

## MINI PROJECT-II

(2018-19)

# Stock Price Prediction



## Software Requirement Specifications

## Institute of Engineering & Technology

### Team Members

Aryavrat

(University RollNo: 161500133)

Ashish Varshney

(University RollNo: 161500143)

Raghav Arora

(University RollNo: 161500421)

Rakhi Agrawal

(University RollNo: 161500436)

### *Mini Project Guide*

**Mr. Vivek Kumar**

*Astt. Professor*

*Department of Computer Engineering & Applications*

## Table of Contents

<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 PURPOSE .....	1
1.2 SCOPE .....	1
1.3 REFERENCES .....	1
1.4 Overview.....	2
<b>2. GENERAL DESCRIPTION.....</b>	<b>3</b>
2.1 PRODUCT PERSPECTIVE.....	3
2.2 PRODUCT FUNCTIONS .....	3
2.3 USER CHARACTERISTICS.....	3
2.4 GENERAL CONSTRAINTS .....	3
2.5 ASSUMPTIONS AND DEPENDENCIES .....	4
<b>3. SPECIFIC REQUIREMENTS.....</b>	<b>5</b>
3.1 EXTERNAL INTERFACE REQUIREMENTS .....	5
3.1.1 <i>Software Requirements</i> .....	5
3.1.2 <i>Hardware Requirements</i> .....	5
3.1.3 <i>Software Interfaces</i> .....	5
3.2 USE CASES .....	6
3.2.1 <i>Use Case #1</i> .....	6
3.4 Non-Functional Requirements.....	8
3.4.1 <i>Performance</i> .....	8
3.4.2 <i>Reliability</i> .....	8
3.4.3 <i>Safety</i> .....	8
3.4.4 <i>Usability</i> .....	8

3.5 Design Constraints .....	8
<b>4. ANALYSIS MODELS.....</b>	<b>9</b>
4.1 DATA FLOW DIAGRAMS (DFD).....	9
4.2 FLOW CHART OF ARIMA MODEL .....	10
<b>5. CHANGE MANAGEMENT PROCESS.....</b>	<b>11</b>

## 1. Introduction

### 1.1 Purpose

This SRS describes the software functional and nonfunctional requirements for the project **Stock Price Prediction**, a Machine Learning based project. This document is intended to be used by the members of the project team that will implement and analysis the correct functioning of the scripts.

### 1.2 Scope

The project deals with the financial data of companies and on the basis of past stock return it will predict the future trend in the field of finance and help user to produce more accurate forecasting of future stock returns with the help of different models and with every model the rate of error also decreases.

### 1.3 References

Project Development	<ol style="list-style-type: none"><li>1. <a href="http://www.Udemy.com">www.Udemy.com</a></li><li>2. <a href="http://www.youtube.com">www.youtube.com</a></li><li>3. <a href="http://www.stackoverflow.com">www.stackoverflow.com</a></li></ol>
Dataset	<ol style="list-style-type: none"><li>1. <a href="https://www.google.com/finance">https://www.google.com/finance</a></li><li>2. <a href="https://www.quandl.com">https://www.quandl.com</a></li></ol>

#### **1.4 Overview**

- (1) All the functional/non-functional requirements their corresponding DFDs, Flow Charts and Use case diagrams have been organized in the following document.
- (2) The general description of the project followed by the functionalities has been listed initially. Later on the application has been described diagrammatically with the help of Use case, Data Flow Diagrams and Flow Charts.

## **2. General Description**

### **2.1 Project Perspective**

The aim of the project is to learn and examine different forecasting models to predict future stock returns based on past returns and numerical news indicators to construct a portfolio of multiple stocks in order to diversify the risk. We do this by applying supervised learning methods for stock price forecasting .

### **2.2 Product Functions**

The project provides the user to apply different stock forecasting model to predict future stock return from past returns. It will help in predicting the future stock return in a more accurate way with the help of Supervised learning. This project also provides the facility of graphs that will help user to analyze .

### **2.3 User Characteristics**

The project basically supports only one type of user whose task is to provide the dataset to the project for further computation. After the processing and computation on the provided dataset user will get the graph containing predicted value curve and the original value curve.

