**CHAPTER 1**

**INTRODUCTION**

* 1. **Data science** is a [multi-disciplinary](https://en.wikipedia.org/wiki/Multi-disciplinary) field that uses scientific methods, processes, algorithms and systems to extract [knowledge](https://en.wikipedia.org/wiki/Knowledge) and insights from structured and unstructured [data](https://en.wikipedia.org/wiki/Data). Data science is the same concept as [data mining](https://en.wikipedia.org/wiki/Data_mining) and [big data](https://en.wikipedia.org/wiki/Big_data). use the most powerful hardware, the most powerful programming systems, and the most efficient algorithms to solve problems. Data science has recently become a popular term among business executives. However, many critical academics and journalists see no distinction between data science and [statistics](https://en.wikipedia.org/wiki/Statistics), whereas others consider it largely a popular term for data mining and big data.
  2. **Python** is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed. Python is a great general-purpose programming language on its own, but with the help of a few popular it becomes a powerful environment for scientific computing.
  3. **Machine learning** is the [scientific study](https://en.wikipedia.org/wiki/Branches_of_science) of [algorithms](https://en.wikipedia.org/wiki/Algorithm) and [statistical models](https://en.wikipedia.org/wiki/Statistical_model) that  [computer systems](https://en.wikipedia.org/wiki/Computer_systems) use to perform a specific task without using explicit instructions, relying on patterns and [inference](https://en.wikipedia.org/wiki/Inference) instead. It is seen as a subset of [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence). Machine learning algorithms build a [mathematical model](https://en.wikipedia.org/wiki/Mathematical_model) based on sample data, known as [training data](https://en.wikipedia.org/wiki/Training_data), in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in a wide variety of applications, such as [email filtering](https://en.wikipedia.org/wiki/Email_filtering) and [computer vision](https://en.wikipedia.org/wiki/Computer_vision), where it is difficult or infeasible to develop a conventional algorithm for effectively performing the task.

**CHAPTER 2**

**COMPANY PROFILE**



Company Name : Inventeron Technologies and Business Solutions LLP

Company Category : IT Software/Embedded.

Contact Person : Mohammed Atha H.K.

Designation : CEO.

Email id : athai@inventeron.com

Website : http://www.inventerontechnologies.com

Year Stand Up : 2013.

Registered office : #14/1, Ground Floor, Magadi Road Toll gate,

Magadi main road, Bengaluru.

Phone : (080) 23146545.

**CHAPTER 3**

**ABOUT THE COMPANY**

Inventeron Technologies and Business Solutions LLP, or ITABS, is an Indian based engineering and electronics company headquartered in Bangalore, Karnataka, India. It is both product and service oriented software company having its products in wireless communication Technology and provides quality service to its valuable clients in its domain.

Inventeron core products are Embedded components (including IC’s, control boards, Controllers, Microprocessors, fuel systems, Water Level controllers, Security Systems, Biometric login systems, Wireless devices etc), industrial products (Network Tower Management systems, ) and Apps (E-commerce apps, website design and development).

**Purpose :** To be a leader in the software Industry by providing enhanced services, relationship and profitability.

**Vision :** To provide quality services that exceeds the expectations of our esteemed customers.

**Mission :** To build long term relationships with our customers and clients and provide exceptional customer services by pursuing business through innovation and advanced technology.

**Core values :**

• We believe in treating our customers with respect and faith.

• We grow through creativity, invention and innovation.

• We integrate honesty, integrity and business ethics into all aspects of our business functioning.

**Goals :**

• Regional expansion in the field of Electronics engineering and software field and develop a strong base of key clients.

• Increase the assets and investments of the company to support the development of services.

A company’s production department plays an integral role in the life cycle of a product. While the department usually is separable from sales, finance, HR,R&D and other divisions, the functions of these areas are related and often require collaboration. The production department is responsible for converting inputs into output through the stages of production process. The production manager is responsible for making sure that raw materials are provided and made into finished goods effectively.

Manager must make sure that work is carried out smoothly, and must supervise procedures for making work more efficient and more enjoyable, As a manager has number of responsibilities, but the primary one is to get product out the door. The goal is to deliver result to the customer, or market and do everything necessary to achieve this, To do this the manager need to make sure the development team is able to work as efficiently as possible and this means making sure they have clear goals, both short term and long term and that nothing prevents them from doing their work. From initial project scope to deploying out to customer sites, each step is the responsibility of a manager. The manager should delegate as much as he can but be ready to check that things are being done as required and be ready to jump in if it is not.

**Machine Learning Engineer Responsibilities:**

* Designing and deploying machine learning and deep learning systems.
* Running machine learning test and experiments.
* Implementing appropriate ML algorithms,
* Study and transform data science prototypes.
* Design machine learning systems.
* Research and implement appropriate ML algorithm and tool.
* Develop machine learning application according to requirements.
* Select appropriate datasets and data representation methods.
* Run machine learning test and experiments.
* Perform statistical analysis and fine-tuning using test results.
* Train and retain system when necessary.
* Extend existing ml libraries and framework.

**3.1 Key Strength**

Our team is the key strength of our Skills because of which we are experts in Software Development, mobile Apps development, Embedded Apps development, Web design and development, PCB designing, Online and Multimedia, Cloud computing, ERP,E-Commerce and Much more.

**Training -** The company also focuses on providing training to the individual on various platforms. They are as follows:

* **Machine Learning:** Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence. Company provides training on developing Machine Learning Projects from Scratch.
* **Android Application Development:** Android Application is a special software Application designed for mobile phones and tablets which runs on android operating system. Company provides training on the creation and deployment of android apps for interested individual. Web Application Development:
* **Web Application Development:** is the creation of application programs that reside on remote servers and are delivered to the user’s device over internet. A web application does not need to be downloaded and is instead accessed through a network. An end user can access a web application through web browser such as Google Chrome, Safari or Mozilla Firefox. A majority of web applications can be written in JavaScript, cascading style sheets and HTML5.Company provides training on developing web Projects from Scratch.
* **Internet of Things:** The IOT depends on a whole host of technologies – such as application programming interfaces (APIs) that connect devices to the Internet.Other key IOT technologies are Big Data management tools, predictive analytics, AI and machine learning, the cloud, and radio-frequency identification (RFID).
* **Embedded product design**: is a specialized area that requires trained talent, expensive tools and modern equipment. In addition, there’s also the need for various support services such as industrial design, mechanical design, thermal engineering along with regulatory certifications.

**CHAPTER 4**

**CLIENTS AND PRESENT PROJECTS**

**4.1 Clients**

* Robert Bosch.
* Next Power Systems Pvt. Ltd.
* NSK Electronics.Bio-needs India Pvt. Ltd.FabelluS.

