**IMPLEMENTATION OF STOCK MARKET PREDICTION USING DEEP LEARNING CASE OF FUSION FOREX.**

**PETER KARIUKI THUKU**

**SCT221-0767/2021**

A Project Proposal Prepared And Submitted To The Faculty Of Computing And Information Technology, Jomo Kenyatta University Of Agriculture And Technology In Partial Fulfilment For The Award Of Bachelor Of Science In Information Technology.

**2024**

# DECLARATION

This proposal project is my original work and has not been presented for a

Degree in any other University

Student Name: PETER KARIUKI THUKU

…………………. …………………

Signature Date

This proposal project has been submitted for examination with my approval

As University Supervisor

Supervisor Name: MR PHILLIP OYIER

……………… ……………….

Signature Date

# ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to Almighty God for giving me a graceful bearing till this time of my life. Secondly, I thank my family for support in the process of carrying out my project. I am also grateful to my project supervisor, MR Phillip Oyier for guiding and helping me to carry out my project and acquire several skills that will help me in the project areas of related tasks.

# ABSTRACT

Finances represent one of the key requirements to perform any useful activity for humanity. Financial markets, e.g., stock markets, forex provide the opportunity to anyone to invest and generate finances. However, to reap maximum benefits from these financial markets, effective decision making is required to identify the trade directions, e.g., going long/short by analyzing all the influential factors, e.g., price action, economic policies, and supply/demand estimation, in a timely manner. In the dynamic realm of financial investments, it is often challenging for investors to accurately predict stock market fluctuations, leading to potential financial losses. Developing a web-based application project to facilitate accurate prediction and analysis of stock market trends through the implementation of machine learning algorithm., Stock prediction aims to provide investors with timely, reliable, and data-driven insights into the performance of various stocks, allowing them to make informed investment decisions. By leveraging advanced machine learning and data analysis techniques, Stock prediction offers users real-time access to comprehensive stock market data, trend predictions, and actionable insights. This platform effectively minimizes the need for extensive manual analysis and costly consultations with financial advisors, empowering users to make informed investment choices independently. With its user-friendly interface and customizable features, Stock predictor caters to both novice and experienced investors, equipping them with the tools needed to navigate the complexities of the stock market.

TABLE OF CONTENTS

[DECLARATION ii](#_Toc175858025)

[ACKNOWLEDGEMENT iii](#_Toc175858026)

[ABSTRACT iv](#_Toc175858027)

[LIST OF ACCRONYMS x](#_Toc175858028)

[CHAPTER ONE 1](#_Toc175858029)

[1.0 INTRODUCTION. 1](#_Toc175858030)

[1.1 BACKGROUND. 2](#_Toc175858031)

[1.2PROBLEM STATEMENT. 3](#_Toc175858032)

[1.3 PROPOSED SOLUTION 4](#_Toc175858033)

[1.4 OBJECTIVES 5](#_Toc175858034)

[1.5 RESEARCH QUESTIONS. 5](#_Toc175858035)

[CHAPTER TWO 10](#_Toc175858041)

[LITERATURE REVIEW 10](#_Toc175858042)

[2.1 introduction 10](#_Toc175858043)

[2.2 THEORETICAL REVIEW 10](#_Toc175858044)

[2.3Application of research area to the problem domain 14](#_Toc175858049)

[2.4integration and architecture 15](#_Toc175858050)

[2.5 RESEARCH GAPS 16](#_Toc175858052)

[CHAPTER THREE 17](#_Toc175858053)

[SYSTEM ANALYSIS AND DESIGN 17](#_Toc175858054)

[3.1 Introduction 17](#_Toc175858055)

[3.2 System Development Methodology 17](#_Toc175858056)

[3.3 Feasibility Study 19](#_Toc175858057)

[3.3.1Economic Feasibility 19](#_Toc175858058)

[3.3.2 Technical feasibility 20](#_Toc175858059)

[3.3.3 Operational Feasibility 20](#_Toc175858060)

[3.3.4 Legal and Ethical Feasibility 21](#_Toc175858061)

[3.3.5 Schedule Feasibility 21](#_Toc175858062)

[3.4 Requirement Elicitation 21](#_Toc175858063)

[3.5 Data Analysis 24](#_Toc175858064)

[3.6 SYSTEM SPECIFICATIONS 26](#_Toc175858065)

[3.7 Requirements Analysis and Modelling 27](#_Toc175858068)

[3.8 Logical design 31](#_Toc175858070)

[3.9 Physical design 33](#_Toc175858072)

[CHAPTER FOUR 37](#_Toc175858075)

[SYSTEM IMPLEMENTATION AND TESTING, CONCLUSIONS AND RECOMMENDATIONS 37](#_Toc175858076)

[4.0 Introduction 37](#_Toc175858077)

[4.1 Environment and Tools 37](#_Toc175858078)

[4.1.1 System Code Generation 38](#_Toc175858079)

[4.1.2 Snapshots of the system 41](#_Toc175858080)

[4.2 TESTING 43](#_Toc175858081)

[4.2.1 Unit Testing 43](#_Toc175858083)

[4.2.2. GUI Testing 44](#_Toc175858084)

[4.2.3. Integration Testing 45](#_Toc175858085)

[4.2.4. Acceptance Testing 46](#_Toc175858086)

[4.2.5. Testing predictions 46](#_Toc175858087)

[REFERENCES 48](#_Toc175858089)

[QUESTIONAIRE 50](#_Toc175858090)

[APPENDIX: 57](#_Toc175858091)

[PROJECT SCHEDULE 57](#_Toc175858092)

[GANT CHART 58](#_Toc175858093)

[BUDGET 58](#_Toc175858094)

[APPENDIX 59](#_Toc175858095)

**LIST OF FIGURES**

[Figure 1 Activity diagram 28](#_Toc175893584)

[Figure 2 Use case Diagram 29](#_Toc175893585)

[Figure 3 Sequence diagram 30](#_Toc175893586)

[Figure 4 logical Design 31](#_Toc175893587)

[Figure 5 System Architecture 32](#_Toc175893588)

[Figure 6 User interface home page 33](#_Toc175893589)

[Figure 7 Prediction page 34](#_Toc175893590)

[Figure 8 system mockup 35](#_Toc175893591)

[Figure 9 predicted trend 36](#_Toc175893592)

[Figure 10 Calculated moving averages form 36](#_Toc175893593)

[Figure 11 Stock data fetching 38](#_Toc175893594)

[Figure 12 Fetching stock news and fundamental data 39](#_Toc175893595)

[Figure 13 Model training 40](#_Toc175893596)

[Figure 14 Successful epoch's 40](#_Toc175893597)

[Figure 15 system snapshot home page 41](#_Toc175893598)

[Figure 16 snapshot for stock predictor 42](#_Toc175893599)

**LIST OF TABLES**

[Table 1 Project schedule 57](#_Toc175894025)

[Table 2 Gant chart 58](#_Toc175894026)

# LIST OF ACCRONYMS

AI Artificial intelligence

ANN Artificial Neural Networks

ARIMA Autoregressive Integrated Moving Average

DNN Deep neural networks

EMA Exponential Moving Average

GARCH Generalized Autoregressive Conditional Heteroscedasticity

LSTM Long Short- Term Memory

MA Moving Average

NLP Natural language Processing

RNN Recurrent neural networks

SMA Simple moving average

SVR Support vector regression

# CHAPTER ONE

## INTRODUCTION.

Stock trading, the heartbeat of the financial world, offers a pathway to prosperity, but it's riddled with uncertainty. It's a marketplace where investors buy and sell shares of companies, hoping to secure financial gains. In this world, timing and information are paramount.

The stock trend prediction system will be your compass in this dynamic landscape. It's built to decipher the intricate dance of stock prices. Leveraging historical data, advanced analytics, and real-time news sentiment analysis, the app equips you with reliable predictions for stock trends.

Whether you're a seasoned trader or a newcomer, the user-friendly platform ensures you stay informed. It's time to make your stock trading decisions confidently and embark on a path to financial success with the stock trend prediction System.

My client fusion forex is a trading firm that mainly trades in place of its clients using the various trading approaches and also providing mentorship to its clients. The institution offers workshops led by experienced traders and investors helping people gain insights into specific trading techniques.

Top of Form

## 1.1 BACKGROUND.

Stock trading is the pulse of global financial markets. It is the process of buying and selling ownership shares (stocks) of publicly-traded companies. These transactions occur on stock exchanges, where investors, ranging from individuals to institutions, engage in a perpetual battle of supply and demand. The objective of stock trading is to capitalize on price fluctuations, either for short-term gains or long-term investment, with the ultimate aim of growing wealth.

However, stock trading is not merely a straightforward exchange of shares. It is a complex arena influenced by an array of factors, including economic conditions, company performance, geopolitical events, and investor sentiment. The need to anticipate these multifaceted dynamics and make ill-informed decisions is paramount, as stock trading inherently carries risk and potential rewards.

The purpose of this project is to bridge the gap between the intricacies of stock trading and the everyday investor.. Unfortunately, the tools and resources available often fall short in offering the accuracy, accessibility, and user-friendliness necessary for confident decision-making.

This project introduces a pioneering stock trend prediction app as a solution to this critical issue. Drawing on the historical performance of stocks, advanced technical indicators, and real-time analysis of financial news sentiment, the app endeavors to provide users with precise forecasts of stock trends. The ultimate goal is to empower investors and traders, irrespective of their level of experience, with the knowledge and insights required to make sound financial decisions.

Economic conditions, company performance, geopolitical events, and investor sentiment all contribute to the volatility of stock prices, making it challenging for investors to accuratelypredict trends and capitalize on market movements. Despite the availability of various tools and resources, many fall short in providing the precision, accessibility, and user-friendliness necessary for effective decision-making. This tend to be great challenges that the company fusion forex is facing.

## 1.2PROBLEM STATEMENT.

With reference to my client fusion forex in the ever-fluctuating world of financial markets, individuals, investors, and traders face a difficult challenge when it comes to making timely and well-informed decisions regarding stock investments. This challenge is primarily rooted in the inherent complexity and uncertainty surrounding stock price movements. Current stock trend prediction tools often fall short in terms of delivering the accuracy, accessibility, and user-friendliness that users require to navigate this complex landscape.

One of the pressing issues is the limited accuracy of existing prediction models. While various forecasting techniques have been employed, there remains a substantial gap between the predictions offered by these tools and the actual outcomes in the stock market. The financial well-being of individuals, as well as the profitability of investments, is highly dependent on the precision of these predictions. Inaccurate forecasts can lead to significant financial losses and missed opportunities for investors, ultimately affecting their confidence in the market.

Accessibility is another critical concern. Many stock prediction tools are overly complex and may require advanced knowledge of financial markets and data analysis. This complexity acts as a barrier, preventing a wider audience from taking advantage of these tools. Investors, especially those who are relatively new to the stock market, may be overwhelmed by intricate jargon and technicalities, leading to underutilization of these tools.

Furthermore, user-friendliness plays a crucial role in the adoption of such apps. A stock trend prediction tool must be intuitive and easy to use, regardless of a user's level of experience in stock trading. Without a user-friendly interface, individuals may find it challenging to leverage the app effectively and make timely investment decisions.

Top of Form

## 1.3 PROPOSED SOLUTION

I propose developing of a deep learning-based stock market prediction system.

The proposed system that aims to leverage the power of machine learning algorithms to analyze vast amounts of historical market data, identify underlying patterns, and make predictions about future market movements. By utilizing machine learning, the system can effectively capture complex relationships and nonlinear dynamics that may exist in the market data, enabling more accurate and reliable predictions.

Machine learning algorithms excel at recognizing patterns and extracting insights from large datasets, making them well-suited for stock market prediction. These algorithms can learn from historical market data, adapt to changing market conditions, and identify subtle indicators and signals that may influence stock prices.

By utilizing historical market data, the system can identify patterns and correlations between various market factors, such as stock prices, trading volumes, economic indicators, and news sentiment. It can uncover hidden relationships and dependencies that may not be apparent to human analysts, providing a comprehensive and data-driven approach to stock market forecasting.

