |
|---|
| **1. WORK DONE**<br>[Please write the details of the work done in the last fortnight.]<br><br>Development of Trade Bot |
| **2. WORK TO BE DONE**<br><br>Testing and debugging |
| **3. PROBLEMS ENCOUNTERED**<br><br>Encountered problem when trying to integrate prediction model and sentiment analysis into the trade bot |
| **4. SELF EVALUATION OF THE PROGRESS**<br><br>- |

Supervisor's signature

Student's signature

# FINAL YEAR PROJECT WEEKLY REPORT

*(Project II)*

| | |
|---|---|
| **Trimester, Year:** Y3S3 | **Study week no.:** 12 |
| **Student Name & ID:** Gan Yu Nyuk 21ACB06785 | |
| **Supervisor:** Ts Dr Ooi Joo On | |
| **Project Title:** Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation | |

**1. WORK DONE**

[Please write the details of the work done in the last fortnight.]

Report Writing

**2. WORK TO BE DONE**

Submission of Report

Presentation

**3. PROBLEMS ENCOUNTERED**

-

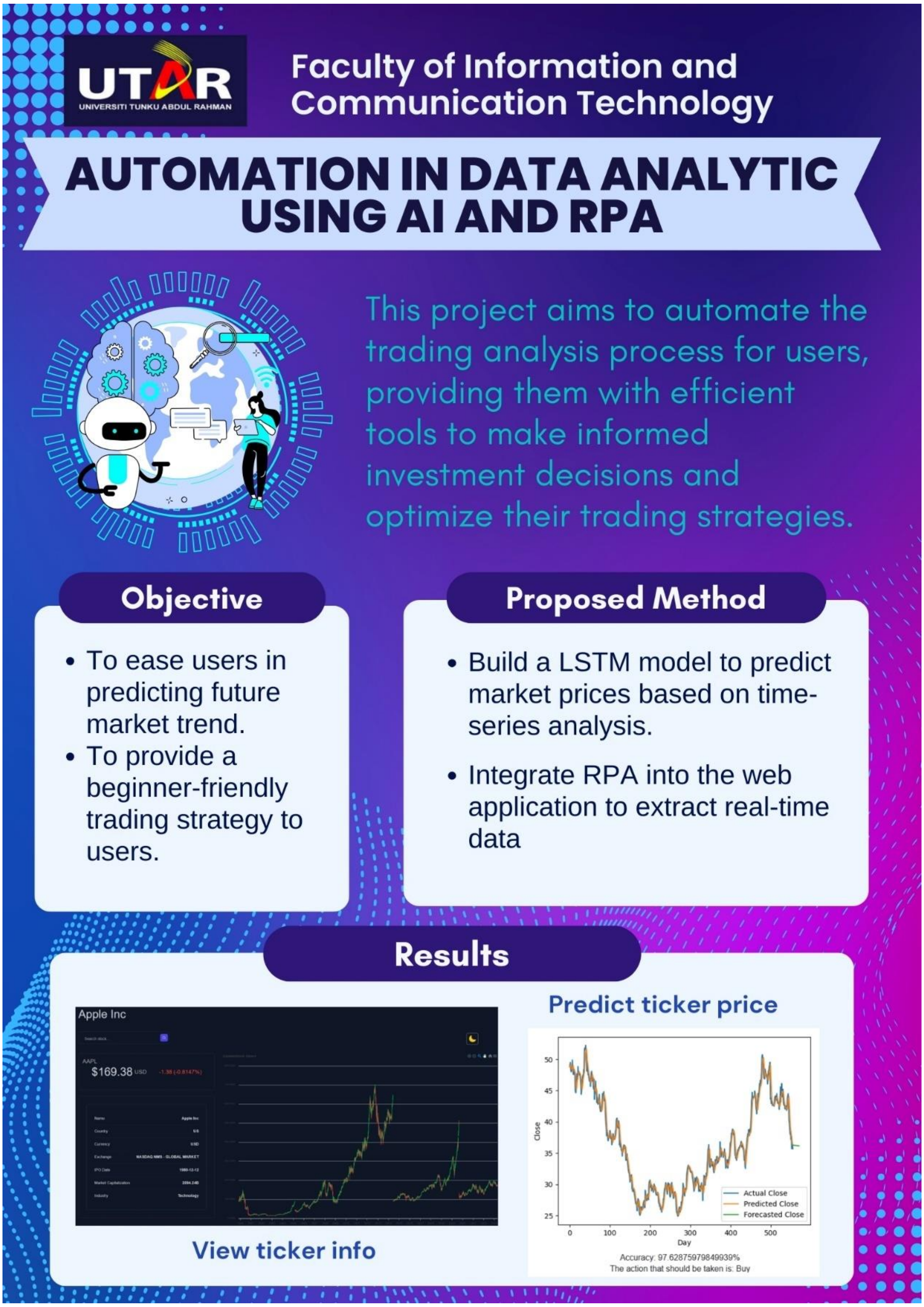
**4. SELF EVALUATION OF THE PROGRESS**

-

Supervisor's signature

Student's signature

# POSTER


UTAR
UNIVERSITI TUNKU ABDUL RAHMAN
Faculty of Information and Communication Technology
AUTOMATION IN DATA ANALYTIC USING AI AND RPA
This project aims to automate the trading analysis process for users, providing them with efficient tools to make informed investment decisions and optimize their trading strategies.
Objective
• To ease users in predicting future market trend.
• To provide a beginner-friendly trading strategy to users.
Proposed Method
• Build a LSTM model to predict market prices based on time-series analysis.
• Integrate RPA into the web application to extract real-time data
Results
Apple Inc
AAPL
$169.38 USD
View ticker info
Predict ticker price
Close
Day
Actual Close
Predicted Close
Forecasted Close
Accuracy: 97.62875979849939%
The action that should be taken is: Buy

# PLAGIARISM CHECK RESULT

FYP2_2106785

ORIGINALITY REPORT

| 10% | 8% | 3% | 5% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to Universiti Tunku Abdul Rahman<br>Student Paper | 2% |
| 2 | eprints.utar.edu.my<br>Internet Source | 2% |