**4.2 Present Projects**

* The company presently working for Robert Bosch in Industrial Automation project.
* We are working for Febellus in developing E-commerce application.
* The company also working on Embedded products and Android apps and web development.

**4.3 Company Domains**

* **IT Consulting:** We have tied up with many companies and we train the candidates with good skills by giving them real time exposure and consult them to many top MNC’s.
* **Outsourcing** :We do outsource many projects to our associated companies and have build a very good network among the companies.
* **HR Management:** We do have a separate HR department for training and Recruiting Purpose. We do offer HR management skills for the needed.
* **Food and Beverages:** Since Food and Beverages is the fastest and evergreen Industry always moving together with the growing technology we do give technical support in many aspects to the industry.
* **Retails:** We do develop software applications needed for smooth maintenance of the accounts and Transactions in the Retails and Wholesale Industry.
* **Health Care:** Healthcare is the ever needed and very important sector always seeking contribution of Technology for its efficient monitoring. We give smart solutions in many ways like software application development, Product development, service and Maintenance.
* **Education:** WE have extended our Arms into the field of Education and brought a revolutionary change in the field by launching various software projects which are suitable in the educational institutes in various aspects such as student data base.

**4.4 Products**

* Smart Surveillance system.
* Safety and Security Systems.
* Biometrics.
* Industrial Automation.
* Smart Traffic Systems.
* Vehicle Tracking Systems.
* Tower Management System.
* Education Management Systems.
* Hotel Management System.
* Personal safety Equipment’s.
* Wireless Communication Devices.
* Software Applications.
* Website Design and Development.

**4.5 Roles and Responsibilities of Individuals**

**4.5.1 Department manager**

A design department manager bridges relationship between creative staff and management. The manager oversees the organization of strategic design projects. This individual is a comfortable Managing people and process as at monitoring the profitability of the work. Once project is awarded, design department manager works with the design director and account director to determine how to best allocate the budget and project hours so that project is profitable. Department managers are responsible for all the functions of department and oversees the staff and implement strategies to increase productivity. Managers also conduct training and seminars to motivate and build skills of their team members.

**4.5.2** **Functions**

* Mentor and train employees.
* Provide a comfortable working Environment.
* Implement business strategies to increase sales.
* Maintain and improve company standards.
* Keep staff motivated.
* Exceed customer satisfaction objectives.
* **IT department:** Goal is to design, maintain, and support an organization’s information technology infrastructure, thus allowing the organization to leverage both information and technology in an efficient, productive and secure manner.
* **Team lead:** The term team lead is responsible for creating an environment oriented to trust, open communication, creative thinking and cohesive team effort. He motivates and inspires team members by setting a good example, thus providing the team with vision if the project objectives A team leader is someone who provides guidance, instruction, direction and leadership to a group of other individual for the purpose of achieving key result or group of aligned result.
* **Project Manager:** The project manger holds the responsibility of the team working on the projects under him. His team may consists of programmers, business analyst, team leads, project leads, technical architects and other employees. The team may report to the project manager regarding progress of the projects undertaken. Project managers ensure project is completed on the time and within budget, that project’s objectives are met and that everyone else is doing their job properly. Projects are usually separate to usual day-today business activities and require the group of people to work together to achieve a set of specific objectives.
* **Software Developer:** Software developer plays a key role in the design, installation, maintenance, testing of software systems. The programs that they create are likely to be more efficient and provide a better service and responsible for reviewing current system, presenting ideas for system improvements, producing detailed specifications, coding the solution for defined problems, preparing training manual for users.
* **Human Resource Department:** Consists of HR Manager and HR Executives. They are responsible for recruiting, screening, interviewing and placing work force in an organization and handle employee relations, payroll, benefits, and training. HR managers plan, direct and coordinate the administrative functions of an organization, shown in figure 4.6.
* **System Designer:** Create detailed design documentation for the development and integration of computer systems to meet the needs of business. They work with analysts on the feasibility of the conceptual design by taking technical specifications prepared by the analyst and designing the system components to meet the requirements. They are responsible for designing and monitoring and performance measurement processes.
* **Business Analyst:** Takes the responsibility of taking to the business users of computers system to understand their needs. The business analyst produces the list of requirements which clearly state the business needs and align with business processes.
* **Tester**: A software tester is involved in the quality assurance stage of software development and deployment. They conduct automated and manual tests to ensure the software created by developers is suitable for this purpose. The role is integral to the

creation of software systems and is employed to find bugs and issues before it gets deployed in real environment.

**4.6 Structure Organization of Company**

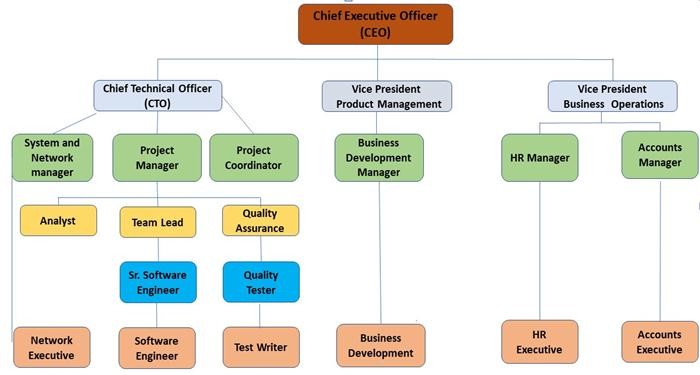
****

Figure 4.1: Structure organization of Company

* The software development team consists of
  + Requirement Analysts.
  + Project Managers.
  + Developers.
* Developer Department team Consists of
* Full-stack developer.
* Front-end developer.
* Back-end developer.

**CHAPTER 5**

**TASKS PERFORMED**

**5.1 Introduction**

The use of algorithms to make trading decisions has become a prevalent practice in major stock exchanges of the world. Algorithmic trading, sometimes called high-frequency trading, is the use of automated systems to identify true signals among massive amounts of data that capture the underlying stock market dynamics. Machine Learning has therefore been central to the process of algorithmic trading because it provides powerful tools to extract patterns from the seemingly chaotic market trends.

This project, in particular, learns models from Bloomberg stock data to predict stock price changes and aims to make profit over time. In this project, we examine two separate algorithms and methodologies utilized to investigate Stock Market trends and then iteratively improve the model to achieve higher profitability as well as accuracy via the predictions. Prediction of stock price level movement is thought to be a difficult task of monetary statistic prediction. Associate degree correct prediction of stock worth movement might yield profits for investors. As a result of the quality of exchange information, development of Economical models for predicting is incredibly troublesome. Statistical strategies and neural networks are usually used for statistic prediction. Since stock markets are complicated, nonlinear, dynamic and chaotic. The vital part of machine learning is the dataset used. The dataset should be as concrete as possible because a little change in the data can perpetuate massive changes in the outcome. In this project, supervised machine learning is employed on a dataset obtained from Apple.enc. This dataset comprises of following five variables: open, close, low, and high. Open, close, low and high are different bid prices for the stock at separate times with nearly direct names. The volume is the number of shares that passed from one owner to another during the time period. The model is then tested on the test data.

