

# **“STOCK PRICE PREDICTION USING MACHINE LEARNING”**

**A**

***Project Report***

***submitted***

***in partial fulfillment***

***for the award of the Degree of***

***Bachelor of Technology***

***in Department of Information Technology***



**Project Mentor:**

Mr Vipin Jain

Associate Professor

**Submitted By :**

Aditya Baghela (19ESKIT004)

Anosh Field (19ESKIT010)

Atishay Jain (19ESKIT020)

Dhruv Raj Naruka (19ESKIT024)

**Department of Information Technology  
Swami Keshvanand Institute of Technology, M & G, Jaipur  
Rajasthan Technical University, Kota  
Session 2022-2023**

**Swami Keshvanand Institute of Technology,  
Management & Gramothan, Jaipur  
Department of Information Technology**

**CERTIFICATE**

This is to certify that **Mr Aditya Baghela**, a student of B.Tech(Information Technology), 8th semester has submitted his Project Report entitled “**Stock Price Prediction Using Machine Learning**” under my guidance.

**Mentor**

Name: Mr Vipin Jain

Associate Professor

Signature.....

**Coordinator**

Name: Mrs Sanju Choudhary

Associate Professor

Signature.....

**Swami Keshvanand Institute of Technology,  
Management & Gramothan, Jaipur  
Department of Information Technology**

**CERTIFICATE**

This is to certify that **Mr Anosh Field**, a student of B.Tech(Information Technology), 8th semester has submitted his Project Report entitled “**Stock Price Prediction Using Machine Learning**” under my guidance.

**Mentor**

Mr Vipin Jain

Associate Professor

Signature.....

**Coordinator**

Mrs Sanju Choudhary

Associate Professor

Signature.....

**Swami Keshvanand Institute of Technology,  
Management & Gramothan, Jaipur  
Department of Information Technology**

**CERTIFICATE**

This is to certify that **Mr Atishay Jain**, a student of B.Tech(Information Technology), 8th semester has submitted his Project Report entitled “**Stock Price Prediction Using Machine Learning**” under my guidance.

**Mentor**

Mr Vipin Jain

Associate Professor

Signature.....

**Coordinator**

Mrs Sanju Choudhary

Associate Professor

Signature.....

**Swami Keshvanand Institute of Technology,  
Management & Gramothan, Jaipur  
Department of Information Technology**

**CERTIFICATE**

This is to certify that **Mr Dhruv Raj Naruka**, a student of B.Tech(Information Technology), 8th semester has submitted his Project Report entitled “**Stock Price Prediction Using Machine Learning**” under my guidance.

**Mentor**

Mr Vipin Jain

Associate Professor

Signature.....

**Coordinator**

Mrs Sanju Choudhary

Associate Professor

Signature.....

## DECLARATION

We hereby declare that the report of the project entitled "Stock Price Prediction Using Machine Learning" is a record of an original work done by us at Swami Keshvanand Institute of Technology, Management and Gramothan, Jaipur under the mentorship of "Mr. Vipin Jain"(Dept.of Information Technology) and coordination of "Mrs. Sanju Choudhary" (Dept.of Information Technology). This project report has been submitted as the proof of original work for the partial fulfillment of the requirement for the award of the degree of Bachelor of Technology (B.Tech) in the Department of Information Tehnology.It has not been submitted anywhere else, under any other program to the best of our knowledge and belief.

### Team Members

(Aditya Baghela,19ESKIT004)  
(Anosh Field,19ESKIT010)  
(Atishay Jain,19ESKIT020)  
(Dhruv Raj Naruka,19ESKIT024)

### Signature

# Acknowledgement

A project of such a vast coverage cannot be realized without help from numerous sources and people in the organization. We take this opportunity to express our gratitude to all those who have been helping us in making this project successful.

We are highly indebted to our faculty mentor Mr Vipin Jain. He has been a guide, motivator source of inspiration for us to carry out the necessary proceedings for the project to be completed successfully. We also thank our project coordinator Mrs Sanju Choudhary for her co-operation, encouragement, valuable suggestions and critical remarks that galvanized our efforts in the right direction.

We would also like to convey our sincere thanks to Prof. Dr. Anil Choudhary, HOD, Department of Information Technology, for facilitating, motivating and supporting us during each phase of development of the project. Also, we pay our sincere gratitude to all the Faculty Members of Swami Keshvanand Institute of Technology, Management and Gramothan, Jaipur and all our Colleagues for their co-operation and support.

Last but not least we would like to thank all those who have directly or indirectly helped and cooperated in accomplishing this project.

## **Team Members:**

(Aditya Baghela,19ESKIT004)

(Anosh Field,19ESKIT010)

(Atishay Jain,19ESKIT020)

(Dhruv Raj Naruka,19ESKIT024)

# Table of Content

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Problem Statement and Objective . . . . .	1
1.2	Literature survey . . . . .	1
1.3	Introduction to Project . . . . .	3
1.4	Proposed Logic / Algorithm / Business Plan / Solution / Device . . .	4
1.5	Scope of the Project . . . . .	4
<b>2</b>	<b>Software Requirement Specification</b>	<b>6</b>
2.1	Overall Description . . . . .	6
2.1.1	Product Perspective . . . . .	6
2.1.1.1	System Interfaces . . . . .	7
2.1.1.2	User Interfaces . . . . .	7
2.1.1.3	Software Interfaces . . . . .	7
2.1.1.4	Communications Interfaces . . . . .	7
2.1.1.5	User Characteristics . . . . .	8
<b>3</b>	<b>System Design Specification</b>	<b>9</b>
3.1	System Architecture . . . . .	9
3.2	High Level Design Diagrams . . . . .	10
3.2.1	Use Case Diagram . . . . .	11
3.2.2	Activity Diagram . . . . .	12
3.2.3	Data-Flow Diagram . . . . .	13
3.2.4	Sequence Diagram . . . . .	14
<b>4</b>	<b>Methodology and Team</b>	<b>15</b>
4.1	Overview of Existing Work . . . . .	15
4.2	Proposed Work . . . . .	17
4.2.1	Working of LSTM model . . . . .	18



4.2.2	Working of model: . . . . .	20
4.3	Team Members, Roles & Responsibilities . . . . .	21
<b>5</b>	<b>Centering System Testing</b>	<b>24</b>
5.1	Functionality Testing . . . . .	24
5.2	Performance Testing . . . . .	26
<b>6</b>	<b>Test Execution Summary</b>	<b>27</b>
<b>7</b>	<b>Project Screen Shots</b>	<b>28</b>
7.1	Dataset Detail . . . . .	28
7.2	Tool Technologies . . . . .	28
7.2.1	Python . . . . .	28
7.2.2	Numpy . . . . .	29
7.2.3	Scikit Learn . . . . .	30
7.2.4	Tensorflow . . . . .	30
7.2.5	Keras . . . . .	31
7.2.6	Compiler Option . . . . .	31
7.2.7	Jupyter Notebook . . . . .	32
7.3	Results . . . . .	32
<b>8</b>	<b>Project Summary and Conclusions</b>	<b>40</b>
<b>9</b>	<b>References</b>	<b>41</b>

# List of Figures

---

2.1	System Interface . . . . .	7
3.1	System Architecture . . . . .	9
3.2	High Level Design Diagram . . . . .	10
3.3	Use Case diagram . . . . .	11
3.4	Activity Diagram . . . . .	12
3.5	Data flow diagram . . . . .	13
3.6	Sequence diagram . . . . .	14
4.1	Proposed Workflow . . . . .	17
4.2	LSTM model . . . . .	18
4.3	Model . . . . .	20
7.1	Read Dataset . . . . .	32
7.2	Graph of Close Price history . . . . .	33
7.3	Data Scaling . . . . .	33
7.4	LSTM Summary . . . . .	35
7.5	Approach of LSTM . . . . .	35
7.6	Training Process . . . . .	36
7.7	Predicted Close Price . . . . .	37
7.8	Plot for Real vs Predicted value for Amazon using LSTM . . . . .	38
7.9	Plot for Real vs Predicted value for Facebook using LSTM . . . . .	38
7.10	Plot for Real vs Predicted value for Google using LSTM . . . . .	39
7.11	Plot for Real vs Predicted value for Amazon using LSTM . . . . .	39

# Chapter 1

## Introduction

---

### 1.1 Problem Statement and Objective

The stock market appears in the news every day. You hear about it every time it reaches a new high or a new low. The rate of investment and business opportunities in the Stock market can increase if an efficient algorithm could be devised to predict the short term price of an individual stock. Previous methods of stock predictions involve the use of Artificial Neural Networks and Convolution Neural Networks which has an error loss at an average of 20%. In this report, we will see if there is a possibility of devising a model using Recurrent Neural Network which will predict stock price with a less percentage of error. And if the answer turns to be yes, we will also see how reliable and efficient will this model be.

### 1.2 Literature survey

There are following three types of survey that are required to be done on stock market to predict the future price of any stock which are as follows:

#### **Time Series Analysis**

The definition of forecasting can be like this the valuation of some upcoming result or results by analyzing the past data. It extends different areas like industry and business, economics and finance, environmental science. Forecasting problems can be classified as follows:

- Long term forecasting (estimation beyond 2 years)
- Medium-term forecasting (estimation for 1 to 2 years)
- Short term forecasting (estimation for weeks or months, days, minutes, few seconds)

The analysis of time consists of several forecasting problems. The designation of a time series is a linear classification of observations for a selected variable. The variable of the stock price in our case. Which can weather multivariate or univariate? Only particular stock is included in the univariate data while more than one company for various instances of time is added in multivariate. For investigating trends, patterns and cycle or periods the analysis of time series advantages in the present data. In spending money wisely an early data of the bullish or bearish in the case of the stock market. Also, for categorizing the best-performing companies the analysis of patterns plays its role for a specific period. This makes forecasting as well as time series analysis an important research area.