The machine learning-based approach also offers the advantage of continuous learning and adaptation. As new market data becomes available, the system can update its models and incorporate the latest information, ensuring that predictions remain relevant and accurate in real-time or near-real-time scenarios. This adaptability allows the system to respond to changing market dynamics and adjust its predictions accordingly. The required data will be generated from Yahoo finance.

## **1.4 OBJECTIVES**

**General objective**

To develop an advanced and reliable algorithmic model that can accurately forecast the future movement and trends of stock prices in financial markets

Specific objective

* Design and implement a user-friendly interface using Streamlit that caters to both novice and experienced users.
* To conduct a comprehensive research on machine and deep learning algorithms on incorporation to stock prediction.
* To test the developed prototype in real world scenario.

## **1.5 RESEARCH QUESTIONS**.

1. What are the methods used for stock prediction?
2. What are the methods used to develop a stock prediction prototype?
3. What are the methods of testing that will be used?

## PROJECT JUSTIFICATION.

This project driven by a pressing need for accurate and reliable tools to navigate the dynamic world of finance and investment. It seeks to empower individuals and institutions with the means to make informed financial decisions. The stock market, as a principal avenue for wealth creation and financial security, can greatly benefit from the provision of accurate predictions, which can have a profound impact on investment outcomes.

Moreover, the inherent risk associated with stock market investments underscores the importance of this endeavor. By enabling effective risk assessment and management, accurate stock market predictions help investors safeguard their financial assets and make prudent choices, reducing potential financial losses.

Accessibility to stock markets, often considered a complex and daunting realm, is another key rationale behind this project. While experienced traders may navigate these markets with confidence, newcomers are frequently deterred by the intricacies. Therefore, the development of a user-friendly stock prediction app serves as a means to level the playing field, making stock trading accessible to a wider audience.

Additionally, the economic implications of stock markets extend far beyond individual investments. Accurate stock predictions can provide valuable insights to policymakers, financial institutions, and businesses, aiding in decision-making processes with broader economic implications.

In an era shaped by data and technology, this project serves as a prime example of the power of data-driven decision-making. It showcases how advanced algorithms and machine learning can be harnessed to extract valuable insights from financial data, highlighting the practical applications of data science in the financial sector.

Lastly, ethical responsibility is a cornerstone of this project. It upholds ethical standards by ensuring that users are ill-informed and aware of the risks associated with stock trading, promoting transparency and ethical practices in the financial sector.

## Scope

This project aims at predicting public listed companies‟ stock price. This means that the scope of study revolves around PLC that already has several years of historical data. The project will try to find several PLCs‟ data available in the internet and uses it to feed the data mining machine to give the desired results.

This project focuses on American stock market because the nature and systems used in each market is different from one region to another such as the financial terms, the American rules and regulations and the indexes used by the NYSE is different with other stock markets.

The existing idea was to get the data from the NYSE. However, due to certain difficulties, the data cannot be obtained. Therefore, the next alternative would be finding free source data from the internet. Sources will include yahoo finance and google finance.

The free source data must comprise of three years back-dated data for the purpose of the research. Since there are a lot of websites that provide the free data, the project will be scoped down into finding these websites and extract the given data for the system.

The three years data is to be fed into the machine to learn and recognize the patterns of the price movements and to predict the movements in the future. Besides than learning, the data will also be used as a test and validation data in order to ensure the accuracy and reliability of the predicted stock prices.

The nature of the stock data that would be used within this project would be the price fluctuations rather than the volume and other quotations that would usually associate with a stock data. This is because the stock price movement is already adequate to find the historical patterns and to predict the future movement of the stock prices.

The scope of this study would further assist in creating the accuracy and reliability that the system proposes. The training model, validation model and testing model each having one year of historical stock data to train, validate and test the patterns and predictions would be integral in order to train the system to be as accurate as possible.

## FEASIBILITY OF THE PROJECT

### OPERATIONAL FEASIBILITY

The proposed system will help in providing a platform to predict the trend of the stock prices for the usage of investors and future investors. The system also suits the tolerance of risk and the need of investors to find the desired trend of stocks that the investors would want to spend in. this would immediately and significantly help investors to get the desired return on investment.

In order to cater the different background and levels of technological acceptance by the users, the system will issue internal training and user guide manual to the users. It also would provide internal troubleshooting advices towards the users of the system. This would greatly assist the investors in order to efficiently use the system.

It is believed that such suggestions being proposed into the system would effectively addresses the problem that investors currently have and to ease the work of financial regulators in the nation. Hence, it is feasible in terms of operation to create the system.

## SYSTEM REQUIREMENTS

**Hardware Requirements**

1. A Laptop or Desktop with the following specifications:
2. Core i7 Processor (4 Cores Minimum)
3. 16GB minimum RAM
4. 1TB Hard-Drive (SSD recommended for faster processing)
5. Windows/Linux/Mac Operating System

**Software Requirements**

The following software and tools will be required for development and deployment of the stock prediction system:

1. Python 3.8+: Programming language used for implementing the system's logic and algorithms.
2. Jupyter Notebook/Google Colab: For experimenting with LSTM model and conducting data analysis.
3. Pandas, NumPy, Matplotlib, Seaborn: Python libraries for data manipulation, numerical computing, and data visualization.
4. TensorFlow/Keras: Deep learning framework for building and training LSTM model for stock prediction.
5. Streamlit: Python library for building interactive web applications. It will be used to create a user-friendly interface for users to interact with the prediction system.
6. Yahoo Finance and Alpha Vantage APIs: For fetching historical stock data and real-time market data.
7. News API: For retrieving latest news related to stocks and financial markets.

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 introduction

Stock market is basically nonlinear in nature and the research on stock market is one of the most important issues in recent years. People invest in stock market based on some prediction. For predict, the stock market prices people search such methods and tools which will increase their profits, while minimize their risks. Prediction plays a very important role in stock market business which is very complicated and challenging process. Employing traditional methods like fundamental and technical analysis may not ensure the reliability of the prediction. To make predictions regression analysis is used mostly. Traditional methods, such as fundamental and technical analysis, do not always guarantee reliable predictions, often leading to the use of statistical methods like regression analysis (Fama, 1970; Malkiel, 2003)

## 2.2 THEORETICAL REVIEW

### Fundamental analysis

This focuses on the financial condition and operating results of a specific company. It means that the analysis is on the actual financial performance of the company and the relation of that with the current stock price. It would be the indicator whether a stock price is undervalued or overvalued in which the market would soon adjust itself towards the current fundamental valuation of the company. In essence, fundamental analysis assists investors to formulate expectations on the future performance of the company and its stocks (Gitman, Joehnk, & Smart, 2011).

Therefore, we can conclude that fundamental analysis is the evaluation of the company’s intrinsic value in which the market would adjust the stock prices according to the intrinsic value. Thus, giving a solid ground to predict the stocks’ future prices. The situation of fundamental analysis is as follows: if the stock price is higher than the company’s intrinsic value (overvalued), then the market would eventually strive to lower down the stock price to create an equilibrium of the stock price and its intrinsic value and vice versa. With fundamental analysis, stockbrokers can predict the future price according to its intrinsic value.

### Technical analysis

Technical analysis on the other hand is to analyze the market behavior rather than the financial condition of the company. It involves studying the various market statistics such as volume of trading, amount of short selling or buy/sell patterns of the stocks. It would try to evaluate the position of its current price to predict the future movement of the stock price. Based on historical data, it would try to recognize the pattern and trend as a basis to predict the future stock price (Gitman, Joehnk, & Smart, 2011).

Base on the excerpt, it can be concluded that the trend of stock price is that it would climb until it reaches it peaks and begin to fall. It would then fall until it reaches the trough where it would begin climbing up again. Hence, the technical analysis would try to evaluate the current position of the stock price. In order to do so, it must use the historical data to plot its current trend and base on that evaluation, try to predict its future movement. Technical analysis is a feasible yet effective way to predict the future stock price movement.

### TIME SERIES

Time series approach means that historical events do have a pattern that would be applicable to predict the future events. Causal approach would usually find a reason behind the historical patterns and uses it to generate the forecast. Judgmental approach is the gathering the knowledge and opinions from experts to predict the future events and lastly, experimental approach would mean experiments are conducted to become the data to rely upon for forecasting. However, time-series forecasting is the most widely used due to its model that only take into consideration that historical events are bound to happen again given the same factors of the events

One widely studied theoretical framework is the Efficient Market Hypothesis (EMH), which posits that stock prices reflect all available information, making it impossible to consistently outperform the market using historical data. However, many studies have challenged the EMH and proposed alternative theories, such as behavioral finance, which suggests that investor psychology and irrational behavior can lead to predictable patterns in stock prices. The Efficient Market Hypothesis (EMH) suggests that stock prices fully reflect all available information, making it impossible to consistently outperform the market using historical data (Fama, 1970). However, behavioral finance challenges EMH by highlighting how psychological biases and irrational behavior can lead to predictable patterns in stock prices (Kahneman & Tversky, 1979; Thaler, 1985).

### DEEP LEARNING

. Algorithms, particularly machine learning algorithms like stacked long short-term memory (LSTM) networks, process the data, learn patterns, and make predictions. LSTM, a type of recurrent neural network (RNN), is specifically designed to capture sequential dependencies in the data. The training phase involves feeding historical data to the model, allowing it to learn and adjust its parameters, becoming familiar with patterns and relationships in the data.

Within the realm of stock market prediction, distinct theoretical divisions and approaches exist. Technical analysis centers on historical price and volume data to identify patterns and trends, positing that historical price movements can predict future price changes. In contrast, fundamental analysis focuses on analyzing a company's financial health, earnings reports, economic indicators, and news events to assess the intrinsic value of a stock. A broader division lies between machine learning and traditional statistical methods. Machine learning techniques, including deep learning models like lstm, are gaining prominence for their capacity to capture intricate patterns and dependencies in the data. Traditional statistical methods, including regression models and time series analysis, have a long history in stock market prediction. Lastly, the distinction between quantitative approaches, which rely on numeric data and quantitative models, and qualitative approaches, which incorporate non-numeric data like news sentiment and qualitative factors, is also noteworthy.

When comparing LSTM models to Feedforward Neural Networks (FNNs), the distinction lies in their treatment of temporal dependencies in stock price data. For instance, if a stock has been steadily increasing in price over the past few days, there is a temporal dependency suggesting that it may continue to rise in the near future. LSTM models are particularly adept at capturing these temporal dependencies because they are designed to process sequential data and retain information over time LSTMs leverage recurrent connections and memory cells to capture sequential patterns over time, making them proficient in grasping long-term dependencies and intricate trends inherent in financial data. Conversely, FNNs process data in a single pass, lacking the ability to consider temporal context. While suitable for simpler tasks, FNNs may falter when tasked with comprehensively capturing the sequential nature of stock prices, hindering their efficiency in forecasting long-term trends.

In the realm of comparing LSTMs to Convolutional Neural Networks (CNNs), the focus shifts to their respective strengths in handling sequential data. LSTMs are purpose-built to apprehend temporal dependencies, rendering them well-suited for time series data analysis, particularly in stock price prediction where understanding patterns over time is paramount. On the other hand, CNNs, primarily designed for image processing tasks, can be repurposed for sequential data analysis by treating data as images. Despite their proficiency in capturing local patterns, CNNs may struggle to capture long-term dependencies as effectively as LSTMs, particularly in tasks reliant on temporal context for accurate predictions.

Machine learning and deep learning techniques, including LSTMs and CNNs, offer advanced capabilities for capturing intricate patterns and dependencies in stock data, surpassing traditional statistical models in accuracy (Zhuge et al., 2017; Bao, Yue, & Rao, 2017). The ability of LSTM models to understand long-term dependencies and complex trends makes them particularly effective for financial forecasting (Chen et al., 2015; Nelson, Pereira, & de Oliveira, 2017).

## 2.3Application of research area to the problem domain

Several researchers implemented their work to provide accurate solutions to this dynamic problem and have proposed various methods for predicting the stock market. Jayanth Balaji performed a deep learning method to predict a company's stock price using 14 different deep learning methods. Similar work is implemented using Artificial Neural Networks (ANN) by Tsong Wuu Lin his work tried to maximize the profitability using this model. Autoregressive models are powerful models for predicting the stock market, they give a strong insight on time series analysis and make very accurate predictions. Sentiment analysis is also one of the strong ways to predict the stock market. Social media analytics plays a vital role in sentiment analysis. ARIMA model helps in sentiment analysis and predicting time series data. Sentiment analysis can also be implemented by using deep learning models like CNN and LSTM.