A machine learning project may not be linear, but it has a number of well known steps:

* Define Problem.
* Prepare Data.
* Evaluate Algorithms.
* Improve Results.
* Present Results.

**5.2 Objectives**

Since the tick-by-tick entries retrieved from Bloomberg happen in non-deterministic timestamps, we attempted to standardize the stock data by discretizing the continuous time domain, from 9:00 am to 5:00 pm when the market closes. Specifically, the time domain was separated into 1-minute buckets and we discarded all granularities within each bucket and treated the buckets as the basic units in our learning algorithms. In the past decades, there is an increasing interest in predicting markets among economists, policymakers, academics and market makers. The objective of the proposed work is to study and improve the supervised learning algorithms to predict the stock price. Problem statement Investors invest money on the product and the outcome may not be as expected. This leads to major drop in the company economic and whole shares will drop down.

**Overview**

1. **Installing the Python and SciPy platform.**

There are 5 SciPy key libraries that you will need to install.

* Scipy Numpy.
* Matplotlib.
* Pandas.
* sklearn .

1. **Loading the dataset.**

We are using pandas to load the data. We will also use pandas next to explore the data

both with descriptive statistics and data visualization.

1. **Summarizing the dataset.**

In this step, we look at the data a few different ways.

* Dimensions of the dataset.
* Peek at the data itself.
* Statistical summary of all attributes.
* Breakdown of the data by the class variable.

1. **Visualizing the dataset.**

We consider different types of plots.

* Uni variety plots to better understand each attribute.
* Multivariate plots to better understand the relationships between attributes.

1. **Evaluating algorithms**.

To create model of the data and estimate their accuracy on unseen data.

* Separate out a validation dataset.
* Build models to predict species from flower measurements to select best Algorithm.

1. **Making some predictions.**

We can fit the model on the entire training dataset and make predictions on the validation dataset and evaluate the predictions by comparing them to the expected results in the validation set, then calculate classification accuracy.

**5.3 Modules**

Modules are described below one by one.

**1)** **Data collection :** The dataset is collected from investment.com

**2**) **Extraction Dataset :** Extraction of Data set contains Close, High, Low, Volume.

**Data is stored in .csv format**

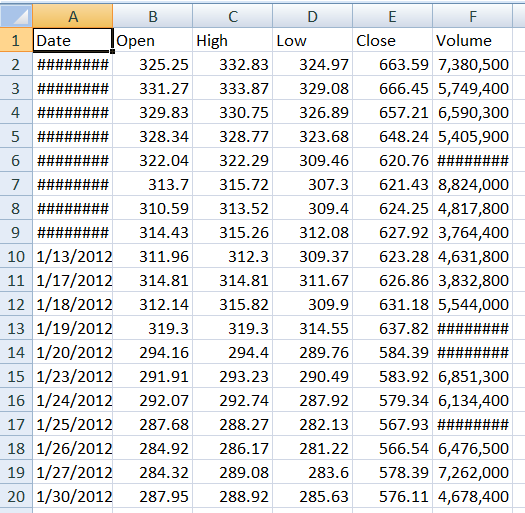


Table 5.1 : Dataset of Apple.inv

**3) Data Preparation :** The collected data will be divide into training set and testing set with the ratio of 80:20.Hence final training set is 7585 and the test set is 1897.

**Training Dataset:**

* The first imports Apple,inc dataset which is already predefined in sklearn module. Iris dataset is basically a table which contains information about stock market price.
* We import LSTM algorithm and train test split class from sklearn and numpy module.
* we encapsulate load\_ Apple,inc() method in Apple,inc\_dataset variable.
* we divide the dataset into training data and test data using train\_test\_spilt method.
* The X denotes feature values and y denotes target values.
* This method divides dataset into training and test data randomly 80:20.
* We fit our training data to LSTM algorithm so that it get trained using this data.

**Testing dataset:**

* We have dimensions of stock in numpy array called x\_new and we want to predict next day price of stock of Iris. We do this by using predict method which takes this array as input and splits out predicted target value as output.
* The predicted target value comes out in the graph blue represents next day• stock market price. Finally we find entropy to reduce the error.

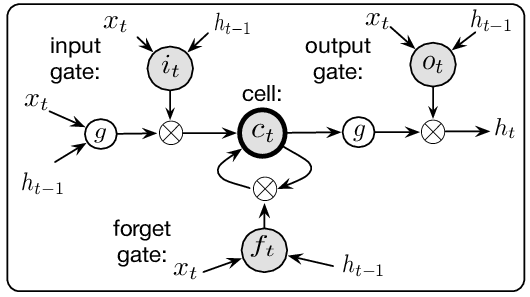


Figure 5.1 : LSTM Diagram

**4) Modeling:** The model used is LSTM(Long Short Term Memory) is an Artificial Recurrent Neural Network. LSTM can not only process the single data, it can process data sequentially and keep in the hidden state.

* Long Short-Term Memory(LSTM): The models are extremely powerful time-series models. An LSTM module (or cell) has 5 essential components which allows it to model long-term data.
* Cell state (ct) - This represents the internal memory of the cell which stores both short term memory and long-term memories Hidden state (ht) - This is output state information calculated w.r.t. current input, previous hidden state and current cell input which eventually use to predict the future stock market prices.
* Input gate (it) - Decides how much information from current input flows to the cell state.
* Forget gate (ft) - Decides how much information from the current input and the previous cell state flows into the current cell state.
* Output gate (ot) - Decides how much information from the current cell state flows into the hidden state, so that if needed LSTM can only pick the long-term memories.