### **Fundamental analysis**

Fundamental Analysts are concerned with the business that reasons the stock itself. They assess a company's historical performance as well as the reliability of its accounts. Different performance shares are created that aid the fundamental forecaster with calculating the validity of a stock, such as the P/E ratio. Warren Buffett is probably the foremost renowned of all Fundamental Analysts.

What fundamental analysis within the stock market is trying to reach, is organizing the true value of a stock, that then will be matched with the worth it is being listed on stock markets and so finding out whether the stock on the market is undervalued or not. Find out the correct value will be completed by numerous strategies with primarily a similar principle. The principle is that an organization is price all its future profits. Those future profits have to be discounted to their current value. This principle goes on the theory that a business is all about profits and nothing else. Differing to technical analysis, the fundamental analysis is assumed as further as a long approach.

Fundamental analysis is created on conviction that hominoid society desires capital to make progress and if the company works well, than it should be rewarded with an additional capital and outcome in a surge in stock price. Fundamental analysis is usually used by the fund managers as it is the maximum sensible, objective and prepared from openly existing data like financial statement analysis.

One more meaning of fundamental analysis is on the far side bottom-up business analysis, it discusses the top-down analysis since initial analyzing the world economy, followed by country analysis and also sector analysis, and last the company level analysis.

### **Technical analysis**

Chartists or the technical analysts are not involved with any other of the fundamentals of the company. The long run price of a stock based generally exclusively on the trends of the past value (a form of time series analysis) that is set by them. The head and shoulders or cup and saucer are various numerous patterns that are employed. Also the techniques, patterns are used just like the oscillators, exponential moving average (EMA), support and momentum and volume indicators. Candlestick patterns, believed to have been initial developed by Japanese rice merchants, are nowadays widely used by technical analysts. For the short-term approaches, the technical analysis is used compare to long-run ones. So, in commodities and forex markets it is more predominant wherever traders target short-term price movements. There are basic rules are used in this analysis, first all significant about a company is already priced into the stock, another being that the value changes in trends and finally that history (of prices) tends to repeat itself that is especially due to the market science.

- Business
- Companies
- Insurance company
- Government Agency
- This application is helpful for stock investors, sellers, buyers, brokers

## **1.3 Introduction to Project**

Due to the high profit of the stock market, it is one of the most popular investments. People investigated for methods and tools that would increase their gains

while minimizing the risk, as the level of trading and investing grew. Two stock exchanges namely- the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE), which are the most of the trading in Indian Stock Market takes place. Sensex and Nifty are the two prominent Indian Market Indexes. Since the prices in the stock market are dynamic, the stock market prediction is complicated.

From gradually the very past years some forecasting models are developed for this kind of purpose and they had been applied to money market prediction. Generally, this classification is done by:

1. Time series analysis
2. Fundamental analysis
3. Technical analysis

A stock market prediction is described as an action of attempting to classify the future value of the company stock or other financial investment traded on the stock exchange. The forthcoming price of a stock of the successful estimation is called the Yield significant profit. This helps you to invest wisely for making good profits.

## **1.4 Proposed Logic / Algorithm / Business Plan / Solution / Device**

The future price of a stock is the main motivation behind the stock price prediction. In various cases like business and industry, environmental science, finance, and economics motivation can be useful. The future value of the company's stock can be determining.

## **1.5 Scope of the Project**

The future price of a stock is the main motivation behind the stock price prediction. In various cases like business and industry, environmental science, finance, and economics motivation can be useful. The future value of the company's stock can be determining.

Chapter 2 contains a literature survey that provides a summary of individual paper.

Chapter 3 provides an overview of existing work for stock price prediction that has been done using LSTM and .

Chapter 4 presents Implementation and its results, tools and technology used to achieve this and dataset detail.

Chapter 5 contains a conclusion about stock price prediction and future work about what you are wanted to do in future.

# Chapter 2

## Software Requirement Specification

---

### 2.1 Overall Description

Businesses primarily run over customer's satisfaction, customer reviews about their products. Shifts in sentiment on social media have been shown to correlate with shifts in stock markets. Identifying customer grievances thereby resolving them leads to customer satisfaction as well as trustworthiness of an organization.

Hence there is a necessity of an unbiased automated system to classify customer reviews regarding any problem. In today's environment where we are justifiably suffering from data overload (although this does not mean better or deeper insights), companies might have mountains of customer feedback collected; but for mere humans, it is still impossible to analyze it manually without any sort of error or bias. Oftentimes, companies with the best intentions find themselves in an insights vacuum. You know you need insights to inform your decision making and you know that you are lacking them, but don't know how best to get them.

Sentiment analysis provides some answers into what the most important issues are, from the perspective of customers, at least. Because sentiment analysis can be automated, decisions can be made based on a significant amount of data rather than plain intuition.

#### 2.1.1 Product Perspective

Businesses primarily run over customer's satisfaction, customer reviews about their products. Shifts in sentiment on social media have been shown to correlate with shifts in stock markets. Identifying customer grievances thereby resolving them leads to customer satisfaction as well as trustworthiness of an organization.



#### 2.1.1.1 System Interfaces

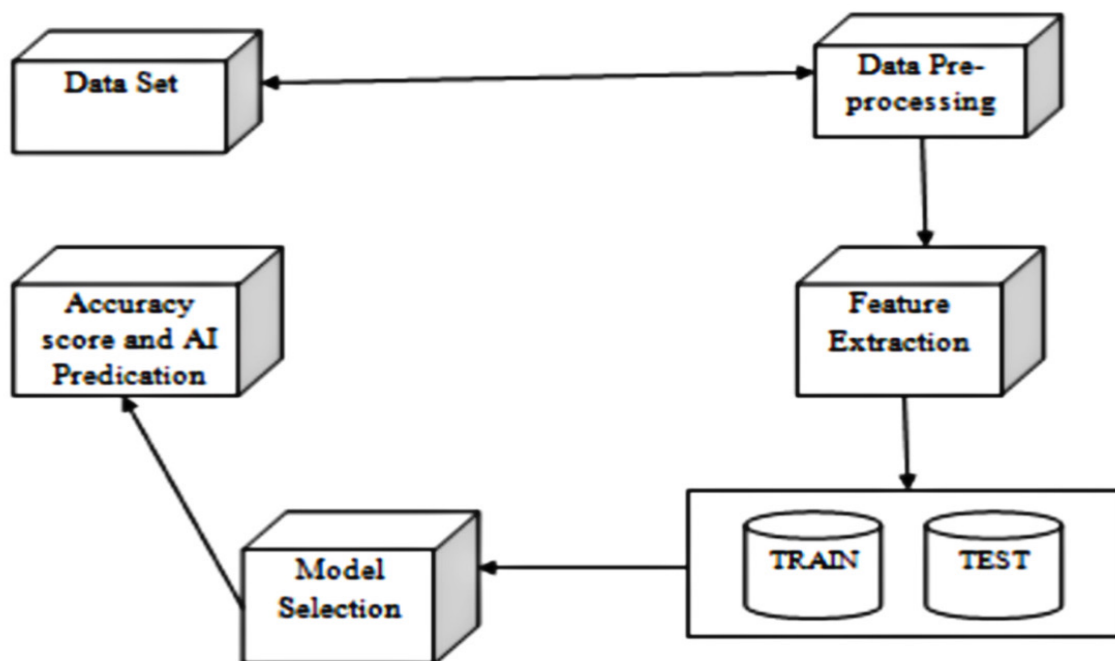


Figure 2.1: System Interface

#### 2.1.1.2 User Interfaces

The user interface for the system shall be compatible with any type of web browser such as Mozilla Firefox, Google Chrome, and Internet Explorer.

#### 2.1.1.3 Software Interfaces

##### Web Server

- Yahoo finance, NASDAQ , OS (Windows)

##### Development End

- Python, Python libraries, HTML, XML, JavaScript, OS(Windows))

#### 2.1.1.4 Communications Interfaces

The System shall be using HTTP/HTTPS for communication over the Internet and for intranet communications, it shall use TCP/IP protocol.