LSTMs and Transformer models are used to generate context from sequences which are long. The efficiency of the LSTM models and Transformer models are significantly reduced when the sequence length is reduced. [26–29]. The sequence length used in the CNN model is considerably lesser than the ones used for LSTM and transformer models.

One common approach is the use of machine learning algorithms, such as support vector machines (SVM), random forests, and neural networks, to train predictive models. These models leverage historical data to identify patterns and relationships that can be used to make future predictions. Researchers have also explored ensemble methods, combining multiple models to improve prediction accuracy.

Additionally, natural language processing (NLP) techniques have been applied to analyze textual data from news articles, social media, and corporate announcements. Sentiment analysis is used to gauge the overall market sentiment, and news sentiment is incorporated into predictive models to capture the impact of news events on stock prices.

A typical architecture includes data collection, preprocessing, feature extraction, model training, and prediction generation (Tsai & Hsiao, 2010). Feature extraction involves creating meaningful features from raw data, while model training optimizes algorithms to improve accuracy (Chong, Han, & Park, 2017). Ensemble methods and machine learning models like LSTM and CNN are commonly used to enhance prediction accuracy (Huang et al., 2005; Tsai et al., 2011).

## 2.4integration and architecture

The integration and architecture of stock market prediction systems involve the combination of various data sources, predictive models, and decision-making components. A typical architecture consists of data collection and preprocessing, feature extraction, model training, and prediction generation.

Data collection involves gathering relevant data from multiple sources, such as stock exchanges, financial databases, news APIs, and social media platforms. Preprocessing techniques, such as data cleaning, normalization, and feature engineering, are applied to ensure the data is suitable for analysis.

Feature extraction involves transforming raw data into meaningful features that capture relevant information for prediction. This step may involve technical indicators, statistical measures, sentiment scores, and other derived features.

Model training involves selecting appropriate machine learning algorithms and training the models using historical data. Various techniques, such as cross-validation and hyper parameter tuning, are employed to optimize model performance. Ensemble methods may also be used to combine multiple models for improved prediction accuracy.

Prediction generation involves applying the trained models to new data to generate forecasts and make investment decisions. The output can be in the form of price movements, buy/sell signals, or risk assessments, depending on the specific objectives of the system.

## SUMMARY

In summary, the literature on stock market prediction systems demonstrates a wide range of theoretical concepts, methodologies, and techniques. Researchers have explored efficient market theory, behavioral finance, technical analysis, fundamental analysis, machine learning, and artificial intelligence to develop predictive models. The application of these approaches to the problem domain involves integrating multiple data sources, employing various predictive models, and generating actionable insights for investors and traders. By incorporating deep learning models such as LSTM it will make analysis easier from short term to long term with a higher efficiency and minimal error rate. By incorporating deep learning models such as LSTM, predictions become more accurate and adaptable, addressing both short-term and long-term trends (Zhuge et al., 2017; Bao et al., 2017).

## 2.5 RESEARCH GAPS

Despite significant advancements, several research gaps remain in the field of stock market prediction systems. Some potential areas for future research include:

a) Incorporating real-time data: Most existing systems rely on historical data, but incorporating real-time data could improve prediction accuracy and responsiveness to market dynamics.

b) Handling market anomalies and rare events: Stock markets are prone to anomalies and rare events, such as financial crises or unexpected news events. Developing models that can effectively handle such events and adapt to changing market conditions is an important area for further investigation.

c) Evaluating prediction uncertainty: Providing measures of prediction uncertainty can help investors make informed decisions and manage risk effectively. Research on quantifying and communicating prediction uncertainty is needed.

d) Data Quality: Inaccurate, incomplete, or unstructured data sources can limit the effectiveness of prediction algorithms.

# CHAPTER THREE

# SYSTEM ANALYSIS AND DESIGN

## 3.1 Introduction

This chapter covered system methodology, requirement elicitation which helped in data collection of user and system requirements, and feasibility study all this which clearly gave me the overall image of how the system should look and operate. System analysis projects are initiated for three reasons: problems, opportunities, and directives. It also covered functional and non-functional system requirements, system logical design using Unified Modelling Language (UML) diagrams, and system physical design, including database design and user interface design.

## 3.2 System Development Methodology

In this project, we employ the CRISP-DM methodology to develop a system for predicting stock prices using machine learning techniques. The CRISP-DM methodology is composed of six distinct phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. Each phase is crucial to ensure the successful development and implementation of the Stock Market Prediction System.

**1. Business Understanding**

- Objective: The primary objective of this phase is to understand the project goals and requirements from a business perspective and then convert these into a data mining problem definition.

- Application: For the Stock Market Prediction System, the goal is to develop a predictive model that can forecast stock prices based on historical data. This involves identifying key business questions, understanding the financial market's needs, and determining the scope of the project.

**2. Data Understanding:**

- Objective: This phase focuses on initial data collection and familiarization. It involves identifying data quality issues and discovering initial insights into the data.

- Application: We collect historical stock data from sources like Yahoo Finance. The data includes stock opening price, closing price, highest price, lowest price, and trading volume. We analyze this data to understand its structure and quality, identifying any anomalies or missing values.

**3. Data Preparation:**

- Objective: This phase involves selecting, cleaning, and transforming data to make it suitable for modeling.

- Application: The data preparation process includes handling missing values, normalizing data, and extracting relevant features. For instance, we might create new features such as moving averages or stock price trends to improve the model's predictive power.

**4. Modeling:**

- Objective: In this phase, various modeling techniques are selected and applied to the prepared data. Parameters are adjusted to optimize model performance.

- Application: We experiment with different machine learning algorithms, such as LSTM (Long Short-Term Memory) networks, to model the stock price data. The chosen models are trained on historical data, and their performance is assessed using metrics like accuracy.

**5. Evaluation:**

- Objective: The evaluation phase aims to assess the model's quality and effectiveness in meeting the business objectives.

- Application: The models are evaluated on a validation dataset to ensure they generalize well to unseen data. We compare the models' performance and select the best one for deployment. Additionally, we review the process to ensure all business objectives are met.

**6. Deployment:**

- Objective: This phase involves deploying the model into a real-world environment where it can be used to make predictions.

- Application: The final model is integrated into the Stock Market Prediction System, allowing users to input new data and receive stock price predictions. We also set up monitoring mechanisms to track the model's performance over time and update it as necessary

By following the CRISP-DM methodology, the Stock Market Prediction System is developed in a structured and systematic manner, ensuring all aspects of the project are thoroughly addressed and the final system meets the business requirements effectively.

## 3.3 Feasibility Study

A feasibility study is a crucial step to assess the practicality and viability of the proposed project. It involves evaluating the economic, technical, and other critical aspects to determine whether the project can be successfully implemented.

## 3.3.1Economic Feasibility

The cost estimates for developing the system are within budget while the benefits include, improved product quality leading to increased profits and customer satisfaction, potential cost savings through automated quality assessment and enhanced competitiveness in the stock market.

In cost Analysis, the implementation cost for developing the system, acquiring necessary hardware, software, and training, was analyzed. The budget allocation was found to be within the defined financial resources of Fusion forex limited. About the return on Investment (ROI), a conservative estimate was calculated, factoring in potential increased efficiency and reduced losses due to better quality control.

Return on Investment (ROI):

Total Costs:

Initial Investment (Development + Deployment): 40,000 KES + 40,000 KES = 80,000 KES

Annual Maintenance Costs: 22,000 KES

Net Annual Benefit:

Additional Trading Profits: 45,000 KES

Maintenance Costs: 22,000 KES

Net Annual Benefit: 45,000 KES - 22,000 KES = 23,000 KES

ROI:

ROI = (Net Annual Benefit / Total Costs) 100%

ROI = (23,000 KES / 80,000 KES) 100% ≈ 28.75% The ROI calculations suggested a favorable return within a reasonable timeframe, making the project economically feasible.

## 3.3.2 Technical feasibility

The technical infrastructure and technologies required for the stock prediction system were assessed. Assessing the availability of data sources needed for stock forecasting, such as historical stock prices, financial statements, economic indicators, news sentiment data, and alternative data sources. It also involved evaluating the viability of data collection, cleaning, and preprocessing methods. Considering the computational resources and infrastructure required to build and maintain the stock prediction system, including hardware, software, and data storage. It was determined that the necessary hardware and software could be readily obtained and integrated into the existing IT infrastructure of the organization. The system was also deemed scalable to handle a growing volume of stocks. Moreover, the system architecture allowed for future upgrades and integration with emerging technologies.

## 3.3.3 Operational Feasibility

The integration of the stock prediction system with existing workflow processes was evaluated. This involved assessing the practicality of integrating the stock prediction system into existing workflows and decision-making processes of target users. Evaluating user acceptance and usability considerations, such as the user interface design, ease of interpretation of predictions, and integration with trading platforms or investment management systems. The study found that the system could be seamlessly integrated into other organizations programs, ensuring minimal disruption to other existing operational flow. The system was seen as a valuable tool to streamline the evaluation process and enhance decision-making.

## 3.3.4 Legal and Ethical Feasibility

The system will comply with all legal and ethical standards regarding data privacy, intellectual property, and usage rights. The study ensured that the system and its operations complied with Investment and Financial Analysts Act related to data privacy, intellectual property, and usage of stock data. The project aims to contribute positively to society by potentially reducing the rate at which new stock investors lose their money due to lack of knowledge.

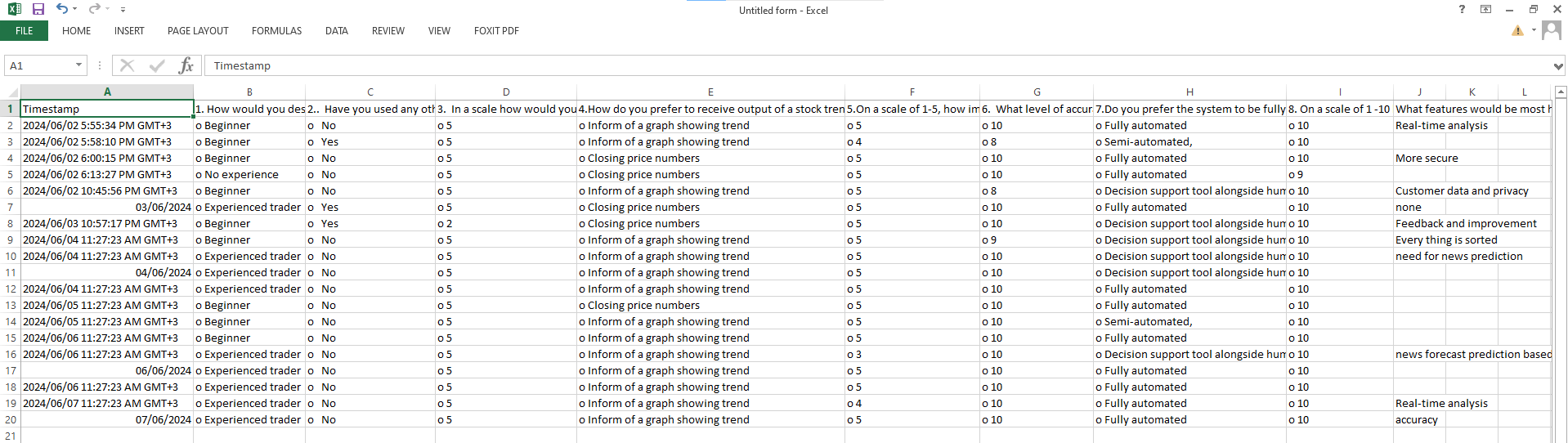
The system does not provide financial advice but generates predictions based on historical data, thus avoiding potential legal issues related to financial advisory services. Furthermore, the system will operate under the regulations of the Capital Markets Authority of Kenya, ensuring compliance with all relevant laws and regulations. (Authority).