### **2.4 General Constraints**

- The application may not work if the required libraries are not installed.
- May not function if the provided dataset contains different attribute.
- May not function if the data set file is empty or have different extension other than .csv .

## 2.5 Assumptions and Dependencies

Following are the Assumptions/Dependencies:

- The application runs on python 3.6.0 and above.
- The provide dataset file contain the following attribute:
  - Date, Open(open rate of the day), High(Highest rate of the day), Low(Lowest value of the day), Close(value at which market close), Volume (Turnover of the day), previous\_day (Average value of the previous day).
- Jupyter notebook is required to run this project and for the debugging purpose .
- Following libraries are required to run project :
  - **Pandas:** *Pandas* is the most popular python library that is used for data analysis. It provides highly optimized performance with back-end source code is purely written in C or Python. We can analyze data in pandas with Series, DataFrame.
  - **Matplotlib:** Matplotlib is an amazing visualization library in Python for 2D plots of arrays. Matplotlib is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader SciPy stack.
  - **Statsmodels:** Statsmodels is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration. An extensive list of result statistics are available for each estimator. The results are tested against existing statistical packages to ensure that they are correct.
  - **Datetime:** The datetime module supplies classes for manipulating dates and times in both simple and complex ways. While date and time arithmetic is supported, the focus of the implementation is on efficient attribute extraction for output formatting and manipulation. For related functionality, see also the time and calendar modules.

### **3. Specific Requirements**

#### **3.1 External Interface Requirements**

##### **3.1.1 Software Requirements**

The system must contains good internet connection. Jupyter notebook is required to run the project and all the required packages must be installed for the proper functioning of the algorithms. User can use anaconda IDE for running the project because it contains all the important libraries and Packaged pre-installed.

##### **3.1.2 Hardware Requirements**

There is only one hardware required for running the project. The hardware must be robust and have high processing capabilities because Anaconda IDE is very heavy and takes lot of time to run.

##### **3.1.3 Software Interfaces**

Operating System	We have chosen Windows 10.
Dataset	For easy processing we have used dataset from Google finance with .csv extension.