#### 2.1.1.5 User Characteristics

There are 3 user Levels in our Stock Price Prediction System:

1. Input as Dataset
2. Pre processing
3. Data splitting
4. Build Model train LSTM
5. Output as Predicted Result

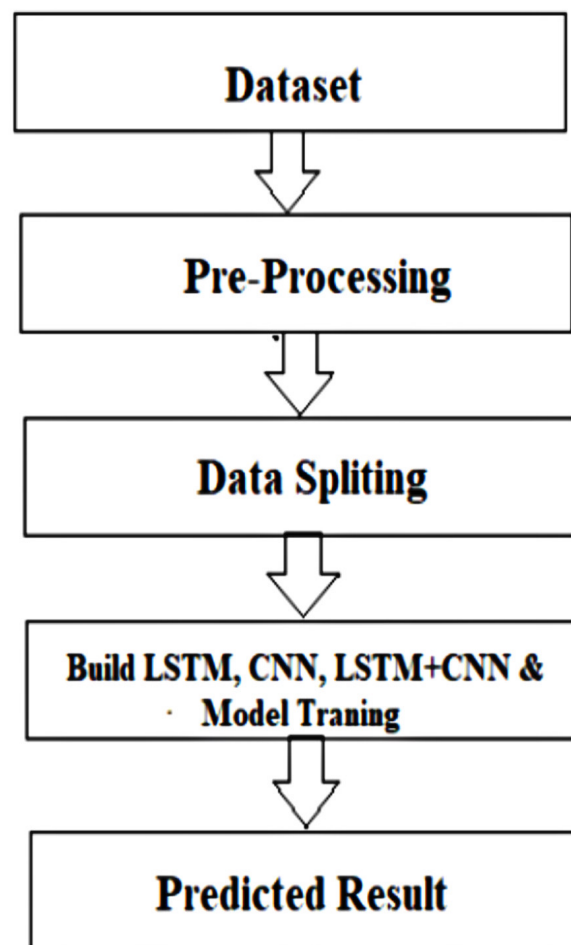
Attributes such as: price of open, high, low, close, adjusted close price taken from a huge dataset are fed as input to the models for training to pre-process the data techniques like normalization one hot encoding applied on dataset. After this data is divided into two sets namely training testing which are ratios of 80:20 respectively. Then, this set is used to train a model using 3 different approaches: LSTM. Finally, all these modules are evaluated using Root mean square error.