| 3 | www.slideshare.net<br>Internet Source | 1% |
| 4 | 123dok.com<br>Internet Source | <1% |
| 5 | dokumen.pub<br>Internet Source | <1% |
| 6 | umpir.ump.edu.my<br>Internet Source | <1% |
| 7 | "Data Science", Springer Science and Business Media LLC, 2020<br>Publication | <1% |
| 8 | fastercapital.com<br>Internet Source | <1% |
| 9 | www.researchgate.net<br>Internet Source | <1% |

| | | |
|---|---|---|
| 10 | ieeexplore.ieee.org<br>Internet Source | <1 % |
| 11 | Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security - Volume 2", CRC Press, 2023<br>Publication | <1 % |
| 12 | Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security", CRC Press, 2023<br>Publication | <1 % |
| 13 | Akshat Bansal, Dency Narendra Patel, Khetan Rishabh, M Sneha. "Stock Prediction Model using Seq2Seq and Bi-directional LSTM", 2022 4th International Conference on Circuits, Control, Communication and Computing (I4C), 2022<br>Publication | <1 % |
| 14 | Submitted to Asia Pacific University College of Technology and Innovation (UCTI)<br>Student Paper | <1 % |
| 15 | Lounnapha Sayavong, Zhongdong Wu, Sookasame Chalita. "Research on Stock Price Prediction Method Based on Convolutional Neural Network", 2019 International | <1 % |

| | | |
|---|---|---|
| | Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2019<br>Publication | |
| 16 | ust.edu.ye<br>Internet Source | <1 % |
| 17 | Submitted to The University of Texas at Arlington<br>Student Paper | <1 % |
| 18 | Submitted to Southern New Hampshire University - Continuing Education<br>Student Paper | <1 % |
| 19 | Xuan Ji, Jiachen Wang, Zhijun Yan. "A stock price prediction method based on deep learning technology", International Journal of Crowd Science, 2021<br>Publication | <1 % |
| 20 | Submitted to University of KwaZulu-Natal<br>Student Paper | <1 % |
| 21 | open-innovation-projects.org<br>Internet Source | <1 % |
| 22 | Abbass Ballit, Tien-Tuan Dao. "Recurrent neural network to predict hyperelastic constitutive behaviors of the skeletal muscle", Medical & Biological Engineering & Computing, 2022<br>Publication | <1 % |