**5.4 Flow Chart**

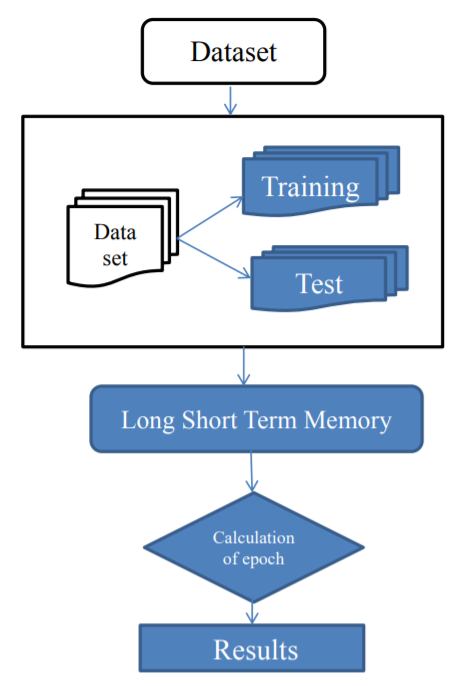


Figure 5.2 : Flow chart

DESCRIPTION:

* DATASET: Dataset is a very basic module and the initial step towards the project. It generally deals with the collection of the right dataset. The dataset that is to be used in the market prediction has to be used to be filtered based on various aspects. Dataset also complements to enhance the dataset by adding more data that are external. Our data mainly consists of the previous year stock prices. Initially, we will be analyzing the dataset and according to the accuracy, we will be using the model with the data to analyze the predictions accurately.
* TRAINING AND TESTING DATA: Training the machine is similar to feeding the data to the algorithm to touch up the test data. The training sets are used to tune and fit the models. The test sets are untouched, as a model should not be judged based on unseen data. The training of the model includes cross-validation where we get a well-grounded approximate performance of the model using the training data. Finally, we will calculate a cross-validated score, for individual sets of hyper parameters. Then, we select the best hyper parameters. The idea behind the training of the model is that we some initial values with the dataset and then optimize the parameters which we want to in the model. This is kept on repetition until we get the optimal values. Thus, we take the predictions from the trained model on the inputs from the test dataset. Hence, it is divided in the ratio of 70:30 where 70% is for the training set and the rest 30% for a testing set of the data.
* LONG SHORT TERM MEMORY: Learning algorithm that utilizes a kind of recurrent neural network (RNN) called Long Short Term Memory (LSTM), where the weights are adjusted for individual data points using stochastic gradient descent. This will provide more accurate results when compared to existing stock price prediction algorithms. The network is trained and evaluated for accuracy with various sizes of data, and the results are tabulated. A comparison with respect to accuracy is then performed against an Artificial Neural Network.

# CHAPTER 6

**REQUIREMENT ANALYSIS AND FEASIBILITY STUDY**

**Feasibility Study**

Simply put, stock market cannot be accurately predicted. The future, like any complex problem, has far too many variables to be predicted. The stock market is a place where buyers and sellers converge. When there are more buyers than sellers, the price increases. When there are more sellers than buyers, the price decreases. So, there is a factor which causes people to buy and sell. It has more to do with emotion than logic. Because emotion is unpredictable, stock market movements will be unpredictable. It’s futile to try to predict where markets are going. They are designed to be unpredictable.

The proposed system will not always produce accurate results since it does not account for the human behaviours. Factors like change in company’s leadership, internal matters, strikes, protests, natural disasters, change in the authority cannot be taken into account for relating it to the change in Stock market by the machine.The objective of the system is to give a approximate idea of where the stock market might be headed. It does not give a long term forecasting of a stock value. There are way too many reasons to acknowledge for the long term output of a current stock. Many things and parameters may affect it on the way due to which long term forecasting is just not feasible.

**Requirement Analysis**

After the extensive analysis of the problems in the system, we are familiarized with the requirement that the current system needs. The requirement that the system needs is categorized into the functional and non-functional requirements. These requirements are listed below:

**Functional Requirements**

Functional requirement are the functions or features that must be included in any system to satisfy the business needs and be acceptable to the users. Based on this, the functional requirements that the system must require are as follows:

• The system should be able to generate an approximate share price.

• The system should collect accurate data from the NEPSE website in consistent manner.

**Non-Functional Requirements**

Non-functional requirement is a description of features, characteristics and attribute of the system as well as any constraints that may limit the boundaries of the proposed system. The non-functional requirements are essentially based on the performance, information, economy, control and security efficiency and services. Based on these the non-functional requirements are as follows:

• The system should provide better accuracy.

• The system should have simple interface for users to use.

• To perform efficiently in short amount of time.

**CHAPTER 7**

**SYSTEM ANALYSIS**

**Problem Statement:** A stock exchange market depicts savings and investments that are advantageous to increase the effectiveness of national economic. The future stock returns have some predictive relationships with the publicly available information of present and historical stock market indices. The investors decide the better time to sell/buy/hold a share in stock market based on the former relationship. Every investor is interested in predicting the future stock prices, whether the investor may be a long-term investor or a day-trader. This possesses a major challenge to design and develop an effective and efficient predictive model that assists the investors to take appropriate decisions.

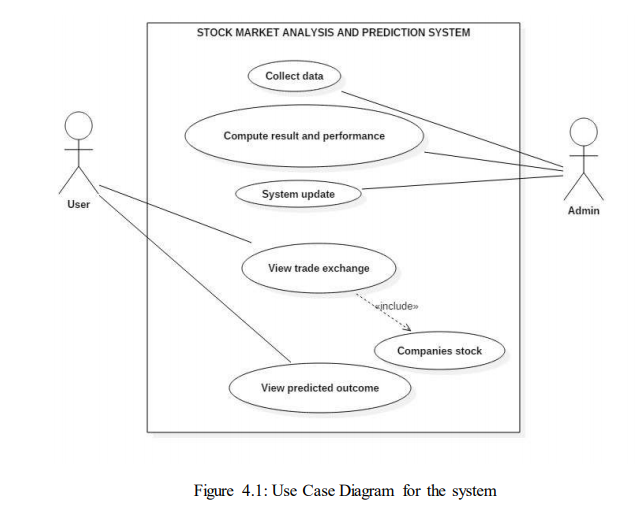
**Existing Systems :**One of the significant financial subject that has engrossed the researcher’s attention for many years is forecasting the stock returns. Investors in the stock market have been attempting to discover an answer to estimate the stock trends in order to decide the better timing to buy or sell or hold a share. Forecasting the stock trends have been done both on qualitative analysis and quantitative analysis. There are many statistical models available for forecasting stock trends and choosing an appropriate model for a particular forecasting application depends on the format of the data.

**Proposed Study :**In this work we propose a prediction model for the time series stock market data. This model will automate the process of change of stock price indices based on technical analysis and provides assistance for financial specialists to choose the better timing for purchasing and selling stocks. Data mining techniques are used to develop the prediction model and R programming language is used for visualization of results.

**CHAPTER 8**

**SYSTEM DESIGN AND ARCHITECTURE**

**Use Case Diagram**

****Figure 8.1: Use case diagram.

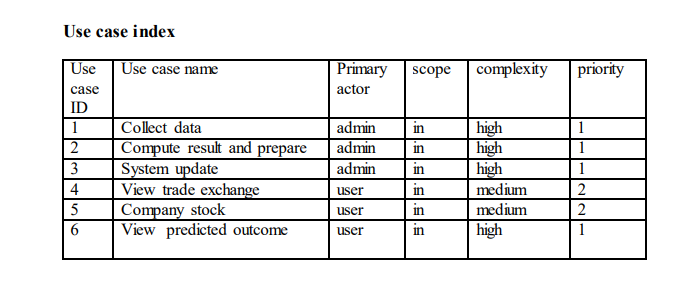
****

Figure 8.2: Use case index.