# Chapter 3

## System Design Specification

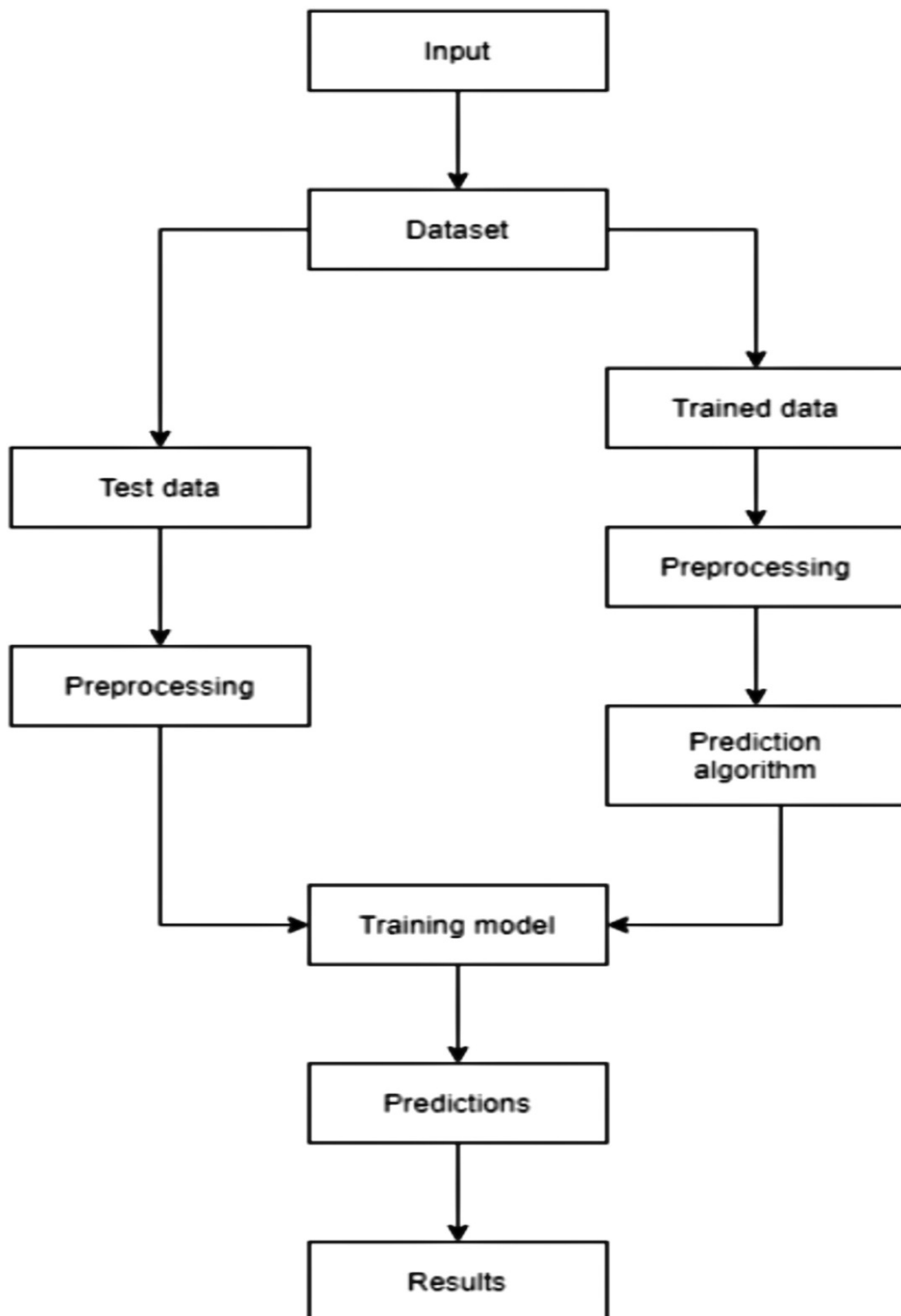
---

### 3.1 System Architecture



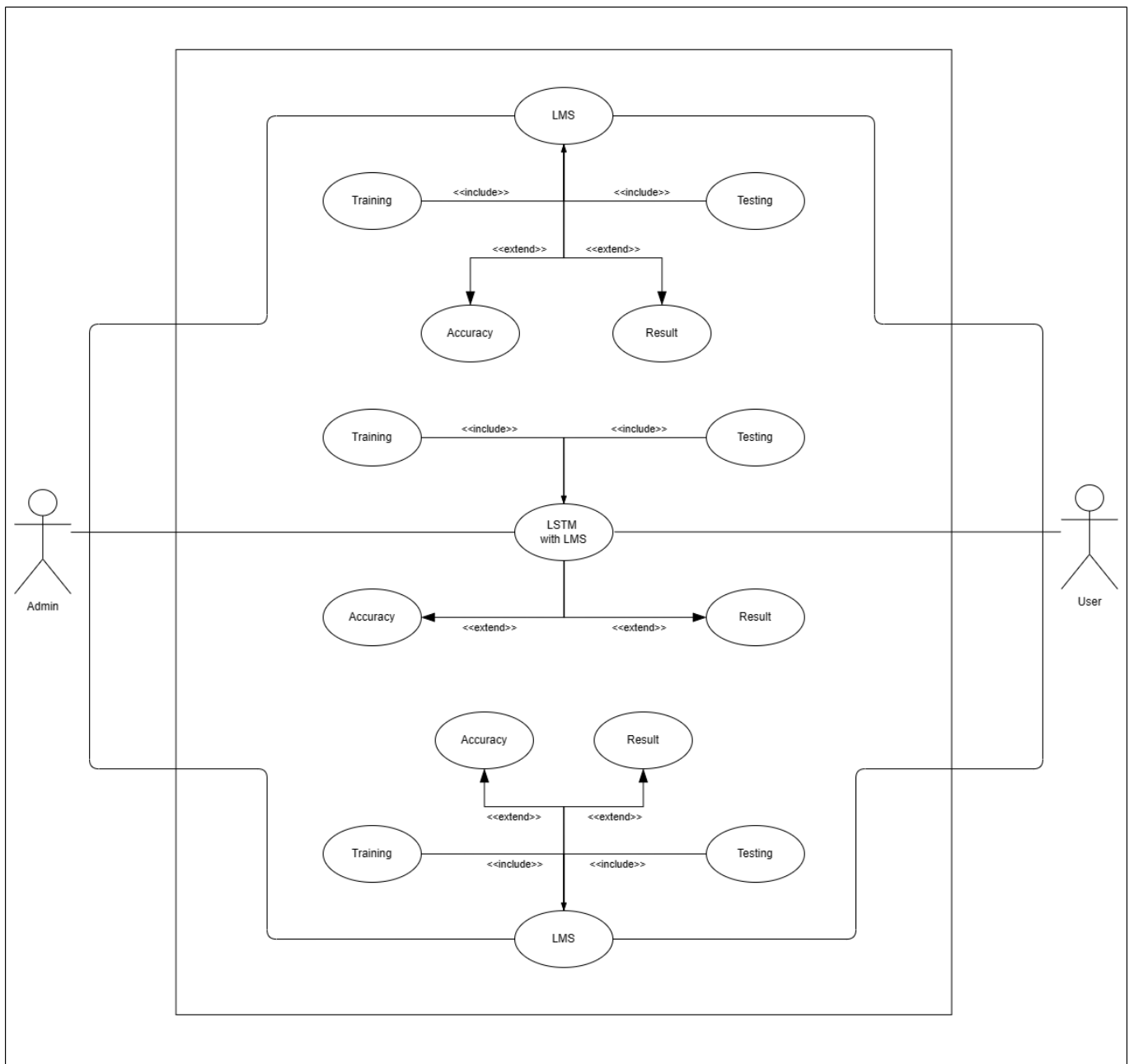
**Figure 3.1:** System Architecture

## 3.2 High Level Design Diagrams



**Figure 3.2:** High Level Design Diagram

### 3.2.1 Use Case Diagram



**Figure 3.3:** Use Case diagram

### 3.2.2 Activity Diagram

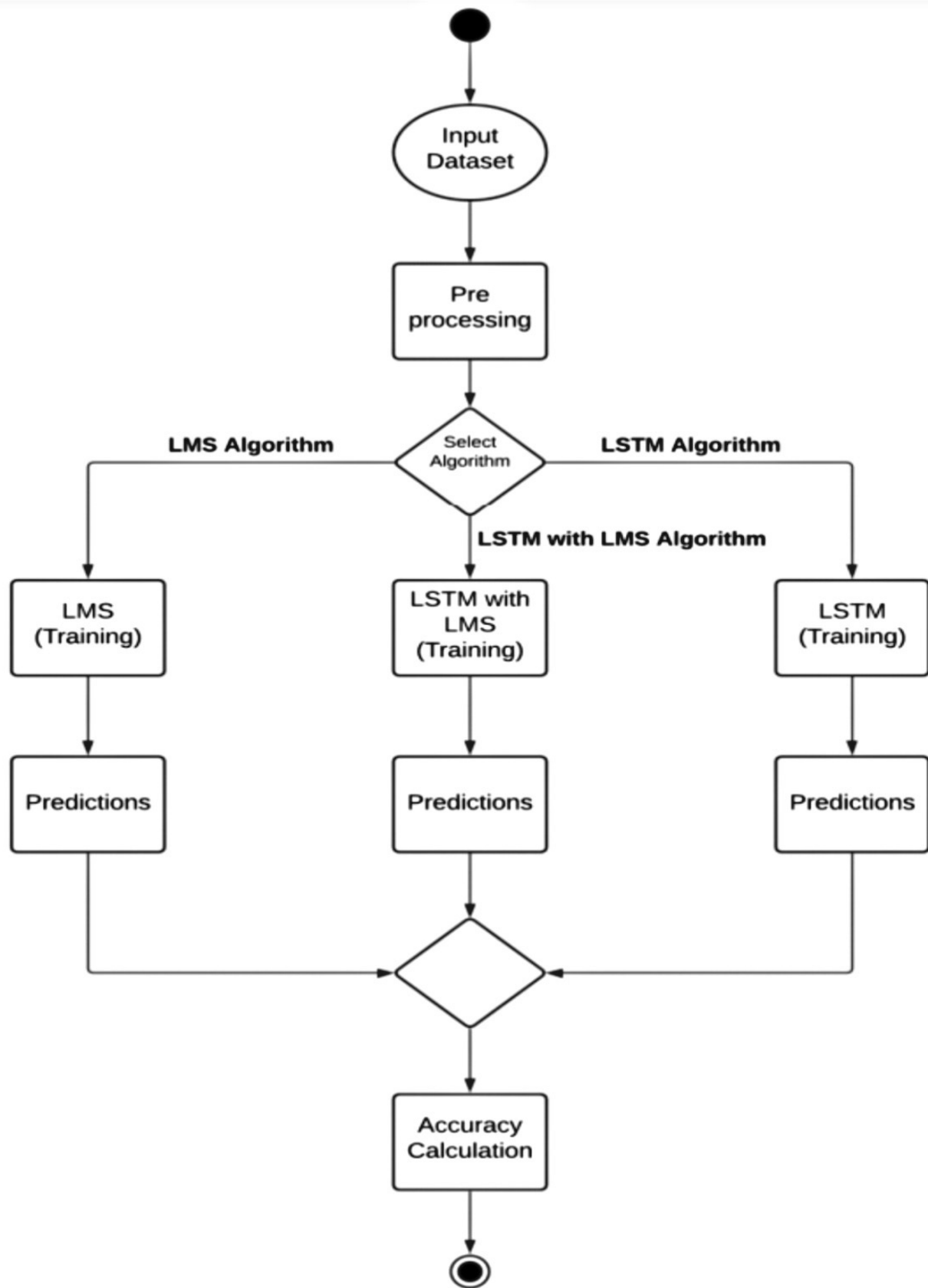
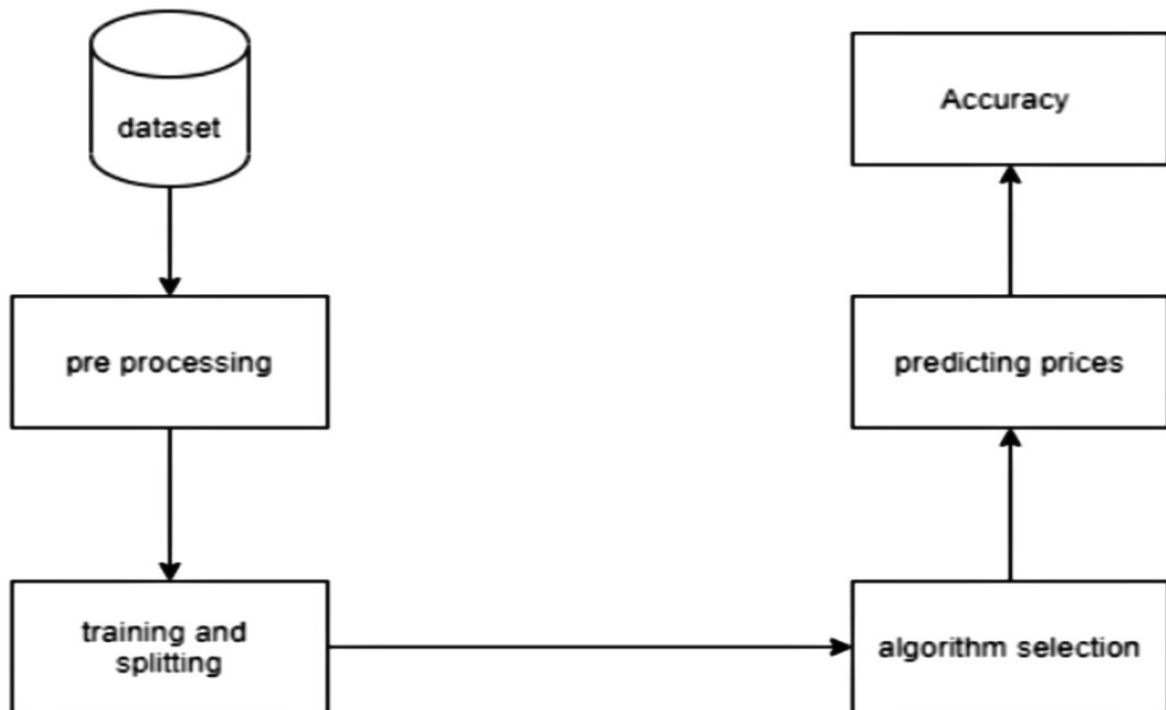