| | | |
|---|---|---|
| 23 | Oluwatobi Adeleke, Sina Karimzadeh, Tien-Chien Jen. "Machine Learning-Based Modelling in Atomic Layer Deposition Processes", CRC Press, 2023<br>Publication | <1 % |
| 24 | Yichuan Xu, Vlado Keselj. "Stock Prediction using Deep Learning and Sentiment Analysis", 2019 IEEE International Conference on Big Data (Big Data), 2019<br>Publication | <1 % |
| 25 | journals.utm.my<br>Internet Source | <1 % |
| 26 | Submitted to Sheffield Hallam University<br>Student Paper | <1 % |
| 27 | Submitted to City University<br>Student Paper | <1 % |
| 28 | Submitted to Quest International University Perak<br>Student Paper | <1 % |
| 29 | cupdf.com<br>Internet Source | <1 % |
| 30 | Submitted to Queensland University of Technology<br>Student Paper | <1 % |
| 31 | Submitted to The University of the West of Scotland | <1 % |

| | | |
|---|---|---|
| | Student Paper | |
| 32 | anyflip.com<br>Internet Source | <1% |
| 33 | journals.orclever.com<br>Internet Source | <1% |
| 34 | mdpi-res.com<br>Internet Source | <1% |
| 35 | www.digitalsilk.com<br>Internet Source | <1% |
| 36 | Submitted to Universiti Teknologi Petronas<br>Student Paper | <1% |

| | | | |
|---|---|---|---|
| Exclude quotes | On | Exclude matches | < 10 words |
| Exclude bibliography | On | | |

| |
|---|
| **Form Title: Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)** |

| | | | |
|---|---|---|---|
| Form Number: FM-IAD-005 | Rev No.: 0 | Effective Date: 01/10/2013 | Page No.: 1of 1 |

**FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY**

| | |
|---|---|
| **Full Name(s) of Candidate(s)** | Gan Yu Nyuk |
| **ID Number(s)** | 21ACB06785 |
| **Programme / Course** | Bachelor of Computer Science (Honours) |
| **Title of Final Year Project** | Automation in Data Analytics using Artificial Intelligence and Robotic Process Automation |

| **Similarity** | **Supervisor's Comments (Compulsory if parameters of originality exceed the limits approved by UTAR)** |
|---|---|
| **Overall similarity index:__10__ %**<br><br>**Similarity by source**<br>Internet Sources: _8__ %<br>Publications: _3__ %<br>Student Papers: _5__ % | Percentage is within acceptance range. |
| **Number of individual sources listed** of more than 3% similarity: _0__ | Ok |

**Parameters of originality required, and limits approved by UTAR are as Follows:**

**(i) Overall similarity index is 20% and below, and**
**(ii) Matching of individual sources listed must be less than 3% each, and**
**(iii) Matching texts in continuous block must not exceed 8 words**

***Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.***

Note: Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

***Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.***

____________________________
Signature of Supervisor

Name: _Ts Dr Ooi Joo On________

Date: __13/9/2024________

____________________________
Signature of Co-Supervisor

Name: __________________

Date: __________________

# FYP 2 CHECKLIST

## UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY (KAMPAR CAMPUS)

**CHECKLIST FOR FYP 2 THESIS SUBMISSION**

| | |
|---|---|
| Student ID | 21ACB06785 |
| Student Name | Gan Yu Nyuk |
| Supervisor Name | Ts Dr Ooi Joo On |

| TICK (√) | DOCUMENT ITEMS<br>Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item. |
|---|---|
| √ | Title Page |
| √ | Signed Report Status Declaration Form |
| √ | Signed FYP Thesis Submission Form |
| √ | Signed form of the Declaration of Originality |
| √ | Acknowledgement |
| √ | Abstract |
| √ | Table of Contents |
| √ | List of Figures (if applicable) |
| √ | List of Tables (if applicable) |
| X | List of Symbols (if applicable) |
| √ | List of Abbreviations (if applicable) |
| √ | Chapters / Content |
| √ | Bibliography (or References) |
| √ | All references in bibliography are cited in the thesis, especially in the chapter of literature review |
| √ | Appendices (if applicable) |
| √ | Weekly Log |
| √ | Poster |
| √ | Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005) |
| √ | I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report. |

*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

____________________

(Signature of Student)

Date: 13/9/2024