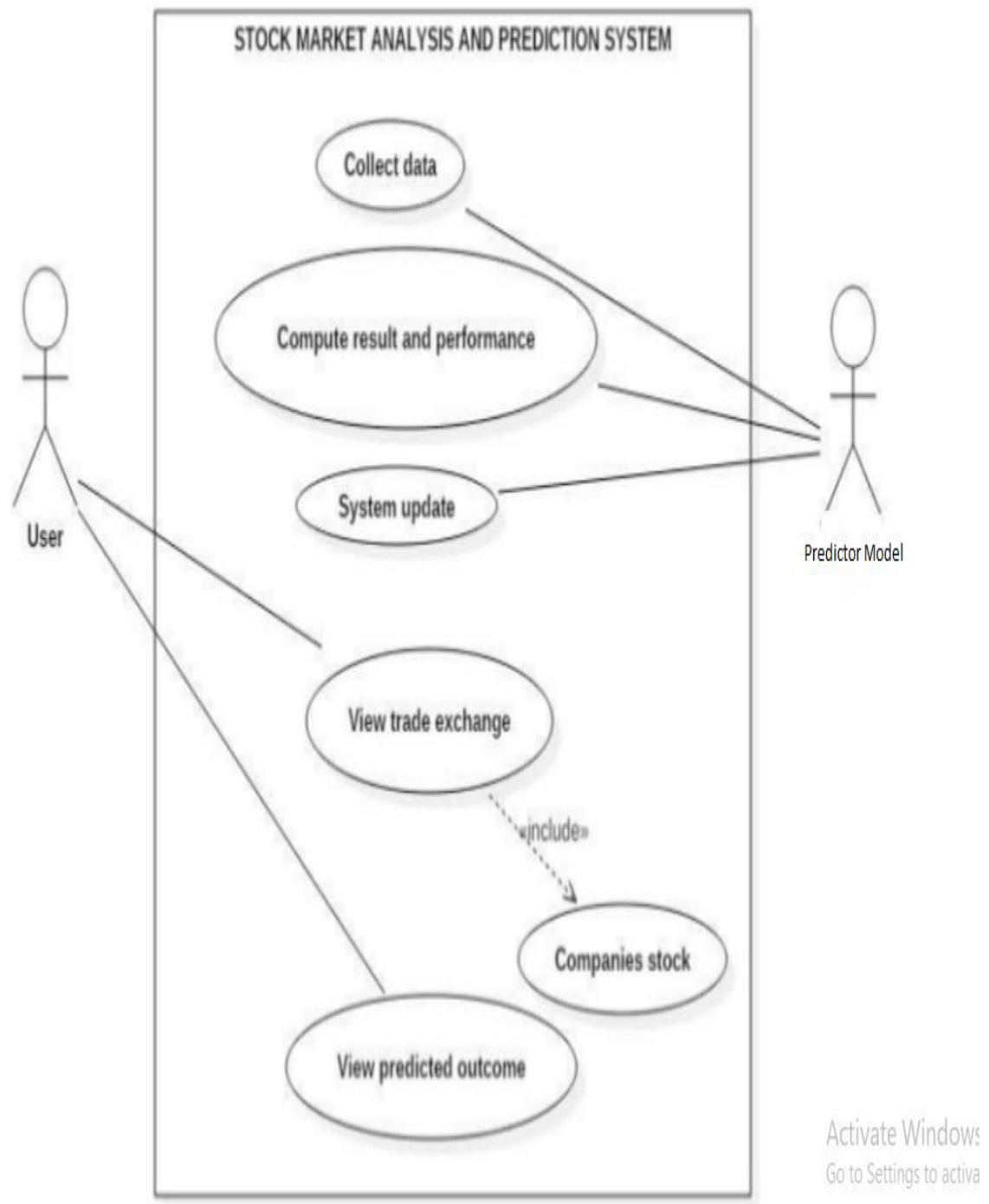


### 3.2 Use Cases

#### 3.2.1 Use Case #1

Objective	To gain knowledge of application functionality.
Primary Actors	User, Predictor Model
Pre-Conditions	To consumer has the application installed on their system and jupyter notebook with all required libraries.
Post-Conditions	Result data is displayed on the jupyter notebook.
Flow of Events	<ol style="list-style-type: none"><li>1. Predicting model collect the data and convert it into DataFeame.</li><li>2. Predicting algorithm preform computation and store RMS value and plot graph.</li><li>3. Model will update the system with recently calculated output.</li><li>4. User view company's trade.</li><li>5. User view all the plotted graphs and RMS value.</li></ol>
Exception Condition	<ol style="list-style-type: none"><li>1. If the user give invalid file format.</li><li>2. If the system doesn't have browse and required packages.</li></ol>

### 3.2.1 Use Case #1



### **3.4 Non-Functional Requirements**

#### **3.4.1 Performance**

The system must be interactive and the delays involved must be less .So in every action-response of the system, there are no immediate delays. In case of popping error messages and saving the settings or sessions there is delay much below 2 seconds. In case of opening file , normalizing dataset, training dataset and evaluation there are no delays and the operation is performed in less than 5 seconds for opening ,normalization , training, computation > 95% of the files.

#### **3.4.2 Reliability**

As the system provides the facility of forecasting it must be made sure that the system is reliable in its operations and for securing the sensitive details.

#### **3.4.3 Safety**

Information transmission should be securely transmitted to server without any changes in information.

#### **3.4.4 Usability**

As the system is easy to handle and navigates in the most expected way with no delays. In that case the system program reacts accordingly and transverses quickly between its states.