Figure 3.4: Activity Diagram

### 3.2.3 Data-Flow Diagram



**Figure 3.5:** Data flow diagram

### 3.2.4 Sequence Diagram

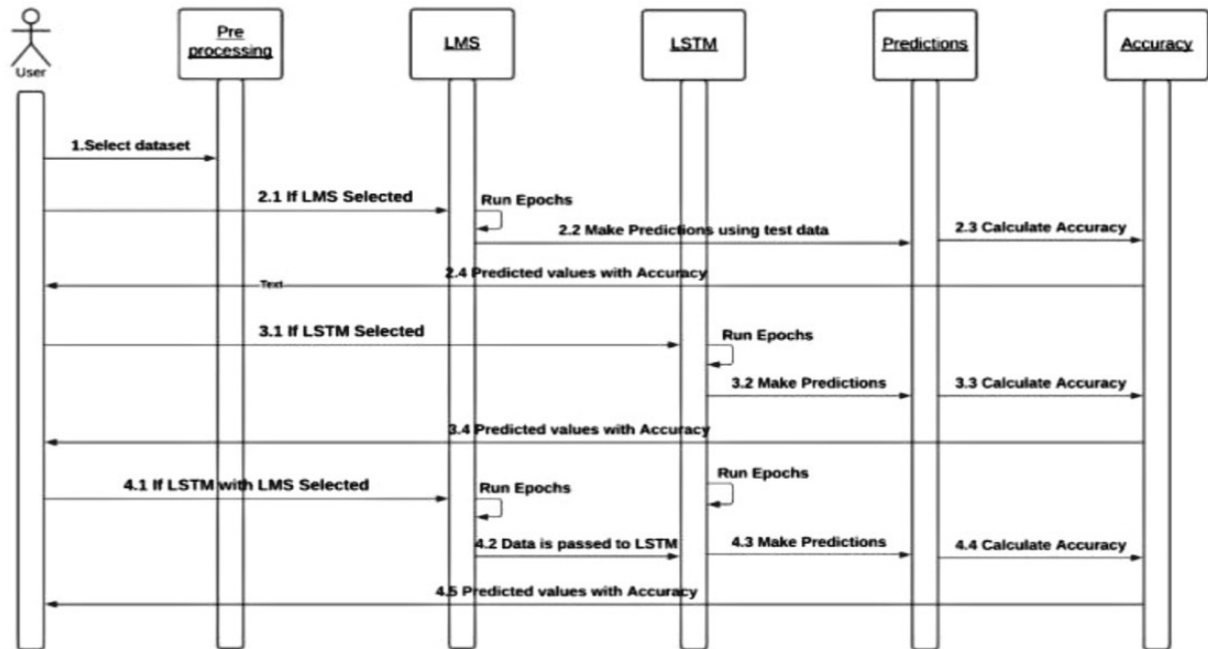


Figure 3.6: Sequence diagram



# Chapter 4

## Methodology and Team

---

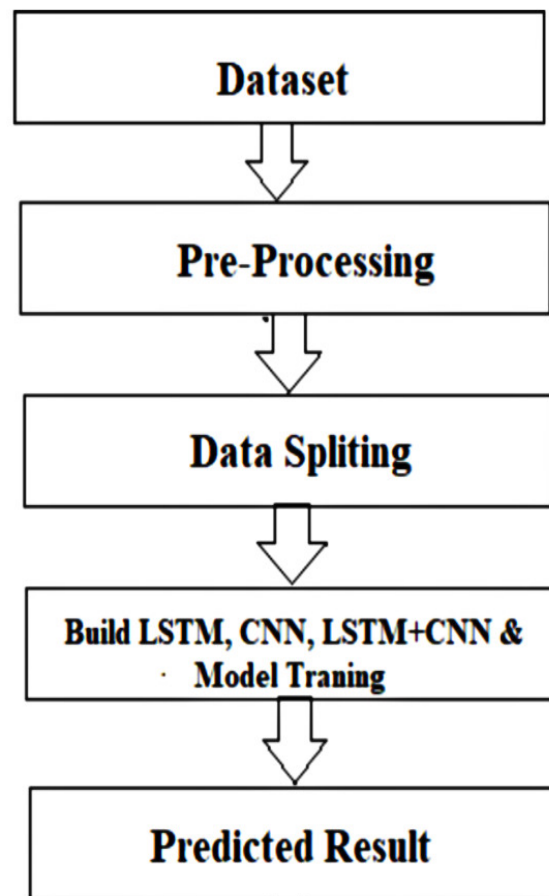
### 4.1 Overview of Existing Work

Stock Price Prediction by Machine Learning present to estimate the stock future value and machine learning technique like LSTM for existing work. This machine-learning algorithm is to perform the best predicting result of the stock future price. LSTM is capable to catching the modifications in the behavior of the stock price for the indicated period in this proposed system.

Propose a machine learning-based normalization for stock price prediction. The dataset utilized for analysis was selected from Yahoo Finance. It consists of approximately 9 lakh records of the required Stock price and other relevant data. The data reflected the stock price at some time intervals for every day of the year. It contains various data like date, symbol, open price, close price, low price, high price and volume. Here, the data for only one company was considered. All the data was available in a file of CSV format which was first read and transformed into a data frame using the Pandas library in Python. The normalization of the data was performed through the sklearn library in Python and the data were divided into training and testing sets. The experiment set was kept as 20% of the available dataset. This paper focuses on two architecture Regression based Model and LSTM. The

Regression-based Model is employed for predicting unbroken values through some given autonomous values. Regression uses a given linear function for predicting continuous values of the most important amongst them and made the predictions using these. LSTM architecture is able to identify the changes in trends which show evident from the result. LSTM is identified as the best model for the proposed methodology. This shows that the proposed system is capable of identifying some interrelation within the data. In the stock market, there may not always follow the same cycle or may not always be in a regular pattern for the changes that are occurred. The period of the existence will differ and the existence of the trend is based on the companies and the sectors. For investors, this type of analysis of trends and cycles will obtain more profit. We must use networks like LSTM as they rely on the current information to analyze various information.

## 4.2 Proposed Work



**Figure 4.1:** Proposed Workflow

The system presented here composes of five modules: -

1. Input as Dataset
2. Pre processing
3. Data splitting
4. Build & Model train LSTM
5. Output as Predicted Result

Attribute such as: price of open, high, low, close, adjusted close price taken from huge dataset are fed as input to the models for training to

pre-process the data techniques like normalization & one hot encoding in applied on dataset. After this data is divided in two sets namely training & testing which are ratio of 80:20 respectively. Then, this set are used to train a model using: LSTM. Finally, all these are evaluated using Root mean square error.

#### 4.2.1 Working of LSTM model

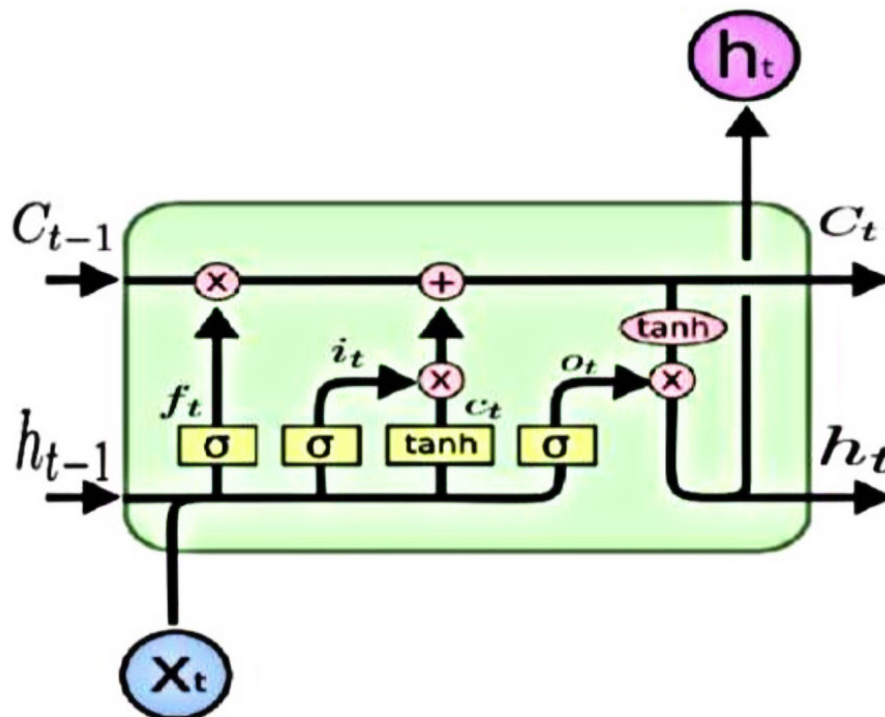


Figure 4.2: LSTM model

Long Short-Term Memory is a kind of recurrent neural network. In RNN output from the last step is fed as input within the present step. It tackled the matter of long-term dependencies of RNN within which the RNN will not predict the word hold on within the long-term memory however can offer additional accurate forecasts from the recent

info. Because the gap length will increase RNN does not offer an economical performance. LSTM will by default retain the knowledge for a long period of time. It is used for processing, predicting, and classifying on the basis of time-series data.

### **Structure of LSTM:**

- LSTM has a chain organization that contains four neural networks and different memory blocks called cells.
- LSTM has a new structure called a memory cell. The memory cell makes the decisions about what information to store, and when to allow reading, writing and forgetting.
- A memory cell contains three main gates:
  1. Input gate- a new value flows into the memory cell.
  2. Forget gate- a value remains in the memory cell.
  3. Output gate- value in the memory cell is used to compute the output.

### **Applications of LSTM includes:**

- Language Modelling
- Machine Translation
- Image Captioning
- Handwriting generation
- Question Answering Chatbot

### 4.2.2 Working of model:

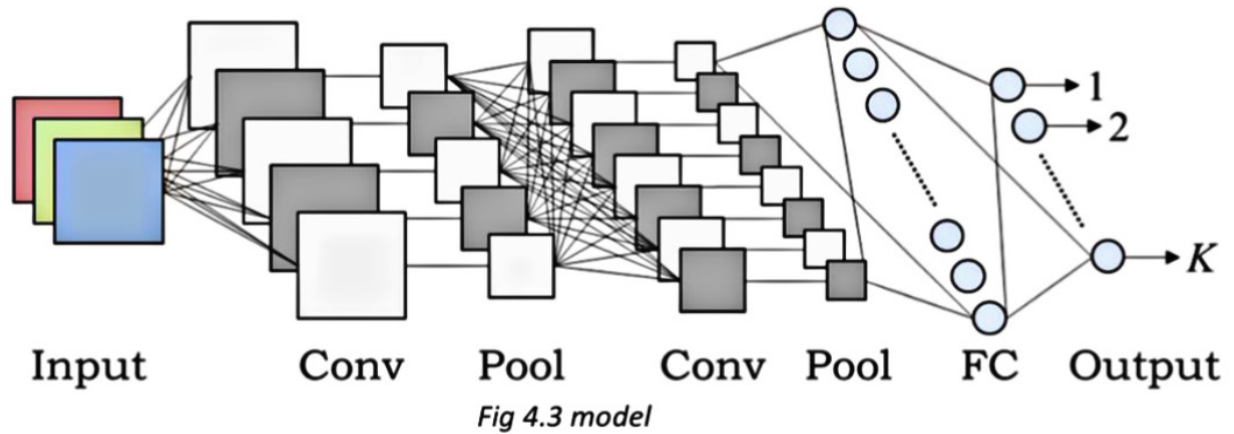


Figure 4.3: Model

- Layer of model:

1. Dropout
2. Flatten
3. Dense
4. Activation

- Dropout: Dropout is randomly selected neurons are ignored during training.
- Dense: A Linear operation in which every input is connected to every output by weight. It followed by a nonlinear activation function.
- Activation: It used sigmoid function and predict the probability 0 and 1. Applications of includes:

- Analyzing Documents

### **4.3 Team Members, Roles & Responsibilities**

Team Member 1:- Aditya Baghela

Roll No:- 19ESKIT004

#### **Responsibility:-**

- Work with project team to develop the project plan and identify resources needed to make project.
- Manage the communication and change required to ensure the project success.
- Works in the documentation of project report.
- Made some part of the project.
- Work collaboratively with other team members towards achieving common project goals/objectives.
- Helps in collecting the dataset required for project.

Team Member 2:- Anosh Field

Roll No:- 19ESKIT010

#### **Responsibility:-**

- Review, and be responsible for, the overall work effort and deliverables of the project.
- Works in the documentation of SRS

- Helps in testing part.
- Work collaboratively with other team members towards achieving common project goals/objectives.
- Helps in making Front end part of the project
- Understand the fundamentals of stock market.