**Use case description:**

**Use case ID:1**

Use case name: Collect data

Description: Every required data will be available in Nepal stock exchange. Admin will be able to collect the data for system.

**Use case ID:2**

Use case name: Compute result and performance

Description: Prediction result will be handled and generated by admin. The system will be built, through which the result of prediction and system performance will be analyzed.

**Use case ID: 3**

Use case name: System update

Description: With the change of market and technology regular update of system is required. Beside there the predict result of stock exchange and their actual price will be updated by admin in regular basis.

**Use case ID: 4**

Use case name: View traded exchange

Description: Company trading which is held at NEPSE can be viewed by user.

**Use Case ID: 5**

Use Case Name: Company Stock

Description: It is extended feature of view traded exchange. This includes the stock value of particular company.

**Use Case ID: 6**

Use Case Name: View predicted outcome

Description: This use case is must important in whole project. The key feature of this project is to predict the stock value of hydropower companies. Thus, this will be available in user interface and viewer can observe them.

**CHAPTER 9**

**SYSTEM FLOW DIAGRAM**

Data Source

Graphical User Interface

Stock Database(SQL Server)

Figure 9.1: System flow diagram

DESCRIPTION:

* DATA SOURCE: A data source, in the context of computer science and computer applications. In a database management system, the primary data source is the database, which can be located in a disk or a remote server.
* STOCK DATABASE(SQL SERVER): Certain principles guide the database design process. The first principle is that duplicate information (also called redundant data) is bad, because it wastes space and increases the likelihood of errors and inconsistencies. The second principle is that the correctness and completeness of information is important. If your database contains incorrect information, any reports that pull information from the database will also contain incorrect information. As a result, any decisions you make that are based on those reports will then be misinformed.

A good database design is, therefore, one that:

* + Divides Company stock price information into subject-based tables to reduce redundant data
  + Provides Access with the information it requires to join the information in the tables together as needed.
  + Helps support and ensure the accuracy and integrity of your information.
  + Accommodates your data processing and reporting needs.
* SOFTWARE FRAMEWORK: A software framework, is a platform for developing software [applications](https://techterms.com/definition/application). It provides a foundation on which software developers can build programs for a specific [platform](https://techterms.com/definition/platform). For example, a framework may include predefined [classes](https://techterms.com/definition/class) and [functions](https://techterms.com/definition/function) that can be used to process [input](https://techterms.com/definition/input), manage hardware devices, and interact with [system software](https://techterms.com/definition/systemsoftware). This streamlines the development process since programmers don't need to reinvent the wheel each time they develop a new application.
* GRAPHICAL USER INTERFACE: The graphical user interface is a form of [user interface](https://en.wikipedia.org/wiki/User_interface)  allows [users](https://en.wikipedia.org/wiki/User_(computing)) to [interact with electronic devices](https://en.wikipedia.org/wiki/Human%E2%80%93computer_interaction) through graphical [icons](https://en.wikipedia.org/wiki/Computer_icon) and audio indicator such as primary notation, instead of [text-based user interfaces](https://en.wikipedia.org/wiki/Text-based_user_interface), typed command labels or text navigation.

# CHAPTER 10

# RESOURCE UTILIZATION

* Effective management of resources is an essential task for companies that are managing different projects .Its important for them to efficiently organize and allocate personal as well as equipment for different projects, same time avoiding idle resources.
* Resource utilization refers to the process of making the most of the resources available to you in order to achieve the objective that you want to. Your individual project and the utilization of your resources for the same could also be called resource utilization.
* During this Internship period, learnt the lesson of resource utilization and its direct benefit on the growth of the company. At the most basic level, utilization is the metric professional services companies use to measure the profitability of the people.

# 10.1 Software Tools

# Anaconda Navigator tool

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda® distribution that allows you to launch applications and easily manage conda packages, environments, and channels without using command-line commands. Navigator can search for packages on Anaconda Cloud or in a local Anaconda Repository. It is available for Windows, macOS, and Linux.

The following applications are available by default in Navigator:

* [JupyterLab](https://jupyterlab.readthedocs.io/en/stable/)
* [Jupyter Notebook](https://jupyter.readthedocs.io/en/latest/)
* [Spyder](https://www.spyder-ide.org/)
* [VSCode](https://code.visualstudio.com/docs)
* [Glueviz](http://glueviz.org/en/stable/)
* [Orange 3 App](http://orange.biolab.si/docs/)
* [RStudio](http://docs.rstudio.com/).

# Jupyter Notebook

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter.

Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

# Spyder

Spyder is a powerful scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.

Spyder offers built-in integration with many popular scientific packages, including NumPy, SciPy, Pandas, IPython, QtConsole, Matplotlib, SymPy, and more.Spyder’s abilities can be extended even further via its plugin system and API. Spyder can also be used as a PyQt5 extension library, allowing you to build upon its functionality and embed its components, such as the interactive console, in your own software.

# 10.2 LANGUAGES USED

**Python**

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable and is Interpreted , Interactive, Object-Oriented Beginner's Language.Python features a dynamic type system and automatic memory management. It supported multiple programming paradigms, including object-oriented, imperative and has large and comprehensive standard library.

Python become popular because of 5 libraries are TensorFlow, NumPy, SciPy, Pandas and Matplotlib. Python is the most widely used programming language today. When it comes to solving data science tasks and challenges, Python never ceases to surprise its users. Most data scientists are already leveraging the power of Python programming every day. Python has been built with extraordinary Python libraries that are used by programmers every day in solving problems.

Python features include:

* + - * **Easy-to-learn** − Python has few keywords, simple structure, and a clearly defined syntax. This allows the student to pick up the language quickly.
      * **Easy-to-read** − Python code is more clearly defined and visible to the eyes.
      * **Easy-to-maintain** − Python's source code is fairly easy-to-maintain.
      * **A broad standard library** − Python's bulk of the library is very portable and cross- platform compatible on UNIX, Windows, and Macintosh.
      * **Interactive Mode** − Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.
      * **Portable** − Python can run on a wide variety of hardware platforms and has the same interface on all platforms.
      * **Extendable** − You can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.
      * **Databases** − Python provides interfaces to all major commercial databases.
      * **GUI Programming** − Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix.
      * **Scalable** − Python provides a better structure and support for large programs than shell scripting.