## 3.3.5 Schedule Feasibility

A detailed project timeline was developed, taking into account the complexity of the development process and potential challenges. The study confirmed that the project could be completed within the stipulated timeframe.

1. **3.4 Requirement Elicitation**

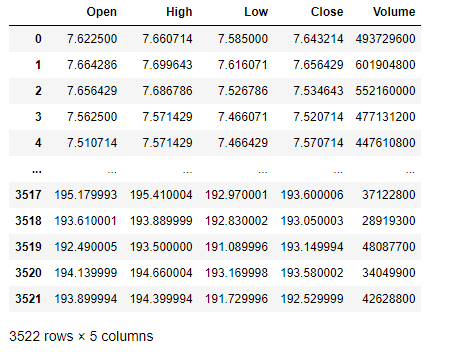
The primary data collection tool for this project was interviews, questionnaires. Data collection also involved getting data from secondary sources of data like online repositories. The questions were structured to be clear, concise, and directly related to the research objectives. The questionnaire was prepared using Google Forms. The data chosen was specifically of the highly traded stocks. This was done to ensure that the data obtained from secondary sources matched the primary data collected from the client. More so, this ensured that the model was trained on the same type of data but from different sources. Data collection was administered by getting views from the client ad also online repositories which is yahoo finance which provided the stock data of the current stock market.



**Yahoo finance**

Yahoo Finance is a comprehensive financial news and data platform provided by Yahoo, a multinational technology company. It offers a wide range of financial information, including stock quotes, financial news, analysis, portfolio management tools, and market data. Users can access real-time stock quotes, historical price charts, financial statements, analyst ratings, and other investment-related information for thousands of publicly traded companies worldwide. The data set for training the module was acquired from this repository.

**Stock dataset**



**Stock company and stock symbols as used in stock market**

|  |  |
| --- | --- |
| COMPANY | SYMBOL |
| Google | GOOGL |
| Apple | AAPL |
| Amazon | AMZN |
| Microsoft | MSFT |
| Tesla | TSLA |

We'll use this data to predict and forecast the worth of days into the future. Then we'll use the simplest Deep Learning Algorithm to predict the longer-term stock values in any company stock or other financial instrument traded on a financial exchange.

## 3.5 Data Analysis

The data was analyzed and interpreted using an Excel spreadsheet and the results displayed as illustrated below this section contains a continuation of the analysis of respondents concerning the non-functional requirements of the system as well as the interactivity of the software.

Sources: Historical stock data was primarily collected from Yahoo Finance, a reliable repository for financial information. The data included stock opening prices, closing prices, highest and lowest prices in a day, and trading volumes.

Dataset Size: The total amount of stock data gathered for this dataset was 3522 files which are grouped into stocks opening price named open, stock closing price as close highest reading Ina day as high, lowest reading in a day as low and the volume of the stock as volume.

Feature Extraction: New features such as moving averages, stock price trends, and volatility indices were found critical to be included in the design. Moving averages, for example, help smooth out price data to identify the direction of the trend.

Visualization: Various graphs and plots will be used to visualize the data. Line charts will likely be used to display stock price trends over time.

Trends Identification: By plotting the historical stock prices, key trends and patterns will be identified. The data might reveal periodic fluctuations corresponding to market cycles.

Correlation Analysis: Correlation analysis will be performed to identify the relationship between different stock parameters. For example, the relationship between trading volume and price movements will be examined to understand how trading activity influences stock prices

Model Selection: Various machine learning models will be considered for predicting stock prices. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are likely to be chosen due to their ability to capture temporal dependencies in time-series data.

Training and Validation: The models will be trained on historical data and validated using a separate validation dataset. Hyperparameter tuning will be performed to optimize model performance.

Questionnaire Results: The responses from the questionnaire will provide valuable insights into user preferences and requirements. For instance, users are expected to express a preference for receiving stock trends in graphical form rather than numeric output.

Experience and Expectations: Users' level of experience with stock markets and their expectations from the stock prediction system will be analyzed to tailor the system's features accordingly. The importance of user-friendly interfaces and high accuracy in predictions will be emphasized.

## 3.6 SYSTEM SPECIFICATIONS

System specification refers to a detailed description of the requirements, functionalities, and characteristics of a system. It outlined the specific features, behaviors, and constraints that the system should possess to meet the desired objectives and fulfill the needs of its users.

1. **Functional Requirements**

Functional requirements refer to what the system does. In this system they include;

1. Fetch historical stock data from yahoo finance
2. Preprocess the collected data by handling missing values and feature extraction
3. Train Deep learning models using the preprocessed historical data to predict future stock prices
4. Generate predictions for future stock prices based on the trained models.
5. Plot both stock data and predicted stock prices(trends)
6. Allow integration with external systems or APIs for accessing real-time data and incorporating it into the prediction models.
7. Display top ten stock news
8. **Non-Functional Requirements:**
9. Performance: The system should be able to handle large volumes of data efficiently and produce predictions within a reasonable timeframe.
10. Accuracy: The prediction models should strive for high accuracy in forecasting stock prices to provide reliable insights to users.
11. Scalability: The system should be scalable to accommodate increasing data volume and user demand without sacrificing performance.
12. Reliability: Ensure high reliability by minimizing system downtime and errors, providing failover mechanisms, and implementing data backup strategies.
13. Security: Implement robust security measures to protect sensitive financial data and user information from unauthorized access, breaches, or cyberattacks.
14. Usability: Design the system with a user-friendly interface and intuitive navigation to facilitate ease of use for users with varying levels of technical expertise.
15. Compatibility: Ensure compatibility with different platforms and devices to allow users to access the system from desktops, tablets, and smartphones.
16. Maintainability: Design the system with clean, modular code and documentation to facilitate easy maintenance, updates, and troubleshooting by developers.
17. Regulatory Compliance: Ensure compliance with relevant regulations and standards governing financial data handling and prediction systems.
18. Ethical Considerations: Incorporate ethical considerations into the system's design and operation, such as transparency in model assumptions, avoiding bias in predictions, and respecting user privacy.

## 3.7 Requirements Analysis and Modelling

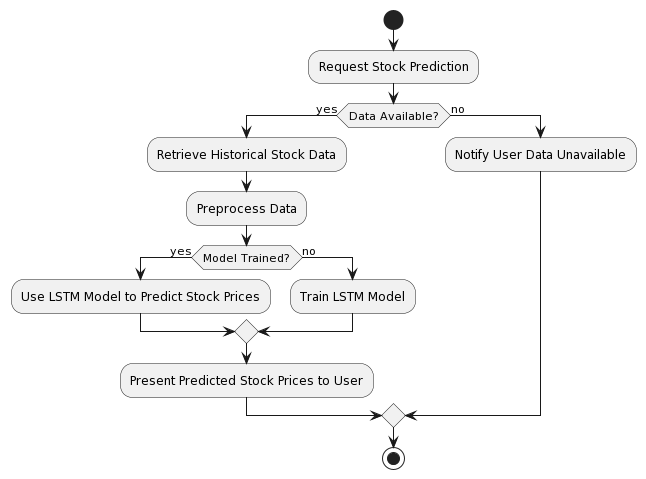
Requirement’s modeling is a technique used in software engineering and systems analysis create

visual representations or models that depict the various aspects of the system's requirements,

such as its functionality. This process of modeling involved using applications such as lucid char and plantuml web server.

**Activity diagram**

This activity diagram captures the main steps involved in the stock prediction process using an LSTM model and provides a clear flow of activities from user input to output.



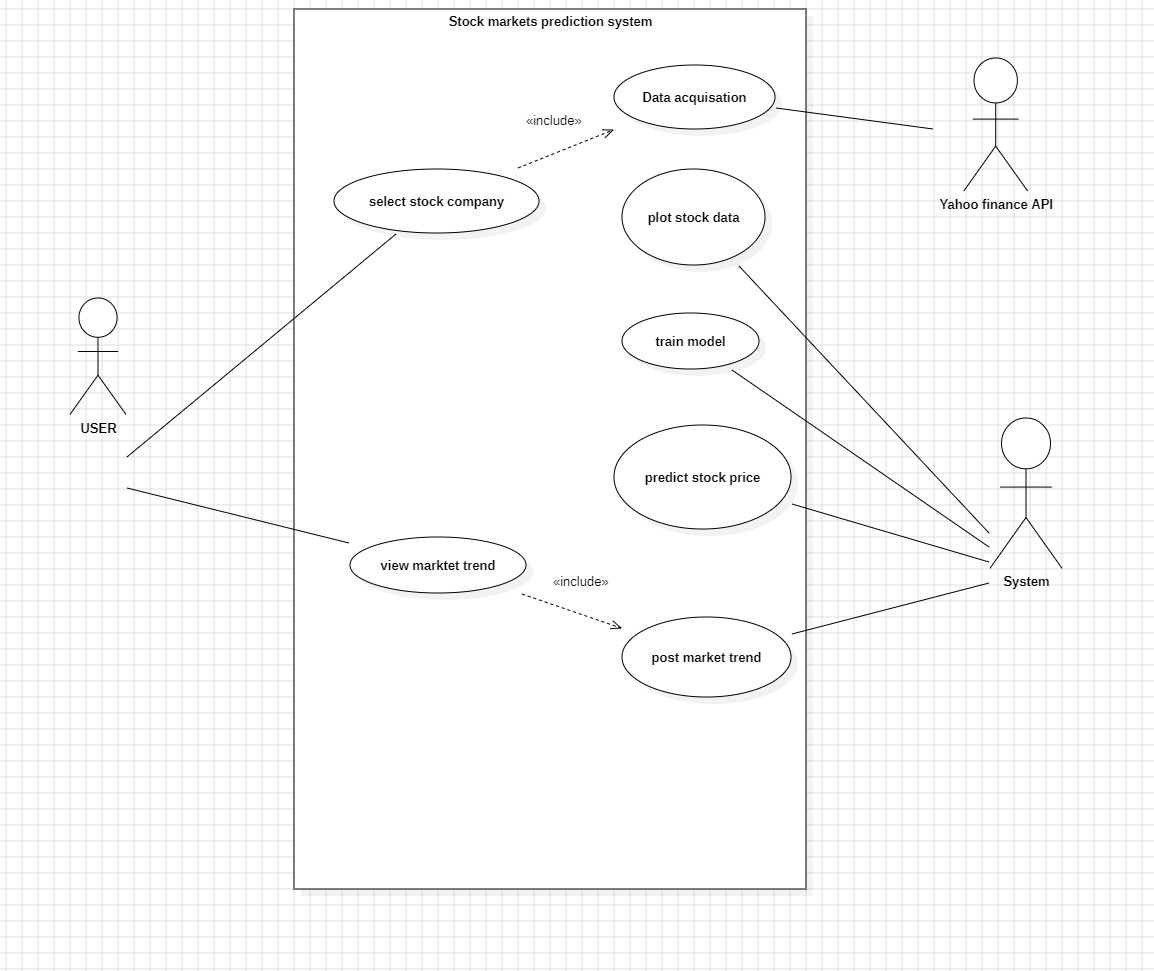
#### Figure 1 Activity diagram

1. The user initiates the process by requesting a stock prediction.
2. The system checks if historical stock data is available. If it is, the system proceeds with retrieving and preprocessing the data.
3. If the LSTM model is already trained, the system directly uses it to predict stock prices. Otherwise, it trains the LSTM model first.
4. Finally, the predicted stock prices are presented to the user. If no data is available, the user is notified accordingly.

**Use Case Diagram**

This diagram depicts the interactions between actors and the system, illustrating the systems

Functionality from the user’s perspective

1. 