Team Member 3:- Atishay Jain

Roll No:- 19ESKIT020

**Responsibility:-**

- Review test strategies and verify that all areas are covered (i.e. unit, functional, performance, stress, acceptance, regression etc.)
- Monitors status to ensure progress is in accordance with plans.
- Work collaboratively with other team members towards achieving common project goals/objectives.
- Works in the documentation of SRS.
- Review and validate all technical project deliverables for their specific line of expertise.
- Helps in making backend part of the project.



Team Member 4:- Dhruv Raj Naruka

Roll No:- 19ESKIT024

**Responsibility:-**

- Review, and be responsible for, the overall work effort and deliverables of the project.
- Works in the documentation of SRS
- Helps in testing part
- Helps in making backend part of the project.
- Analyze and define technical requirements.
- Review and validate all technical project deliverables for their specific line of expertise.

# Chapter 5

## Centering System Testing

---

The designed system has been testing through following test parameters.

### 5.1 Functionality Testing

In testing the functionality of the web sites the following features were tested:

#### 1. Links

- (a) Internal Links: All internal links of the website were checked by clicking each link individually and providing the appropriate input to reach the other links within.
- (b) External Links: Till now no external links are provided on our website but for future enhancement we will provide the links to the candidate's actual profile available online and link up with the elections updates online etc
- (c) Broken Links : Broken links are those links which so not divert the page to specific page or any page at all. By testing the links on our website, there was no link found on clicking which we did not find any page.

## 2. Forms

- (a) Error message for wrong input : Error messages have been displayed as and when we enter the wrong details (eg. Dates), and when we do not enter any details in the mandatory fields. For example: when we enter wrong password we get error message for acknowledging us that we have entered it wrong and when we do not enter the username and/or password we get the messages displaying the respective errors. All the mandatory fields have been marked with a red asterisk (\*) and apart from that there is a display of error messages when we do not enter the mandatory fields. For example: As the first name is a compulsory field in all our forms so when we do not enter that in our form and submit the form we get an error message asking for us to enter details in that particular field.
- (b) Optional and Mandatory fields : All the mandatory fields have been marked with a red asterisk (\*) and apart from that there is a display of error messages when we do not enter the mandatory fields. For example: As the first name is a compulsory field in all our forms so when we do not enter that in our form and submit the form we get an error message asking for us to enter details in that particular field. Error messages have been displayed as and when we enter the wrong details (eg. Dates), and when we do not enter any details in the mandatory fields. For example: when we enter wrong password we get error message for acknowledging us that we have entered it wrong

and when we do not enter the username and/or password we get the messages displaying the respective errors.

### 3. Database

Testing is done on the database connectivity.

## 5.2 Performance Testing

The results from analysis shows that Neural Works Predict offer the ability to predict the stock prices more accurately than the other existing tools and techniques. The accuracy of the predicted output values that lie within 20% of their corresponding target output value. By using this tool one can forecast the stock price of NSE more accurately. This analysis can be used to reduce the error percentage in predicting the future stock prices. It increases the chances for the investors to predict the prices more accurately by reducing the error percentage and hence increase their profit in share markets. Utilizing neural network models together with other forecasting tools and techniques can be considered yet another valuable advancement in the age of technology.

# Chapter 6

## Test Execution Summary

---

Execution Test Summary Report is an overall view of Testing Process from start to end. Test Plan comes at the starting of project while Test Summary Report comes at the end of the testing process. This report is given to the client for his understanding purpose.

### Test Cases

The ML model is designed in such a way that it can be used for prediction of stock for any company. This is done by providing the stock ticker of that company and feeding those in the website that is made using Streamlit module of Python. Stock Tickers like GOOGL for Google, META for Facebook and AAPL for APPLE were entered to predict stock prices.

And in each case the model gave a pretty satisfactory result with high accuracy. The model did not have any problem of underfitting or overfitting.

Screenshots of these test cases are given in Chapter 7.

# Chapter 7

## Project Screen Shots

---

### 7.1 Dataset Detail

The dataset consists of the stock historical data from the Yahoo Finance Website and captures the daily information of each stock. It collects different sectors of stock data, including Banking, Pharma, Petroleum, Software and Textiles etc. and it includes the opening price, the highest price, the lowest price, the closing price, the adjusted closing price, and the volume of stock.

### 7.2 Tool Technologies

#### 7.2.1 Python

The language of select for this project was Python. This was a straightforward call for many reasons:-

1. Python as a language has a vast community behind it. Any problems which may be faced is simply resolved with visit to Stack Overflow. Python is the foremost standard language on the positioning that makes it is very straight answer to any question.
2. Python is an abundance of powerful tools ready for scientific computing Packages. The packages like NumPy, Pandas and SciPy area unit freely available and well documented. These Packages will intensely scale back, and variation the code necessary to write a given

program. This makes repetition fast.

3. Python is a language as forgiving and permits for the program that appear as if pseudo code. This can be helpful once pseudo code give in tutorial papers should be required and verified. Using python this step is sometimes trivial. However, Python is not without its errors. The python is dynamically written language and packages are area unit infamous for Duck writing. This may be frustrating once a package technique returns one thing that, for instance, looks like an array instead of being an actual array. Plus, the standard Python documentation did not clearly state the return type of a method, this can't lead without a lot of trials and error testing otherwise happen in a powerfully written language. This is a problem that produces learning to use a replacement Python package or library more difficult than it otherwise may be.

### **7.2.2 Numpy**

NumPy is python package which provide scientific and higher level mathematical abstractions wrapped in python. It is the core library for scientific computing, that contains a provide tools for integrating C, strong n-dimensional array object, C++ etc. It is also useful in random number capability, linear algebra etc. NumPy's array type augments the Python language with an efficient data structure used for numerical work, e.g., manipulating matrices. NumPy additionally provides basic numerical routines, like tools for locating Eigenvectors.

### 7.2.3 Scikit Learn

Scikit-learn could be a free machine learning library for Python. It features numerous classification, clustering and regression algorithms like random forests, k-neighbors, support vector machine, and it furthermore supports Python scientific and numerical libraries like SciPy and NumPy. In Python Scikit-learn is specifically written, with the core algorithms written in Python to get the performance. Support vector machines are enforced by a Python wrapper around LIBSVM.i.e., linear support vector machines and logistic regression by a similar wrapper around LIBLINEAR.

### 7.2.4 Tensorflow

TensorFlow has an open-source software library for numerical computation using data flow graphs. Inside the graph nodes represent mathematical formulae, the edges of graph represent the multidimensional knowledge arrays (tensors) communicated between them. The versatile architecture permits to deploy the computation to at least one or many GPUs or CPUs in a desktop, mobile device, servers with a single API. TensorFlow was firstly developing by engineers and researchers acting on the Google Brain Team at intervals Google's Machine Intelligence analysis organization for the needs of conducting deep neural networks research and machine learning, but, the system is generally enough to be appropriate in a wide range of alternate domains as well.

Google Brain's second-generation system is TensorFlow. Whereas the



reference implementation runs on single devices, TensorFlow can run on multiple GPUs and CPUs. TensorFlow is offered on Windows, macOS, 64-bit Linux and mobile computing platforms together with iOS and Android.

#### **7.2.5 Keras**

Keras is a high-level neural networks API, it is written in Python and also capable of running on top of the Theano, CNTK, or TensorFlow. It was developed with attention on enabling quick experimentation. Having the ability to travel from plan to result with the smallest amount doable delay is key to doing great research. Keras permits for straightforward and quick prototyping (through user-friendliness, modularity, and extensibility). Supports each recurrent networks and convolutional networks, also as combinations of the 2. Runs seamlessly on GPU and CPU. The library contains numerous implementations of generally used neural network building blocks like optimizers, activation functions, layers, objectives and a number of tools to create operating with text and image data easier. The code is hosted on GitHub, and community support forums embody the GitHub issues page, a Gitter channel and a Slack channel.

#### **7.2.6 Compiler Option**

Anaconda is free premium open-source distribution of the R and Python programming languages for scientific computing, predictive analytics, and large-scale process that aim is to modify package managing and deployment. Package versions unit managed by the package manage-

ment system conda.

### 7.2.7 Jupyter Notebook

The Jupyter Notebook is an open-source web application that enables to making and sharing documents that contain visualizations, narrative text, live code and equations. Uses include: data, data visualization, data transformation, statistical modelling, machine learning, numerical simulation, data cleaning and much more.

## 7.3 Results

### Step 1: Dataset Analysis

Firstly, I have performed Data analysis for stock price of companies. Fig.represent the date, open, close, high, low, adjusted close and volume of stocks details.