# 10.3 LIBRARIES

### TensorFlow

TensorFlow is a library for high-performance numerical computations with around 35,000 comments and a vibrant community of about 1,500 contributors. It’s used across various scientific fields. TensorFlow is a framework for defining and running computations that involve tensors, which are partially defined computational objects that eventually produce a value.

### NumPy

NumPy (Numerical Python) is the fundamental package for numerical computation in Python; it contains a powerful N-dimensional array object. It has around 18,000 comments on GitHub and an active community of 700 contributors. It’s a general-purpose array-processing package that provides high-performance multidimensional objects called arrays and tools for working with them.

### SciPy

SciPy (Scientific Python) is another free and open-source Python library extensively used in data science for high-level computations. SciPy has around 19,000 comments on GitHub and an active community of about 600 contributors.

### Pandas

Pandas (Python data analysis) is a must in the data science life cycle. It is the most popular and widely used Python library for data science, along with NumPy in matplotlib. With around 17,00 comments on GitHub and an active community of 1,200 contributors, it is heavily used for data analysis and cleaning. Pandas provide fast, flexible data structures, such as data frame CDs, which are designed to work with structured data very quickly and intuitively.

### Matplotlib

Matplotlib has powerful yet beautiful visualizations. It’s a plotting library for Python with around 26,000 comments on GitHub and a very vibrant community of about 700 contributors. Because of the graphs and plots that it produces, it’s extensively used for data visualization. It also provides an object-oriented API, which can be used to embed those plots into applications.

**10.4 PLATFORM**

**Machine Learning**

Machine learning is a branch of science that deals with programming the system in such a way that they automatically learn and improve with Experience. Here, learning process includes understanding input data and making wise decision based on supplied data.It is very difficult to gather to all the decision based on all possible inputs. To track the problem, algorithms are developed and these algorithms build knowledge from past experience with the principles of statistics, probability theory, logic, reinforcement learning and control theory.

Machine learning is a vast area and it is quite beyond the scope to cover all the features. There are several ways to implement machine learning technique and commonly used one are Supervised and Unsupervised learning.during the period of Internship, I have learnt Machine Learning Concept related to my project. I mainly focused on Decision Tree Algorithm that predicts type of Iris Species in my Project. I have learnt how it basically built, work, applications, benefits and disadvantages.

# CHAPTER 11

# ADVANTAGES AND DISADVANTAGES

# Advantages

* Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange .
* The successful prediction of a stock's future price could yield significant profit. The efficient-market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable.
* Allowing you to offer employees extra incentives by granting share options - this can encourage and motivate your employees to work towards long-term goals.
* Placing a value on your business.
* Increasing your public profile, and providing reassurance to your customers and suppliers.
* Allowing you to do business - eg acquisitions - by using quoted shares as currency.
* Creating a market for the company's shares.

# Disadvantages

* **Market fluctuations** - your business may become vulnerable to market fluctuations beyond your control - including market sentiment, economic conditions or developments in your sector.
* **Cost** - the costs of flotation can be substantial and there are also ongoing costs of being a public company, such as higher professional fees.
* **Responsibilities to shareholders** - in return for their capital, you will have to consider shareholders' interests when running the company - which may differ from your own objectives.
* **The need for transparency** - public companies must comply with a wide range of additional regulatory requirements and meet accepted standards of corporate governance including transparency, and needing to make announcements about new developments.
* **Demands on the management team** - managers could be distracted from running the business during the flotation process and through needing to deal with investors.

**CHAPTER 12**

**RESULTS**

Collection of the dataset from investment.com the graph is plot to the original values which is obtained from the website this green color represents the original dataset.

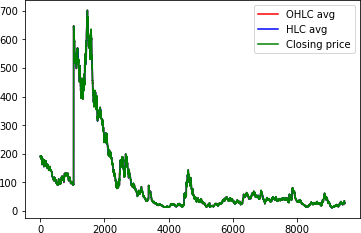


Figure 12.1: Original Dataset Of Google Stock Market.

(Courtesy: Google)

The original dataset the dataset is trained and tested with the LSTM algorithm based on the algorithm they are plotting the graph the red color indicated training dataset and the blur color represents the predicted dataset. The train RMSE value is 7.48 and Test RMSE value is 1.30. After calculation of Epoch to reduce the error for the given dataset the 10times epoch are calculated.

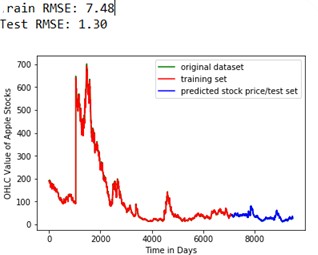


Figure 12.2: After Predicted stock market

**Epoch**

An **epoch** is a term used in [machine learning](https://radiopaedia.org/articles/machine-learning-overview?lang=us) and indicates the number of passes through the entire [training dataset](https://radiopaedia.org/articles/missing?article%5Btitle%5D=training-dataset&lang=us) the machine learning algorithm has completed. If the [batch size](https://radiopaedia.org/articles/batch-size-machine-learning?lang=us) is the whole training dataset (batch mode) then batch size and epoch are equivalent.

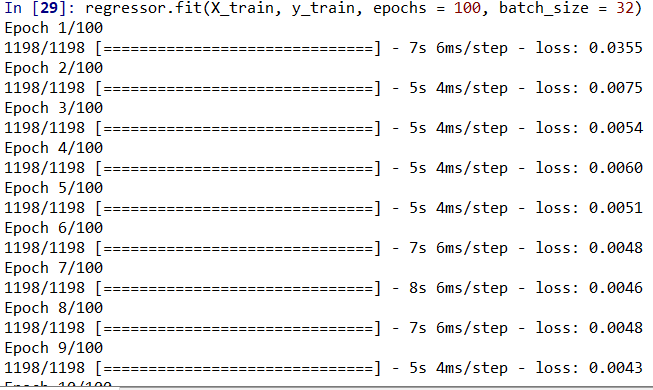


Figure 12.3: Calculation of epoch

**Training Data**

The observations in the training set form the experience that the algorithm uses to learn. In supervised learning problems, each observation consists of an observed output variable and one or more observed input variables. Training set is the data set on which your model is built. Training set is usually manually written and your model follows exactly the same rules and definitions given in the training set.

Some training sets may contain only a few hundred observations; others may include millions. Inexpensive storage, increased network connectivity, the ubiquity of sensor-packed smartphones, and shifting attitudes towards privacy have contributed to the contemporary state of big data, or training sets with millions or billions of examples.