#### Figure 2 Use case Diagram

**Sequence Diagram**

These sequence diagrams show the flow of messages and interactions between objects or

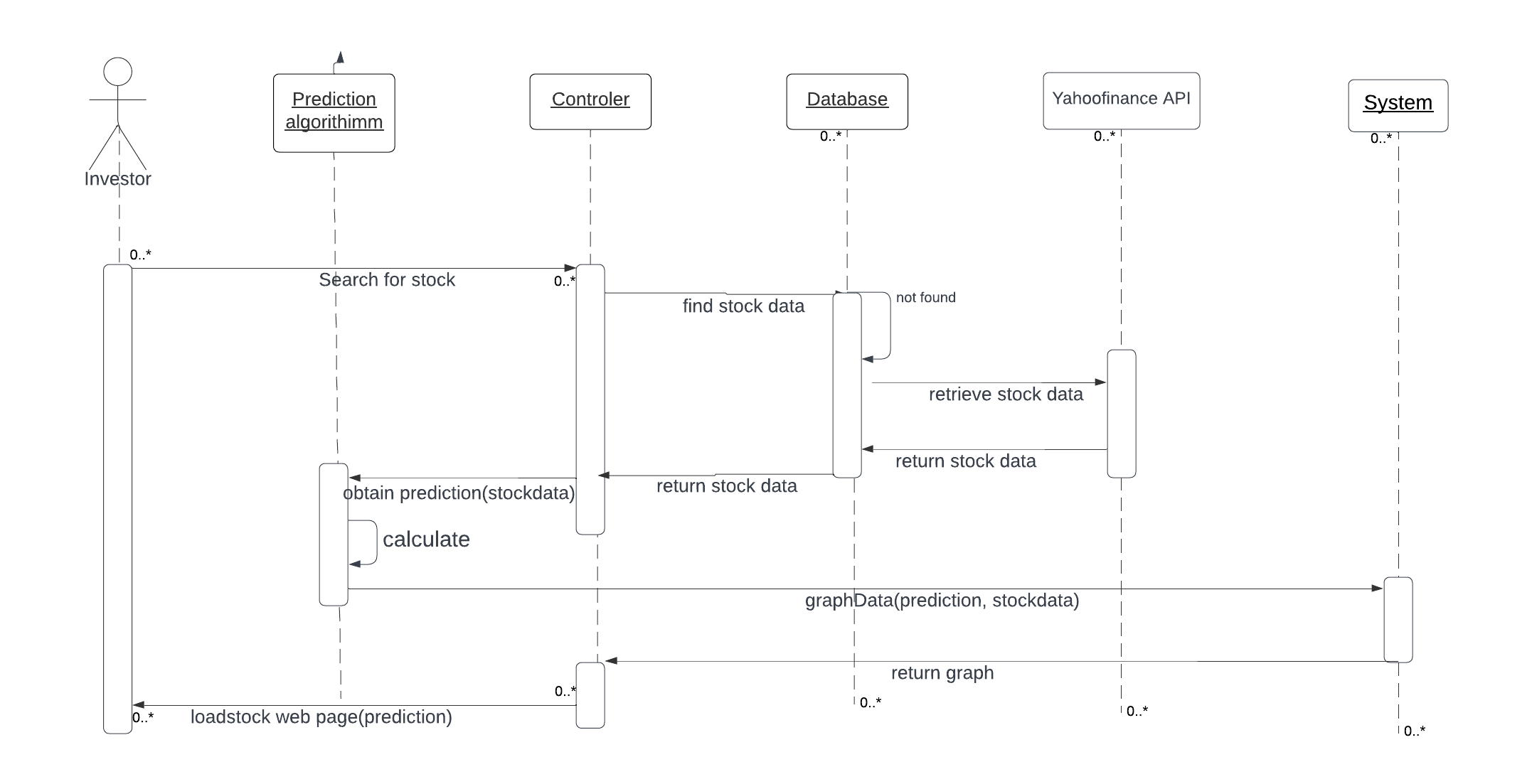
Components over time. They represent the dynamic behavior of a system by illustrating the

Sequence of interactions between different entities. The diagrams representing objects or components, and horizontal arrows representing messages exchanged

Between them. The sequence of messages is shown chronologically, allowing for a clear

Understanding of the order of interactions.

Below are two sequence diagrams on illustrating the order of events for the stock investor.



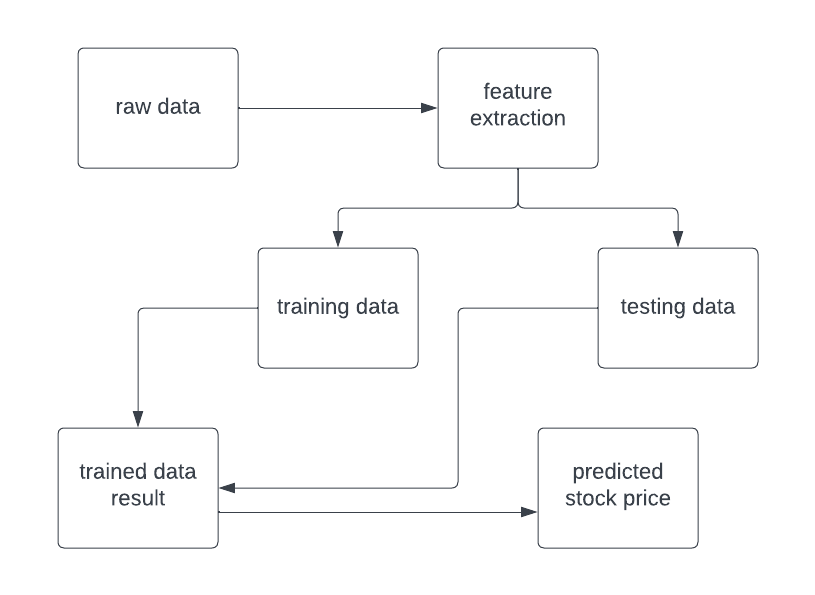
#### Figure 3 Sequence diagram

The sequence diagram shows that the system follows the following steps:

1. The user enters the stock name.
2. The application sends the stock name to the server application.
3. The server application processes the stock ticker to extract relevant features.
4. The server application uses the Deep learning model to predict the stock trend.
5. The server application sends the results back to the client application.
6. The client application displays the results to the user.
7. **System design**

This section gives a complete overview of the system architecture shown in Figure 1 and the process flow diagram of the system gives in detail with a diagram and information

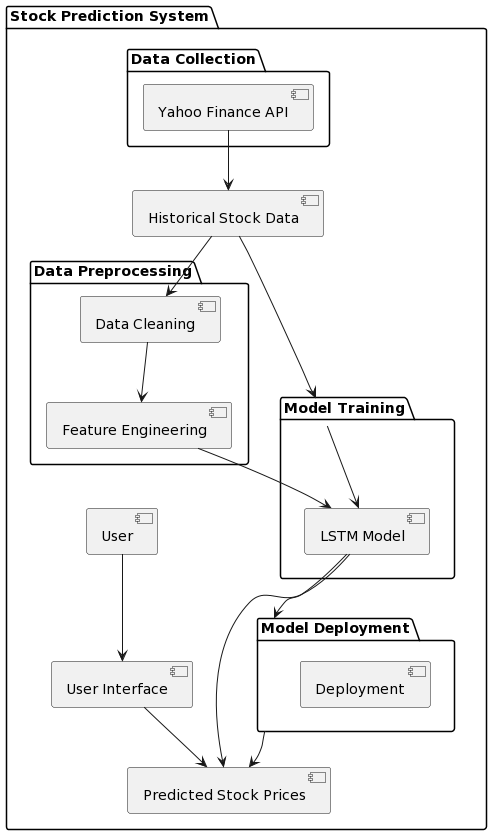
## 3.8 Logical design



#### Figure 4 logical Design

Yahoo finance It is likewise have dataset for distinctive fields. Various statistics scientist competes to create the nice fashions for predicting. It permits the consumer to gather the dataset for the predicting process. To resolve the actual time statistics technological know-how challenge. The dataset used on this challenge has been downloaded from yfinance. The dataset accrued in uncooked format. Stock fee dataset of few agencies has been accrued. The first step is to gather the uncooked inventory charge facts and pSreprocess the facts

1. **System architecture**



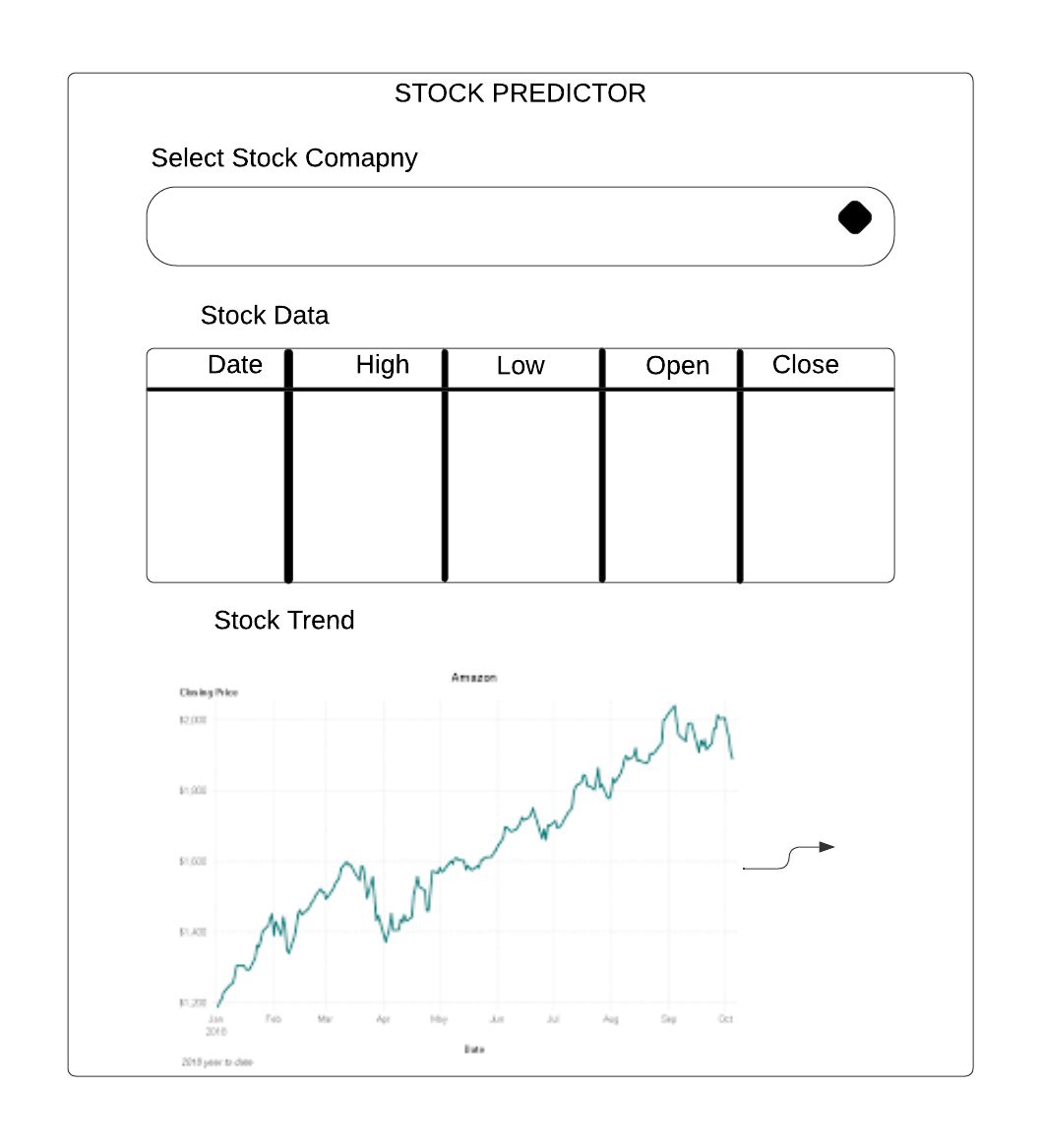
#### Figure 5 System Architecture

1. Data Collection: The Yahoo Finance API collects historical stock data.
2. Data Preprocessing: The collected data goes through cleaning and feature engineering processes.
3. Model Training: The preprocessed data is used to train the LSTM model.
4. Model Deployment: The trained model is deployed for prediction.
5. User Interface: Users interact with the system to get predicted stock prices.
6. Data Flow: Historical stock data flows through data preprocessing and model training to generate predicted stock prices, which are then presented to the user through the user interface.