### Step 2: Read Dataset

	High	Low	Open	Close	Volume	Adj Close
Date						
2009-12-31	7.619643	7.520000	7.611786	7.526071	352410800.0	6.415357
2010-01-04	7.660714	7.585000	7.622500	7.643214	493729600.0	6.515213
2010-01-05	7.699643	7.616071	7.664286	7.656429	601904800.0	6.526476
2010-01-06	7.686786	7.526786	7.656429	7.534643	552160000.0	6.422665
2010-01-07	7.571429	7.466071	7.562500	7.520714	477131200.0	6.410790

**Figure 7.1:** Read Dataset

### Step 3: Graph of Close Price history

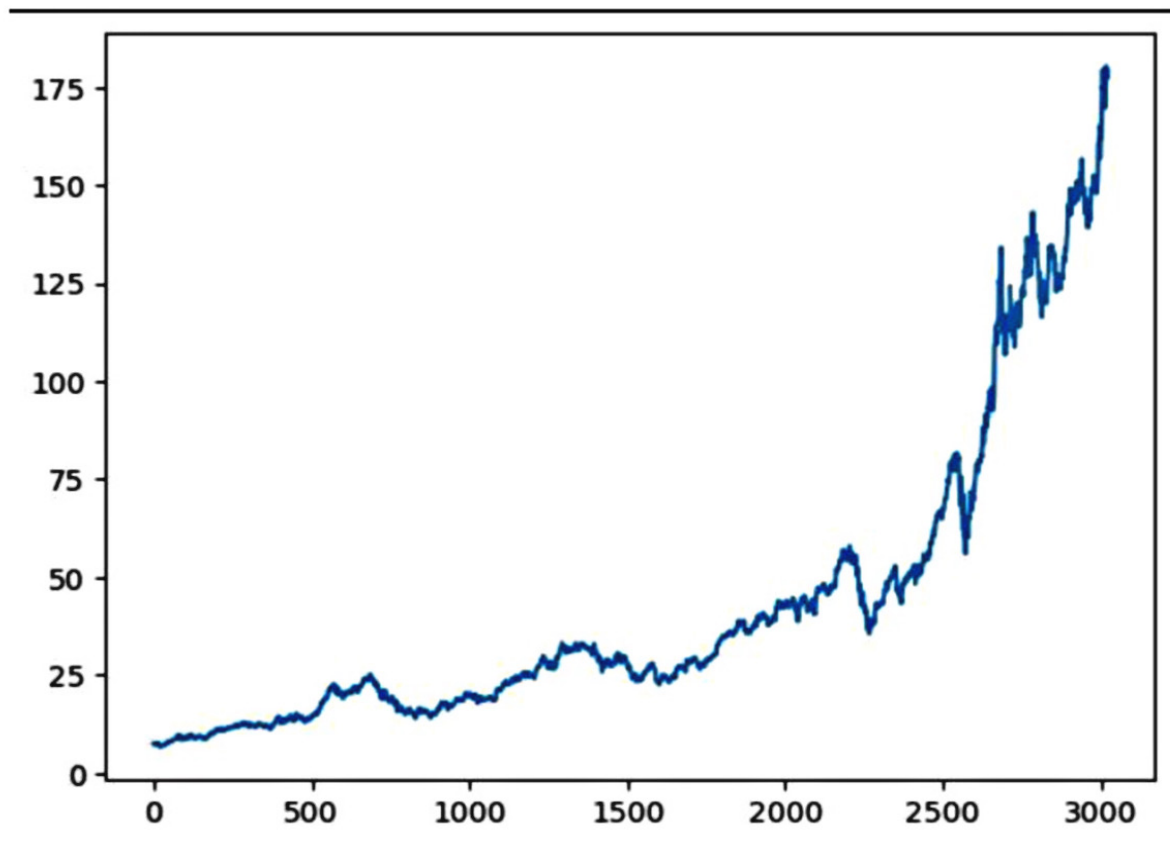


Figure 7.2: Graph of Close Price history

### Step4: Preprocessing

```
array([[0.01641142],  
       [0.0192931 ],  
       [0.01961817],  
       ...,  
       [0.98966821],  
       [0.98837669],  
       [0.9910212 ]])
```

Figure 7.3: Data Scaling

After Dataset reading, I have performed preprocessing operation on the dataset. Here I apply Min-Max Scaler to preprocess the dataset. In preprocessing operation removes the noise into the data and convert data into 0 to 1 form.

## Step 5: Train test Split

After performing preprocessing, I have divided the data set into training and testing set. 80% of the data is used for the training while the remaining 20% of the data is used for testing.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100, 50)	10400
dropout (Dropout)	(None, 100, 50)	0
lstm_1 (LSTM)	(None, 100, 60)	26640
dropout_1 (Dropout)	(None, 100, 60)	0
lstm_2 (LSTM)	(None, 100, 80)	45120
dropout_2 (Dropout)	(None, 100, 80)	0
lstm_3 (LSTM)	(None, 50)	26200
dropout_3 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51
Total params: 108,411		
Trainable params: 108,411		
Non-trainable params: 0		

## Step 6: Model fitting of Long Short-Term Memory

lstm_1 (LSTM)	(None, 60, 50)	10400
lstm_2 (LSTM)	(None, 50)	20200
dense_1 (Dense)	(None, 25)	1275
dense_2 (Dense)	(None, 1)	26
Total params: 31,901		
Trainable params: 31,901		
Non-trainable params: 0		

Figure 7.4: LSTM Summary

Layer (type)	Output Shape	Param #
conv1d_1 (Conv1D)	(None, 60, 64)	256
max_pooling1d_1 (MaxPooling1D)	(None, 30, 64)	0
lstm_1 (LSTM)	(None, 30, 100)	66000
dropout_1 (Dropout)	(None, 30, 100)	0
conv1d_2 (Conv1D)	(None, 30, 32)	9632
max_pooling1d_2 (MaxPooling1D)	(None, 15, 32)	0
flatten_1 (Flatten)	(None, 480)	0
dense_1 (Dense)	(None, 1)	481
activation_1 (Activation)	(None, 1)	0
Total params: 76,369		
Trainable params: 76,369		
Non-trainable params: 0		

Figure 7.5: Approach of LSTM

After generating training dataset, to apply training I have created LSTM network using KERAS. Several variations of this architecture using various numbers of layers and various size of Bottle neck layer.

## Step 7: Apply Training

```
Epoch 11/50
63/63 [=====] - 16s 252ms/step - loss: 0.0072
Epoch 12/50
63/63 [=====] - 16s 258ms/step - loss: 0.0061
Epoch 13/50
...
Epoch 49/50
63/63 [=====] - 12s 190ms/step - loss: 0.0033
Epoch 50/50
63/63 [=====] - 12s 184ms/step - loss: 0.0034

<keras.callbacks.History at 0x2023a43c160>
```

**Figure 7.6:** Training Process

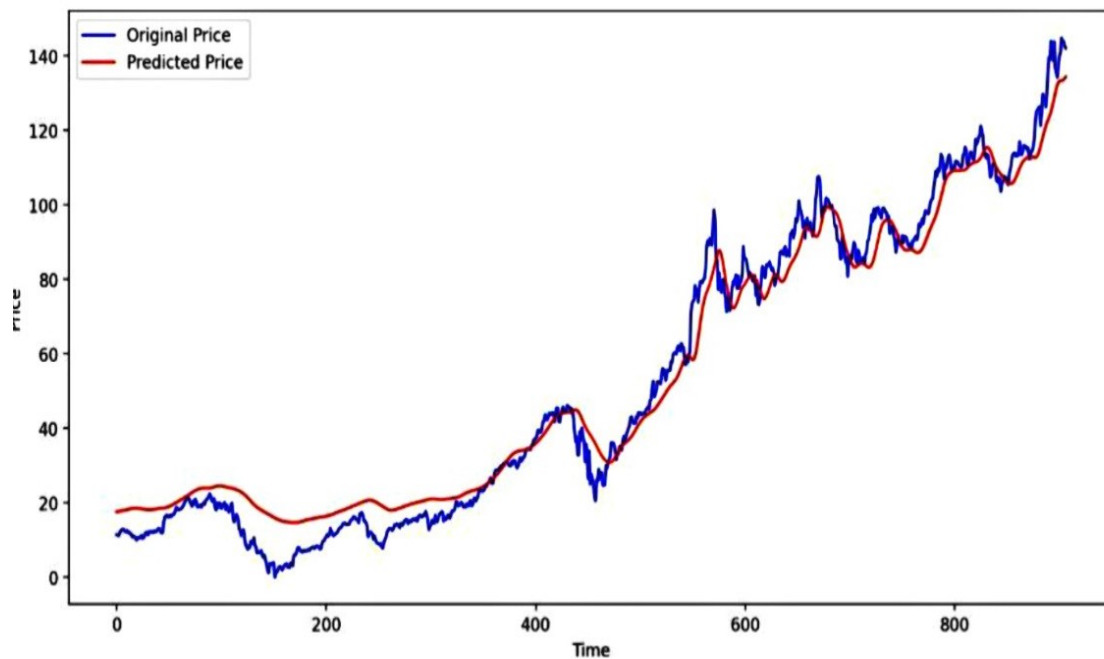
To apply training, from the samples of Training data, 1543 samples are used for training and 460 samples are used for validation. Data is processed in a batch size of 1 and epoch is 1 for the entire training dataset.