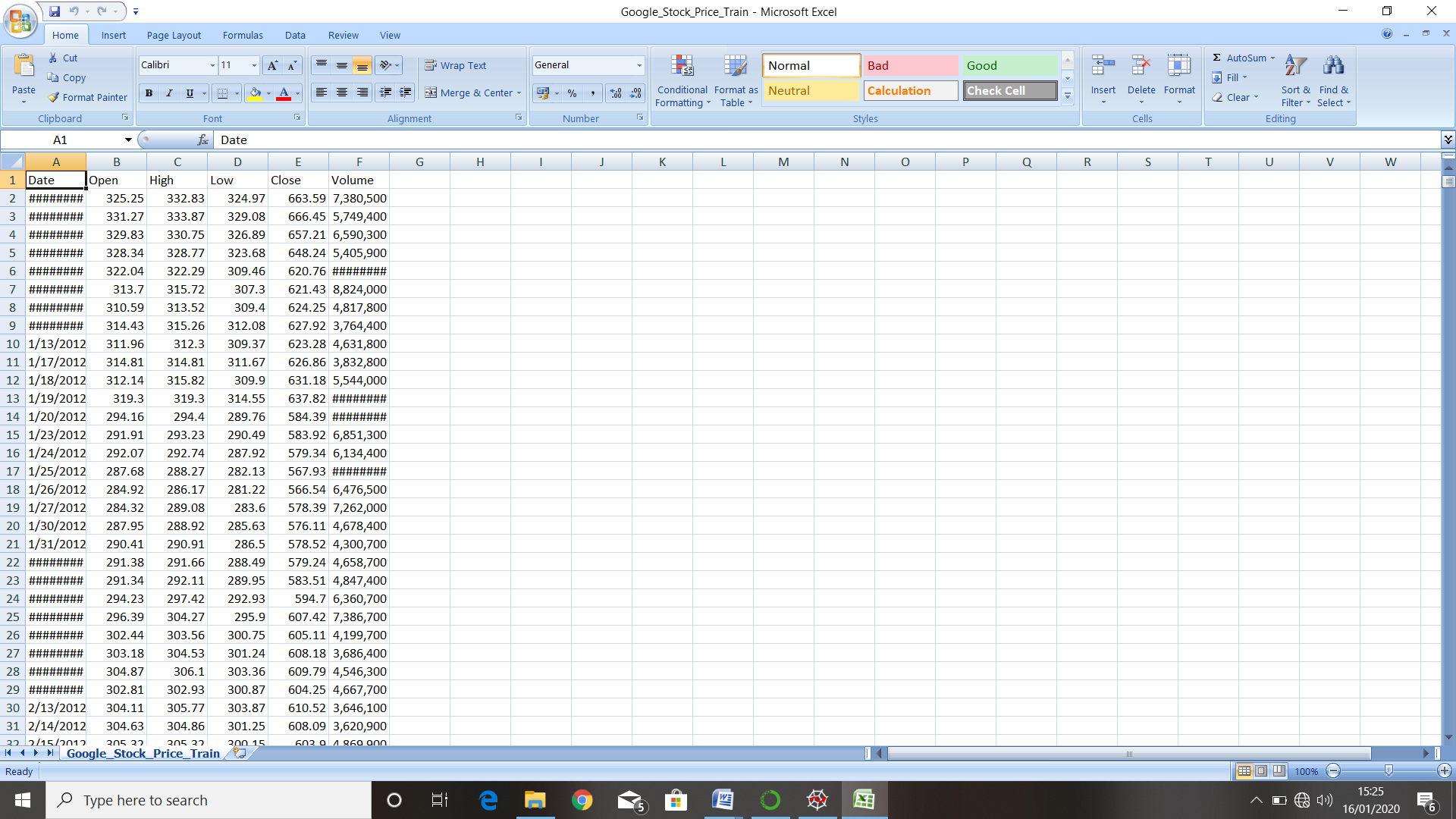


Table 12.1: Google Stock Price Training Data

**Testing Data**

The test set is a set of observations used to evaluate the performance of the model using some performance metric. It is important that no observations from the training set are included in the test set. If the test set does contain examples from the training set, it will be difficult to assess whether the algorithm has learned to generalize from the training set or has simply memorized it.

A program that generalizes well will be able to effectively perform a task with new data. In contrast, a program that memorizes the training data by learning an overly complex model could predict the values of the response variable for the training set accurately, but will fail to predict the value of the response variable for new examples. Memorizing the training set is called over-fitting. A program that memorizes its observations may not perform its task well, as it could memorize relations and structures that are noise or coincidence. Balancing memorization and generalization, or over-fitting and under-fitting, is a problem common to many machine learning algorithms. Regularization may be applied to many models to reduce over-fitting.

In addition to the training and test data, a third set of observations, called a validation or hold-out set, is sometimes required.