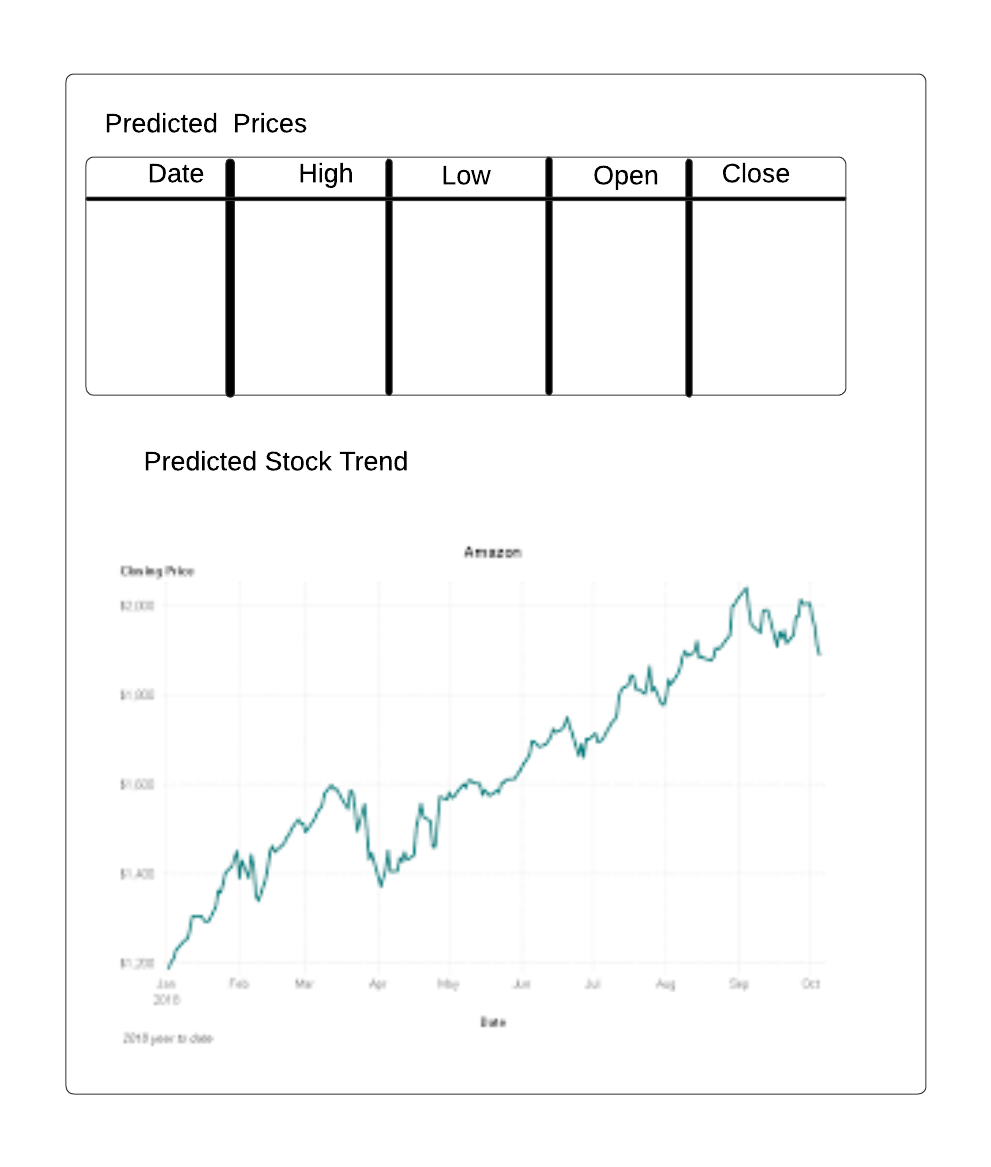
## 3.9 Physical design

1. **User Interface Designs.**

The system will have a web application interface to allow for taking use input that will be sent to the Deep learning model for prediction of trend. The interface will also display results to the user.

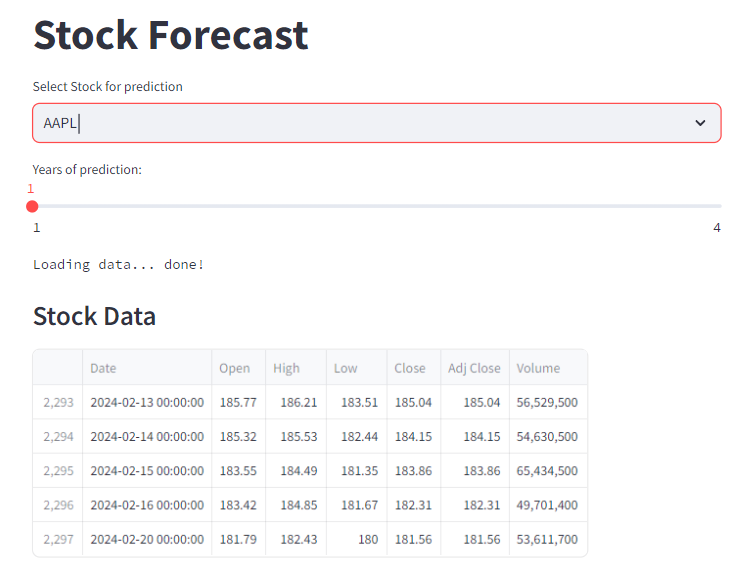


#### Figure 6 User interface home page



#### Figure 7 Prediction page

1. **System mockups.**

****

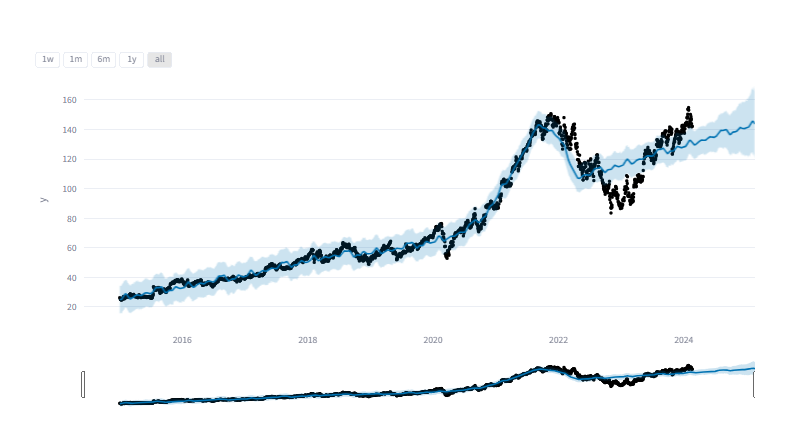
#### Figure 8 system mockup

when the user or investor accesses the system, this will be his/her landing page also the welcome page. In this page the user can select the stock name he or she would like to see the prediction.

After the user selects a stock to predict they will be provided with a forecast table of the predicted closing high and low stock prices.

The system will also provide the user with a graph of the trend. This will be plotted where prediction against the actual stock price.

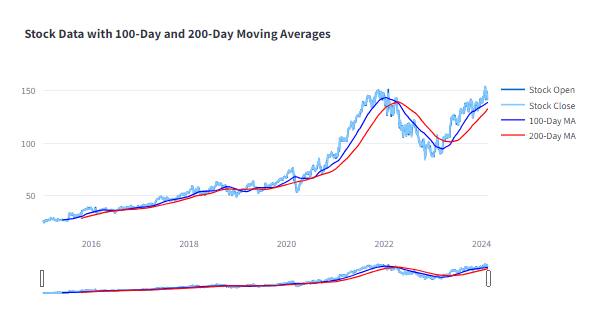
The system will also provide a 100 and 200 days moving average indicator for the user so as to assist in checking the accuracy of the predicted trend as shown below.



#### Figure 9 predicted trend

The system will also provide the user with a graph of the trend. This will be plotted where prediction against the actual stock price.

The system will also provide a 100 and 200 days moving average indicator for the user so as to assist in checking the accuracy of the predicted trend as shown below.



#### Figure 10 Calculated moving averages form

# CHAPTER FOUR

# SYSTEM IMPLEMENTATION AND TESTING, CONCLUSIONS AND RECOMMENDATIONS

## 4.0 Introduction

In this chapter, we delve into the implementation and testing phases of the stock market web application. This includes a detailed overview of the implementation environment, tools used, the process of code generation, testing strategies, and the outcomes. The chapter concludes with a reflection on the project's accomplishments, limitations, and recommendations for future enhancements.

## 4.1 Environment and Tools

The programming languages utilized in the implementation of the system are;

I utilized Streamlit, a Python library that simplifies the creation of web applications for machine learning and data science projects. Streamlit enables the creation of interactive web applications with minimal code, making it easy to showcase machine learning models and visualizations. I used Jupiter note book for data modeling and training of the model. for the user interface I used python programing language.

For the backend, I utilized Python along with frameworks like TensorFlow and Keras for implementing the LSTM model for stock prediction.

TensorFlow and Keras: These libraries provide powerful tools for building and training deep learning models, including LSTM networks for time series forecasting tasks like stock prediction.

Data Integration: I utilized the yfinance library to fetch historical stock price data directly from Yahoo Finance, allowing the model to train on real-world financial data.

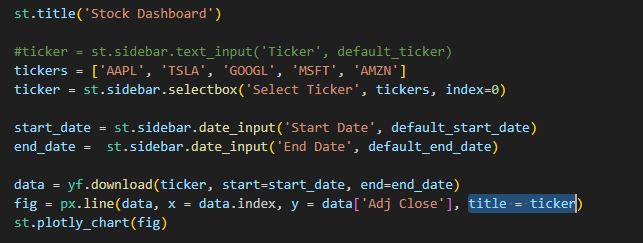
Data Transmission: In this setup, data is processed and predictions are made within Jupyter Notebook using Python. The predictions or relevant data are then served to the Streamlit web application through JSON format for display to the end-user.

## 4.1.1 System Code Generation

The system was written and developed using python. The model was integrated into the web application using streamlit. The following are the snapshots of the system.

**Fetching stock data from yahoo finance platform.**

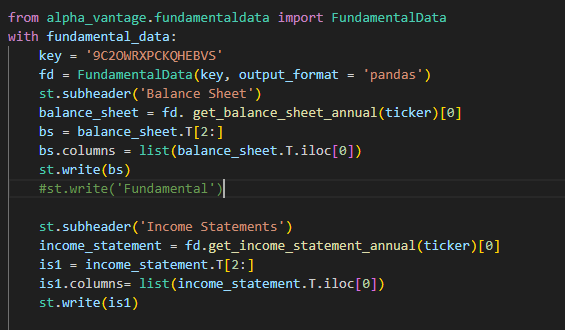
In order to perform predictions the system will download stock data from yahoo finance this will include opening, closing, low, high and the adjacent closing prices of that stock selected. Importing the stock data directly from yahoo finance saves on time and space rather than having a database to store the data.



#### Figure 11 Stock data fetching

**Connecting to alpha vantage**

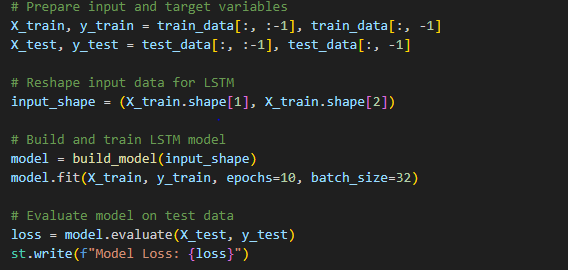
Connecting to alpha vantage grants access to stock fundamental data such as income statements and stock news that may affect the price movement of the stock. Having alpha vantage helps the model train the model with a complete data set therefore predicting stocks with a news sentiment. This connection is shown below.



#### Figure 12 Fetching stock news and fundamental data

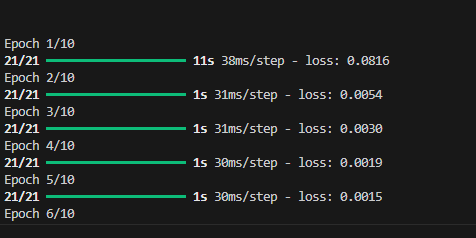
**Model training**

The training data is divided into batches. The neural network processes each batch of data and makes predictions. A loss function is used to calculate the difference (error) between predicted and actual outputs. Using the calculated gradients, optimization algorithms such as Gradient Descent are used to update the model parameters (weights and biases).