### Step8: Predicted Result

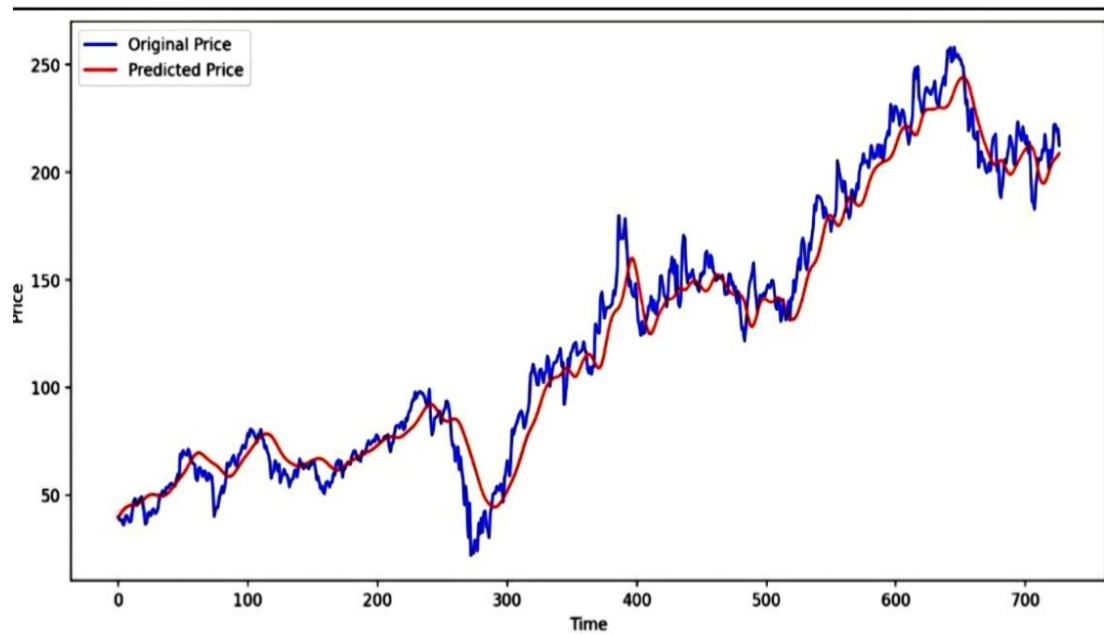
```
array([[0.10753927],
       [0.10837215],
       [0.1090636 ],
       [0.10962331],
       [0.11006683],
       [0.11042947],
       [0.11075294],
       [0.11107945],
       [0.11144108],
       [0.11185017],
       [0.11229828],
       [0.11276951],
       [0.11324644],
       [0.11370483],
       [0.11412087],
       [0.11446837],
       [0.1147238 ],
       [0.1148645 ],
       [0.11488101],
       [0.11477038],
       [0.11453363],
       [0.11417547],
       [0.11371562],
       [0.11318031],
       [0.11260271],
       [0.11201599],
       [0.11145204],
       [0.11092389],
       [0.11045027],
       [0.11005896],
       [0.10978043],
       [0.10962749],
       [0.10959509],
       [0.10968491],
       [0.10988975],
       [0.1101948 ],
       [0.11058071],
       [0.11102289],
```

**Figure 7.7:** Predicted Close Price

## Step 9: Predicted Graph

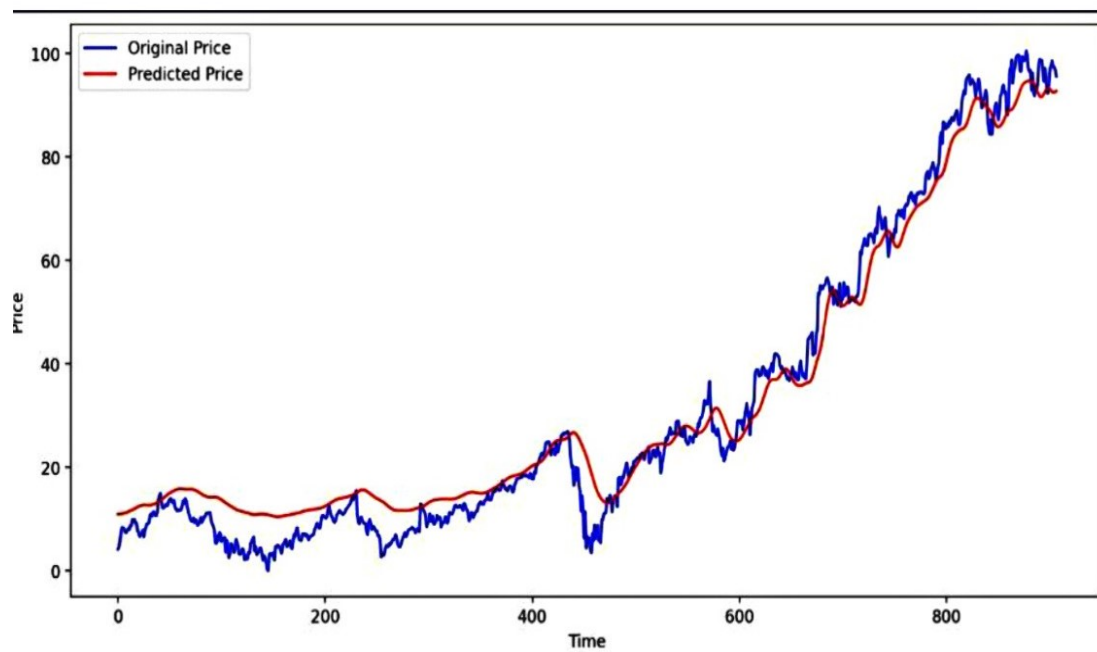


**Figure 7.8:** Plot for Real vs Predicted value for Amazon using LSTM

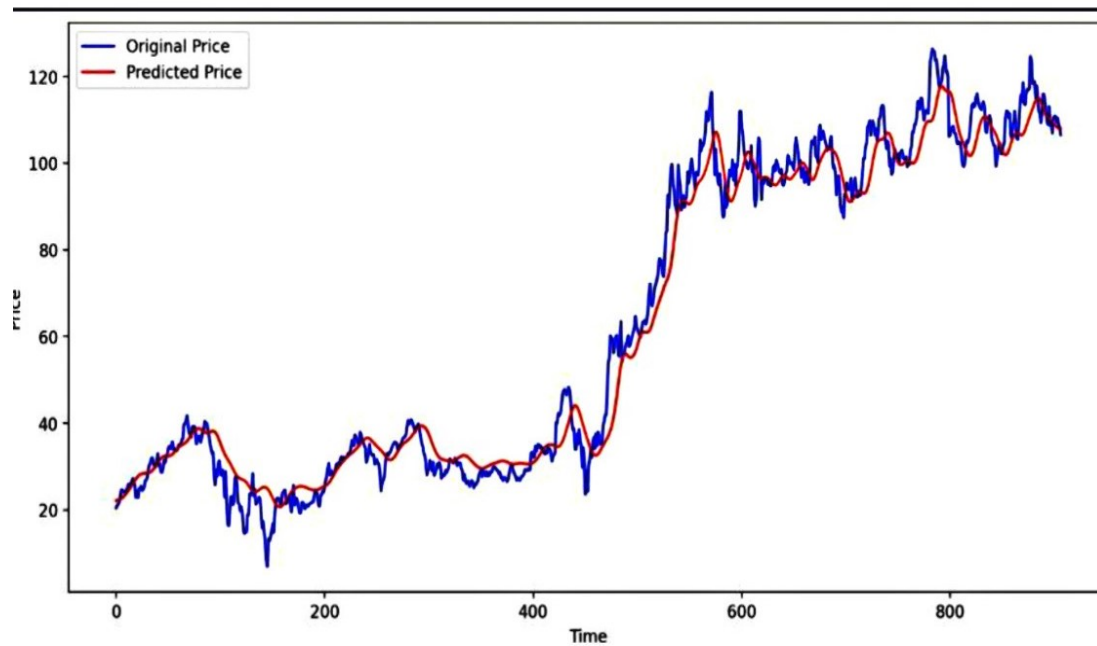


**Figure 7.9:** Plot for Real vs Predicted value for Facebook using LSTM





**Figure 7.10:** Plot for Real vs Predicted value for Google using LSTM



**Figure 7.11:** Plot for Real vs Predicted value for Amazon using LSTM

# Chapter 8

## Project Summary and Conclusions

---

In report, we will compare a machine learning models like LSTM model, the model. We have attendance to train the model using the data of NSE listed companies to predict the stock future value. This is showing the proposed method is capable to distinctive around interrelation with the data. It uses the information given at a specific instant for prediction. Even if the model LSTM are utilized in a lot of other time-dependent data analysis, it is not out performing over the architecture in this case. This is often because of quick changes occur in stock market. The changes in the stock market are not always be in a regular pattern or not always follow the continuous cycle. Based on the companies and sectors, the existence of the trends and the period of their existence will differ. The analysis of this type of cycles and trends can offer a more profit to the investors. In future work, we add more stock market data and compare more model to improve accuracy of predicted stock price.

In the future, for better accuracy model can be trained with more varied and detailed data.

# Chapter 9

## References

---

- [1] “Stock price prediction using LSTM, RNN and -sliding window model - IEEE Conference Publication.” <https://ieeexplore.ieee.org/document/8126078> (accessed Dec. 27, 2019).
- [2] J. Jagwani, M. Gupta, H. Sachdeva, and A. Singhal, “Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend,” in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, Jun. 2018, pp. 462–467, doi: 10.1109/ICCONS.2018.8663035.
- [3] I. Parmar et al., “Stock Dec. 2018, pp. 574–576, doi: 10.1109/IC-SCCC.2018.8703332.
- [4] Y. Lei, K. Zhou, and Y. Liu, “Multi-Category Events Driven Stock Price Trends Prediction,” in 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Nanjing, China, Nov. 2018, pp. 497– 501, doi: 10.1109/CCIS.2018.8691392.
- [5] B. Jeevan, E. Naresh, B. P. V. kumar, and P. Kambli, “Share Price Prediction using Machine Learning Technique,” in 2018 3rd International Conference on Circuits, Control, Communication and Comput-

ing (I4C), Bangalore, India, Oct. 2018, pp. 1–4, doi: 10.1109/CIMCA.2018.8739647.

[6] M. Usmani, S. H. Adil, K. Raza, and S. S. A. Ali, “Stock market prediction using machine learning techniques,” in 2016 3rd International Conference on computer and Information Sciences (ICCOINS), 2016, pp. 322–327.