### **3.5 Design Constraints**

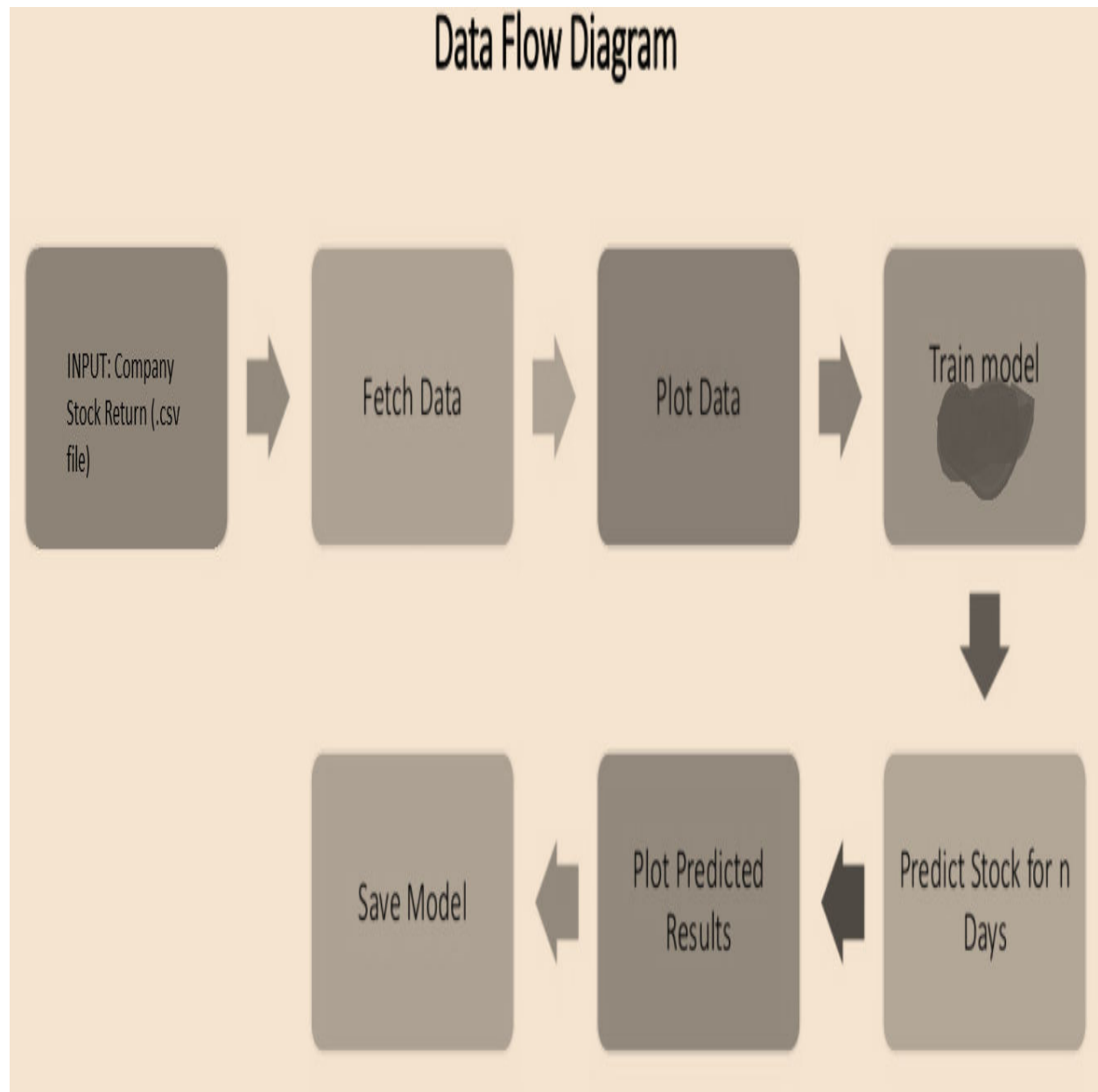
Software Language: All coding will be done in standard python.

Platform dependency: The application is compatible with jupyter notrbook.