#### Figure 13 Model training

One complete pass through the entire dataset (all batches) is called an epoch. After completing an epoch, the model has seen and learned from the entire training dataset once.



#### Figure 14 Successful epoch's

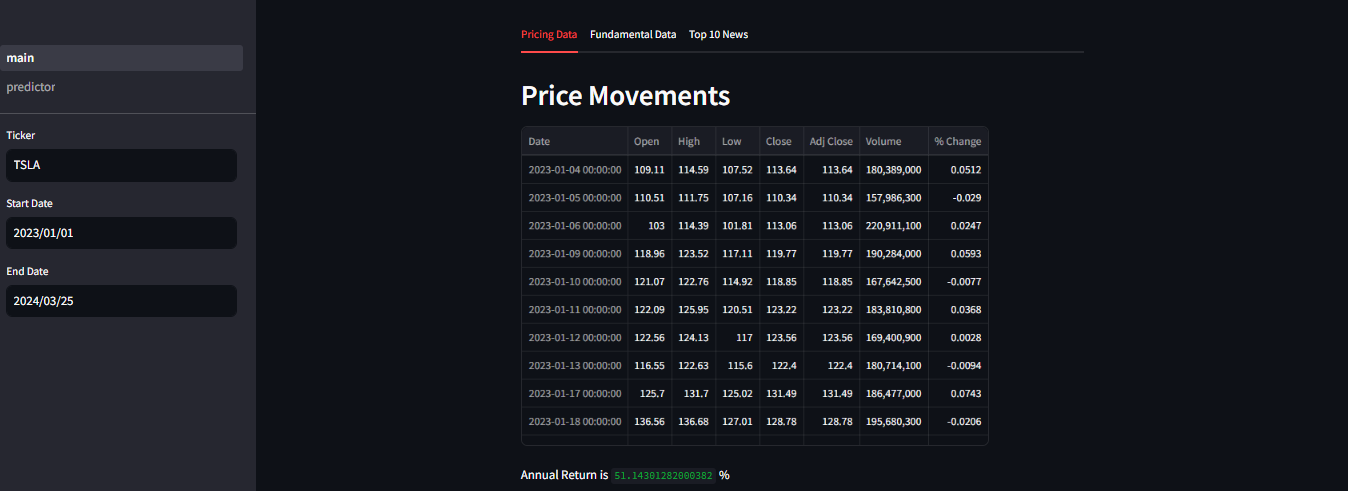
## 4.1.2 Snapshots of the system

**Landing page**

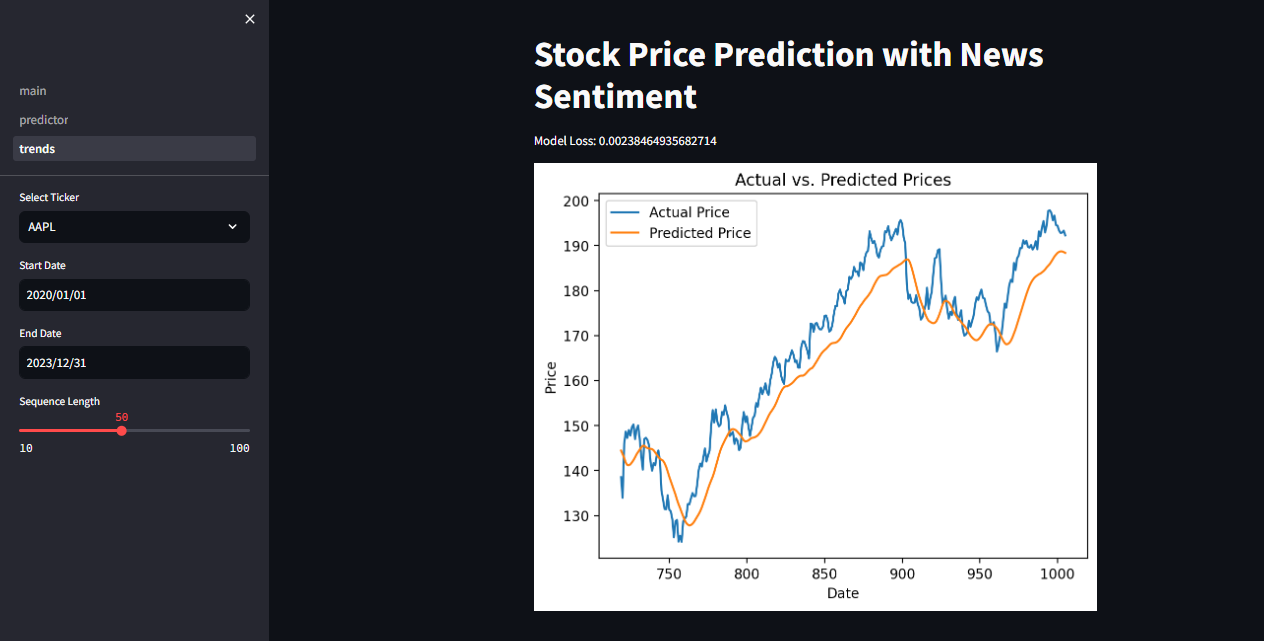


#### Figure 15 system snapshot home page

**Stock data**



**Stock predictions page**



#### Figure 16 snapshot for stock predictor



*Implemented trade for testing*

## 4.2 TESTING

This section brings into light the overall testing strategy used for testing the web-based Application. Testing involved subjecting the system to a valid set of inputs to validate the set of outputs against the predetermined set of outputs.

**Testing Plan**

The testing plan involved doing individual module tests, to validate and verify that the system meets the requirements as guided by the developmental design.

The tests performed included the following; Unit Testing, GUI testing, Validation Testing

## 4.2.1 Unit Testing

This is a type of white box testing and it focuses on testing individual units or modules of software in isolation. It is a test in which the smallest testable portion of the software is tested to verify its functionality against its specification

|  |  |
| --- | --- |
| Task To be Tested | Data fetching Process |
| **Test Case** | Verify the ability to fetch historical stock data and preprocess it for model training |
| **Test Plan** | * Fetch historical stock data using yfinance. * Preprocess the data by handling missing values and scaling. |
| **Expected Results** | The system should successfully fetch historical stock data and preprocess it for model training. |
| **Actual Results** | Historical stock data was fetched and preprocessed without errors. |

## 4.2.2. GUI Testing

GUI (Graphical User Interface) testing involves verifying the functionality, usability, and overall user experience of the application's graphical elements.

|  |  |
| --- | --- |
| Task To be Tested | Testing the user interface for usability and functionality. |
| **Test Case** | Ensure the user interface provides an intuitive experience |
| **Test Plan** | * Interact with the user interface for fetching predictions. * Verify the responsiveness and clarity of displayed information. |
| **Expected Results** | The user interface should be intuitive, responsive, and display predictions clearly. |
| **Actual Results** | The user interface provided a seamless experience with clear and responsive prediction displays. |

## 4.2.3. Integration Testing

Integration testing involves assessing the seamless interaction and cooperation between various elements within the system. This testing phase is crucial to ensure that the individual components, previously developed and tested in isolation, harmonize effectively when integrated into the broader application.

In this testing approach, the primary focus is on evaluating the interplay of components, as opposed to isolated component testing. This emphasis is necessary because defects and issues often emerge at the points where different components converge, which may not become apparent during individual testing

1. **Unified Inter-Testing Approach:** This method involved the integration of all components into a single testing scenario. Here a user interacted with all modules of the system. The goal was to assess the collective behavior of components when combined and identify potential conflicts or inconsistencies in their interactions.
2. **Error Handling and Exception Testing:** It emphasizes scenarios where users provide ambiguous or erroneous inputs. Testing demonstrated the application's capability to gracefully handle errors and provide informative responses.

## 4.2.4. Acceptance Testing

Acceptance testing is performed to validate the software against the user requirements and is done in the final phase of testing. This testing ensure that the software does what the user wants it to do and checks the acceptability of the system

|  |  |  |  |
| --- | --- | --- | --- |
| Test id | Description | Test done or not | result |
| **1** | Does the system perform all the functions it is supposed to, as outlined in the requirements? | **Yes** | Not all |
| **2** | Is the user documentation clear, comprehensive, and easy to follow? | **Yes** | **Ok** |
| **3** | Is it easy to navigate throughout the system? | **Yes** | **Ok** |
| **4** | Are all data inputs accurately processed and errors displayed where necessary? | **Yes** | **Ok** |
| **5** | Can users access only the functions and data they are authorized to use? | **Yes** | **Ok** |
| **6** | Does the system respond appropriately based on your input | **Yes** | **Ok** |
| **7** | Are you satisfied with the rate at which the system responds? | **Yes** | **Ok** |

## 4.2.5. Testing predictions

**Visual Comparison**. this involved Plotting the predicted trends generated by your model alongside the actual trends observed in the market. Use line charts to visualize both the predicted and actual stock prices over time. Visual inspection can provide immediate insights into how well the model's predictions align with reality.

**Back testing:** This involved conducting a back test to assess the profitability of trading strategies based on the model's predictions. Simulate trades using the predicted trends and evaluate their performance against the actual market outcomes. Back testing helps determine the practical utility of the model's predictions in real-world trading scenarios.

**Conclusions**

This section serves to summarize key findings, highlight achievements, and offer insights gained throughout the project. The following are achievements made in this project:

Enhanced interactivity: the implemented system has achieved key principles of interactivity such as feedback, consistency, visibility, and learnability. These interactive features have empowered users to access detailed product information and make informed decisions.

The user-centric approach, intuitive design, and responsive interfaces have amplified user engagement, enhancing the overall experience.

**Recommendations**

Scalability enhancements: Investigate ways to improve the prediction system's scalability so that it can handle more data and user requests. This could include optimizing data processing pipelines, leveraging distributed computing frameworks, or scalability through the use of cloud resources.

* Consider implementing off-chain scaling solutions, such as parallel or asynchronous data processing, to boost system throughput and response time.
* Consider alternative architectures, such as microservices or serverless computing, to improve scalability and meet the growing demand for stock prediction services.

Interoperability Standards:

* Develop interoperability standards and protocols to enable seamless integration with existing financial data sources, trading platforms, and analysis tools.
* Work with industry stakeholders, such as stock exchanges, financial institutions, and regulatory bodies, to create common data formats and APIs for interoperable data sharing.

# REFERENCES

Goddard, J., Kita, A., & Wang, Q. (2015). Investor attention and FX market volatility. *Journal of International Financial Markets, Institutions and Money*, *38*, 79-96.

Reznik, N., & Pankratova, L. (2018). High-Frequency Trade as a Component of Algorithmic Trading: Market Consequences. In *ICTERI Workshops* (pp. 73-83).

Cohen, G. (2022). Algorithmic trading and financial forecasting using advanced artificial intelligence methodologies. *Mathematics*, *10*(18), 3302.

Yang, G. R., & Wang, X. J. (2020). Artificial neural networks for neuroscientists: a primer. *Neuron*, *107*(6), 1048-1070.

Jorgenson, D. W., Weitzman, M. L., ZXhang, Y. X., Haxo, Y. M., & Mat, Y. X. (2023). Can Neural Networks Predict Stock Market?. *AC Investment Research Journal*, *220*(44).

Kouaissah, N., Orlandini, D., Ortobelli, S., & Tichý, T. (2020). Theoretical and practical motivations for the use of the moving average rule in the stock market. *IMA Journal of Management Mathematics*, *31*(1), 117-138.

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, *9*(1), 381-386.

Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G. (2020). A survey on machine learning for stock price prediction: Algorithms and techniques.

Zhang, D., & Lou, S. (2021). The application research of neural network and BP algorithm in stock price pattern classification and prediction. *Future Generation Computer Systems*, *115*, 872-879.

Javed Awan, M., Mohd Rahim, M. S., Nobanee, H., Munawar, A., Yasin, A., & Zain, A. M. (2021). Social media and stock market prediction: a big data approach. *MJ Awan, M. Shafry, H. Nobanee, A. Munawar, A. Yasin et al.," Social media and stock market prediction: a big data approach," Computers, Materials & Continua*, *67*(2), 2569-2583.

Poongodi, M., Vijayakumar, V., & Chilamkurti, N. (2020). Bitcoin price prediction using ARIMA model. *International Journal of Internet Technology and Secured Transactions*, *10*(4), 396-406.

Wang, Y., Liu, H., Guo, Q., Xie, S., & Zhang, X. (2019). Stock volatility prediction by hybrid neural network. *IEEE Access*, *7*, 154524-154534.

Buehler, H., Gonon, L., Teichmann, J., & Wood, B. (2019). Deep hedging. Quantitative Finance, 19(8), 1271-1291.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning in finance. Applied Stochastic Models in Business and Industry, 33(1), 3-12.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1-2), 307-319.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research, 259(2), 689-702.

Medeiros, M. C., Vasconcelos, G. F., Veiga, A., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. Journal of Business & Economic Statistics, 39(1), 98-119.

Nguyen, T. T., Shirai, T., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. Expert Systems with Applications, 42(24), 9603-9611.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1643-1647). IEEE.

Zhang, Y., & Zohren, S. (2020). Deep learning for portfolio optimization. The Journal of Financial Data Science, 2(4), 34-48.

QUESTIONAIRE **Stock prediction using Deep learning approach**

**1. User Interaction and Experience**

**- How comfortable are you with using online tools or applications for financial trading?**

- [ ] Very Comfortable

- [ ] Comfortable

- [ ] Neutral

- [ ] Uncomfortable

- [ ] Very Uncomfortable

**- Would you prefer to input stock information manually, or would you like automated** suggestions based on popular stocks?

- [ ] Manually input stock information

- [ ] Automated suggestions

- [ ] Both

- [ ] Not sure

**2. Data Needs and Preferences**

**- How important is it for you to have access to historical stock data when making trading decisions?**

- [ ] Very Important

- [ ] Important

- [ ] Neutral

- [ ] Not Important

- [ ] Not Sure

**- Do you find it useful to have access to stock-related news articles to help inform your trading decisions?**

- [ ] Very Useful

- [ ] Useful

- [ ] Neutral

- [ ] Not Useful

- [ ] Not Sure

**3. Data Processing and Analysis**

**- Are you interested in the system automatically cleaning and preparing data for analysis (e.g., handling missing values and extracting key features)?**

- [ ] Yes, very interested

- [ ] Yes, somewhat interested

- [ ] Neutral

- [ ] Not interested

- [ ] Not sure what this means

**- Would you like the system to show only the most relevant news (e.g., top ten news articles related to the stock)?**

- [ ] Yes

- [ ] No

- [ ] Maybe

- [ ] Not sure

**4. Predictive Analytics**

**- How important is it to you to have machine learning models predict future stock prices?**

- [ ] Very Important

- [ ] Important

- [ ] Neutral

- [ ] Not Important

- [ ] Not Sure

**- Would you prefer the system to generate detailed predictions for future stock prices based on past trends?**

- [ ] Yes, I prefer detailed predictions

- [ ] No, I do not need detailed predictions

- [ ] I am not sure

**5. Data Visualization**

**- How helpful do you think visualizing both actual stock data and predicted stock prices on charts would be for your trading decisions?**

- [ ] Very Helpful

- [ ] Helpful

- [ ] Neutral

- [ ] Not Helpful

- [ ] Not Sure

**- Would you like to see weekly, daily, and monthly stock trends displayed visually?**

- [ ] Yes, all of them

- [ ] Only weekly

- [ ] Only daily

- [ ] Only monthly

- [ ] Not interested

**6. Integration with Real-Time Data**

**- How valuable is real-time data integration to you for making immediate trading decisions?**

- [ ] Extremely Valuable

- [ ] Valuable

- [ ] Neutral

- [ ] Not Valuable

- [ ] Not Sure

**- Would you like the system to connect with other platforms or APIs to provide up-to-the-minute data and predictions?**

- [ ] Yes

- [ ] No

- [ ] Maybe

- [ ] Not Sure

**7. Overall Preferences and Feedback**

**- What features or functionalities would you consider most essential in a stock market prediction system? (Select all that apply)**

- [ ] User-friendly interface

- [ ] Historical data analysis

- [ ] Real-time data integration

- [ ] Stock news aggregation

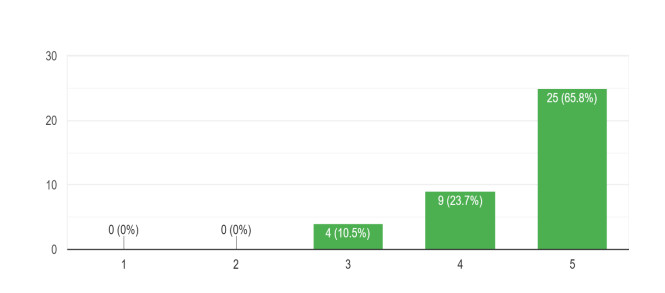
- [ ] Predictive analytics

- [ ] Data visualization

- [ ] Other: [Please specify**]**

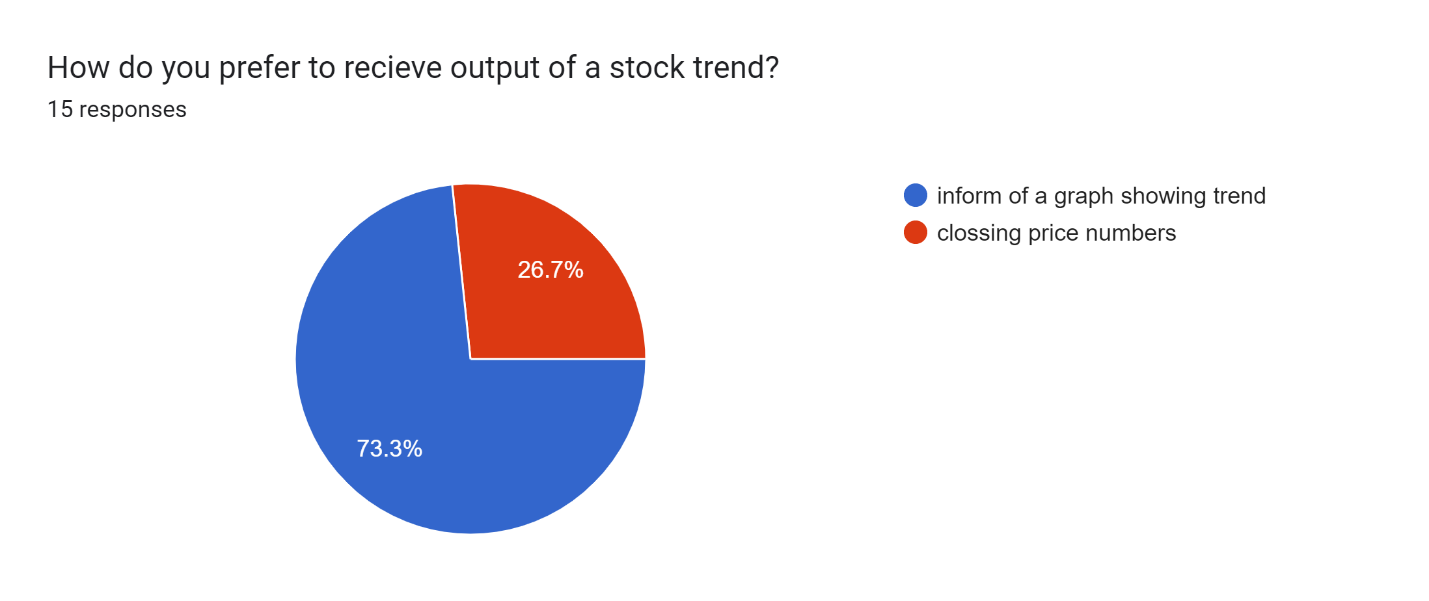
**- Do you have any other suggestions or requirements for the stock market prediction system?**

**On a scale of 1-5, how important is it for you to have a reliable and trustworthy web-based stock prediction system? (1 - Not important at all, 5- Extremely important)**



This analysis deduced that the respondents value an efficient channel that enables them to analyze and predict the stock trend easily.

**How do you prefer to receive output of a stock trend?**



This analysis deduced that the respondents would appreciate a system that will provide for them a stock trend in form of a graph.

**How would you rate the success rate of new stock investors inthe market acording to your yealy data**

It was evident that new stock traders in the market are more likely to blow their accounts dues to lack of assistance and the difficulty in analyzing the stock trend.

**How would you like a stock prediction system to provide output for you**

The findings of this analysis deduced that the responders are truly faced by various challenges in their quest for stock prediction.

# APPENDIX:

## PROJECT SCHEDULE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ACTIVITIES** | **DURATION IN HRS** | **PROPOSED START DATE** | **ACTUAL START DATE** | **PROPOSED END DATE** | **ACTUAL END DATE** | **DELIVERABLES** |
| Problem Identification | **96** | **15/01/2024** | **22/01/2024** | **19/01/2024** | **26/01/2024** | Project Idea |
| Proposal Writing | **144** | **27/01/2024** | **29/01/2024** | **01/02/2024** | **03/02/2024** | Problem Statement |
| Data Collection | **96** | **09/04/2024** | **11/04/2024** | **12/04/2024** | **15/04/24** | Data Report |
| Data Analysis | **48** | **15/04/2024** | **18/04/2024** | **17/04/2024** | **20/04/2024** | Data Report |
| Design | **96** | **22/04/2024** | **24/04/2024** | **25/04/2024** | **27/04/2024** | System Design |
| Coding and Testing | **960** | **1/05/2024** | **14/05/2024** | **21/05/2024** | **10/07/2024** | Testing Result |
| Documentation | **144** | **28/05/2024** | **30/07/2024** | **04/06/2024** | **4//08/2024** | Project Documentation |

##### Table 1 Project schedule

**GANT CHART**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ACTIVITY | JANUARY | FEBRUARY | MARCH | APRIL |
| SOFTWARE  REQUIREMENT |  |  |  |  |
| LOGICAL  DESIGN |  |  |  |  |
| DOCUMENTATION |  |  |  |  |
| PRESENTATION |  |  |  |  |

##### Table 2 Gant chart

**BUDGET**

|  |  |
| --- | --- |
| DEVICE NAME | PRICE (Ksh) |
| Computer | 60,000 |
| Jupiter software | 300 |
| Keyboard and mouse | 3000 |
| Meta trader 5 | 0 |
| Streamlit | 0 |
| Alpha-vantage | 1300 |

*Table 3 Budget*

# APPENDIX

**Interview questions**

**Questions asked at fusion forex ltd**

1. What are your primary objectives in seeking a stock prediction system?
2. Can you describe the specific challenges or pain points you're facing in your current approach to stock trading or investment decision-making?
3. What are the key features or functionalities you expect from the stock prediction system?
4. How do you envision integrating the stock prediction system into your existing workflow or investment strategy?
5. Do you have any preferences or requirements regarding the presentation or visualization of stock prediction results?
6. How would you rate the success of new stock investors according to your yearly data?
7. What level of accuracy or reliability do you expect from the stock prediction system?

**RESOURCES REQUIREMENTS**

**Hardware Requirements**

1. A Laptop or Desktop with the following specifications:
2. Core i7 Processor (4 Cores Minimum)
3. 16GB minimum RAM
4. 1TB Hard-Drive (SSD recommended for faster processing)
5. Windows/Linux/Mac Operating System

**Software Requirements**

The following software and tools will be required for development and deployment of the stock prediction system:

1. Python 3.8+: Programming language used for implementing the system's logic and algorithms.
2. Jupyter Notebook/Google Colab: For experimenting with LSTM model and conducting data analysis.
3. Pandas, NumPy, Matplotlib, Seaborn: Python libraries for data manipulation, numerical computing, and data visualization.
4. TensorFlow/Keras: Deep learning framework for building and training LSTM model for stock prediction.
5. Streamlit: Python library for building interactive web applications. It will be used to create a user-friendly interface for users to interact with the prediction system.
6. Yahoo Finance and Alpha Vantage APIs: For fetching historical stock data and real-time market data.
7. News API: For retrieving latest news related to stocks and financial markets.