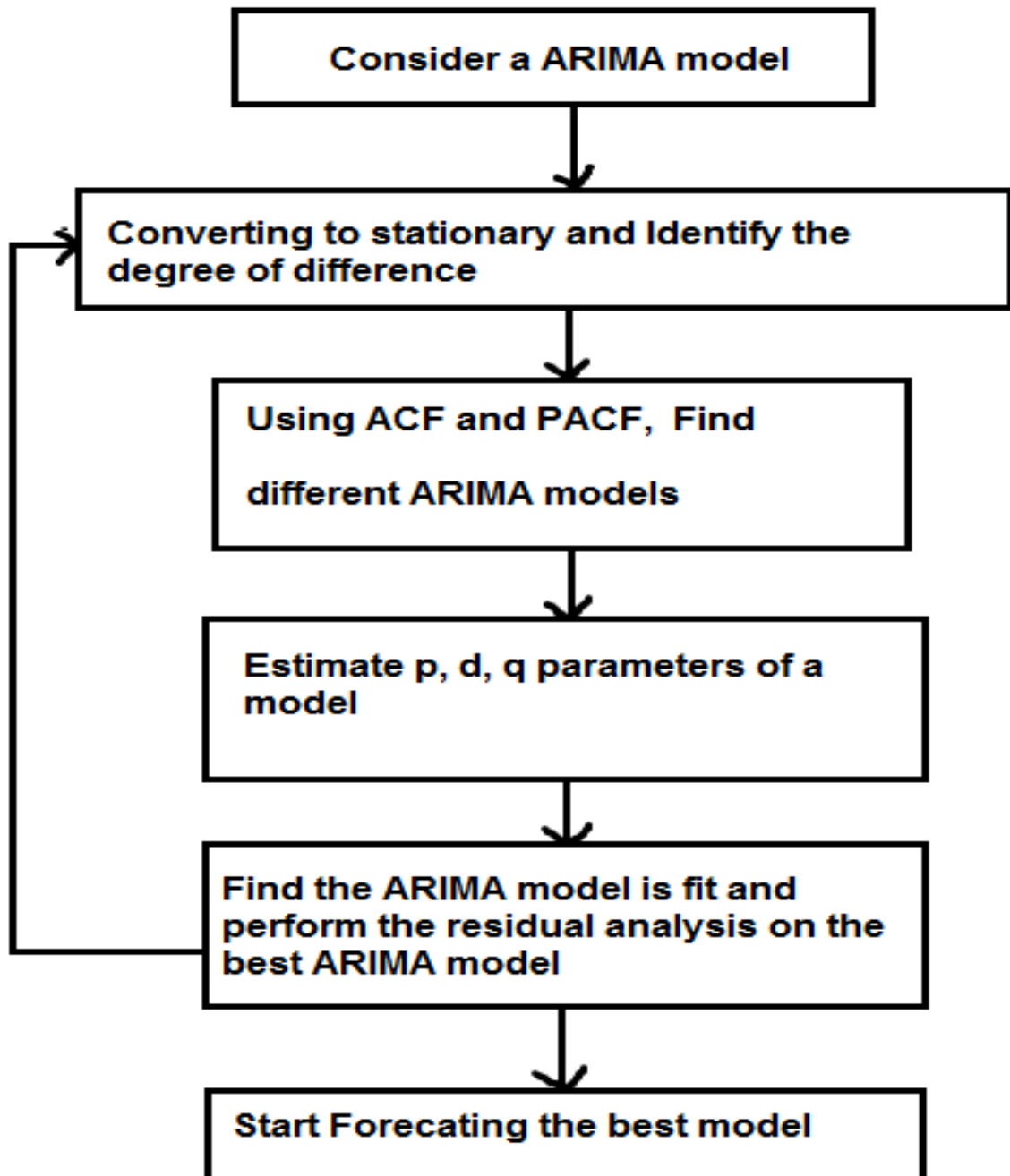
#### **4. Analysis Models**

##### **4.1 Data Flow Diagrams (DFD)**

##### **DFD LEVEL 0**



4.2 Flow chart of ARIMA Model



## **5. Change Management Process**

The change management process establishes an orderly and effective procedure for tracking the submission, coordination, review, evaluation, categorization and approval for release of all changes to the project's baselines.

Change request process flow:

- **Generate Application:** The submitter creates an application to the team.
- **Evaluate Application:** Project team reviews the application and provide an estimated level of effort to process, and to develop a proposed solution for the suggested changes.
- **Authorize:** Approval to move forward with incorporating the suggested change into the project/product.
- **Implement:** If approved, make the necessary adjustments to carry out requested change and communicate the submitter and other stakeholders.