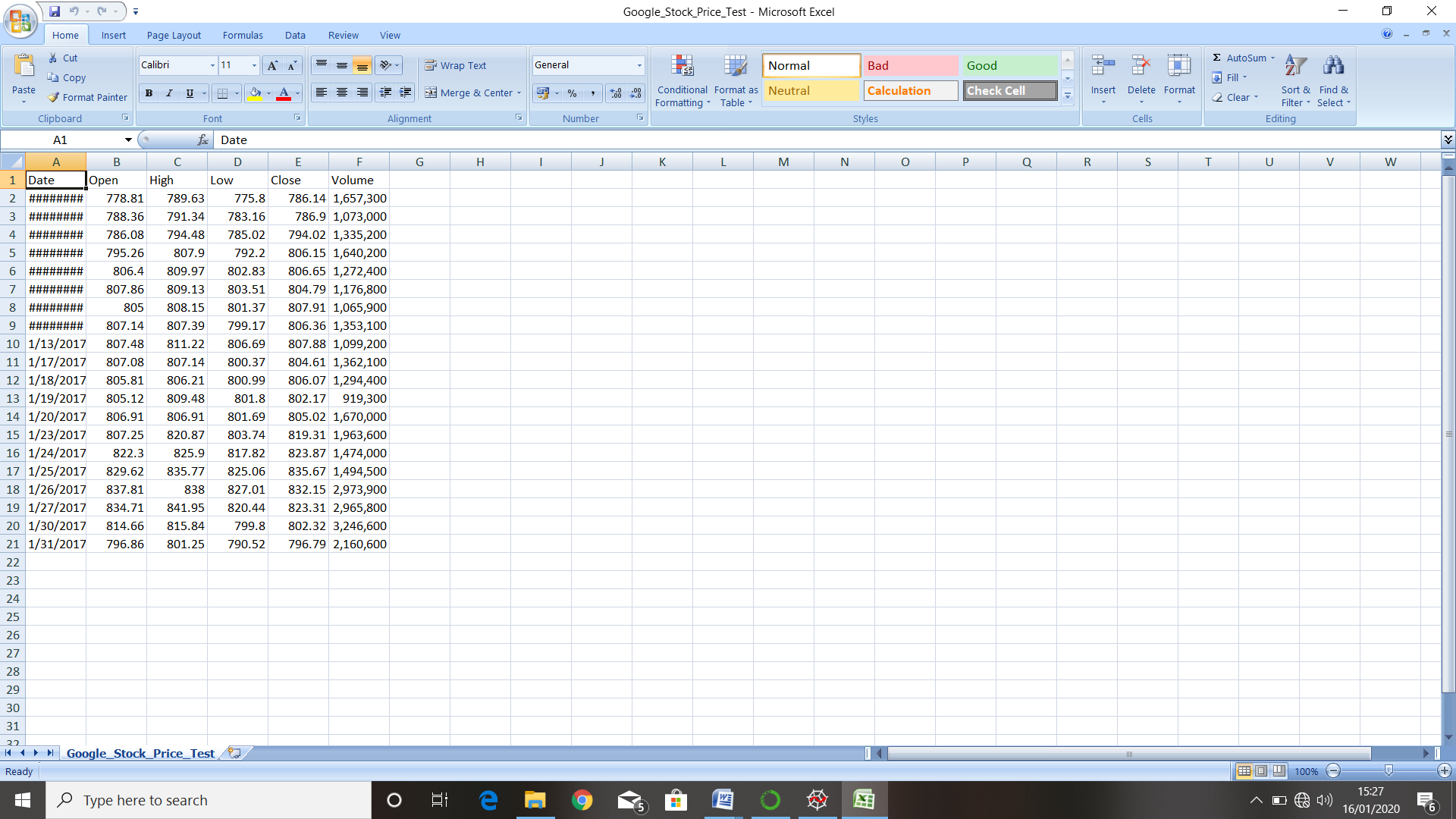


Table 12.2 : Google Stock Price Testing Data

**Output**

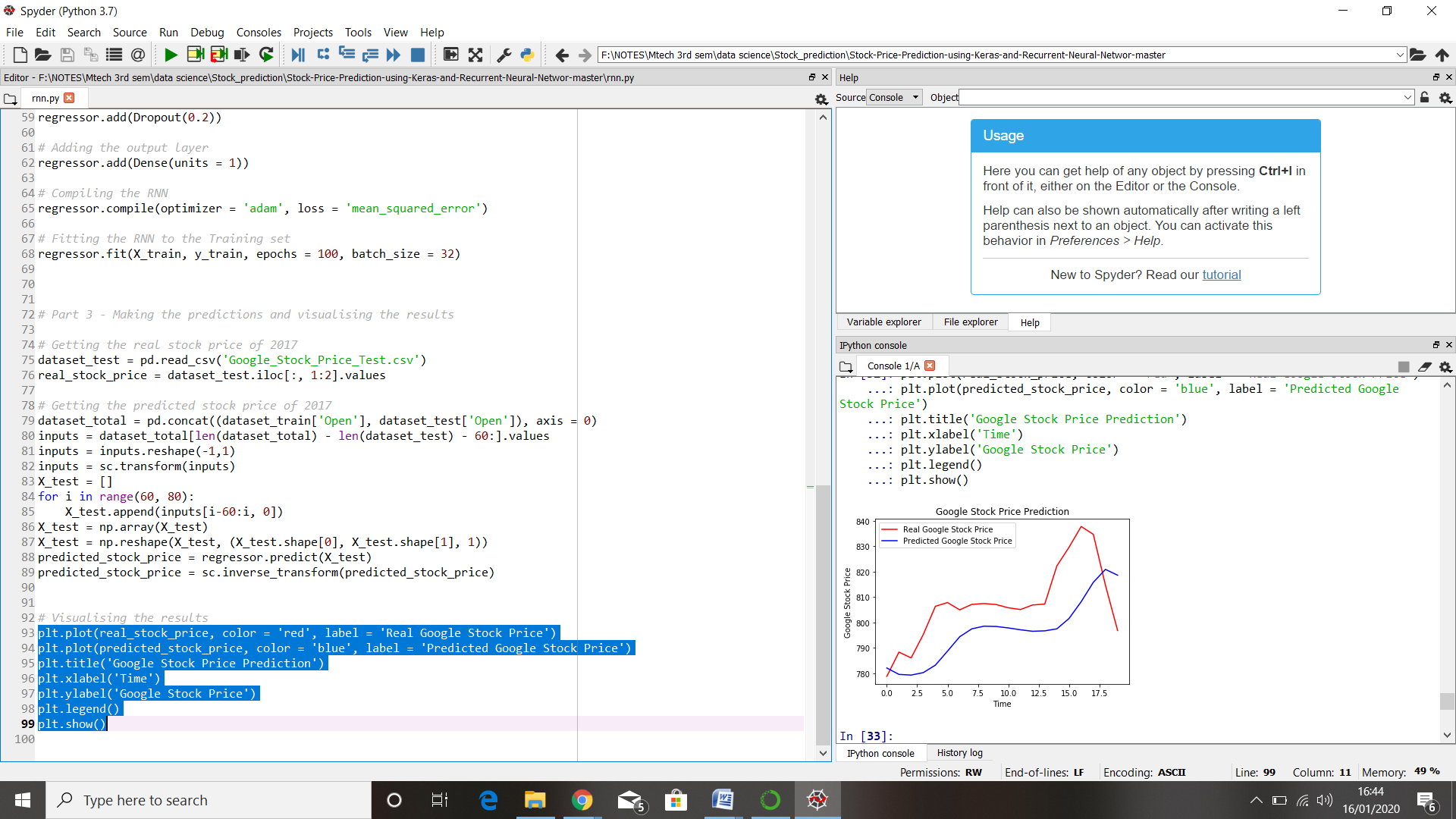


Figure 12.4: Google Stock Price Prediction

**Notation:**

* X-axis represent the time series of the day.
* Y-axis represent the Google Stock Price.
* Red line represent Real Google Price.
* Blue line represent Prediction Google Price.

**CONCLUSION**

Studied different methodologies for Stock Market Prediction which will help the investor for making the correct decision for buy or sell the stocks. Each method is having some limitation and some disadvantage. The limitations can be overcome by selecting suitable prediction techniques for specific domains. In future one can combine the two method and get proper result and output. Made an attempt to evaluate different methods of forecasting the stock market trends by which any invester can find the best method by which they can predict the stock market much more accurately than previously done methods. Based on the technical analysis using historical time series stock market data and data mining techniques. Based on the obtained graph in the final output, going to predict the final results.demonstrated the potential model to predict the stock price indices on short-term basis.

This could guide the investors in the stock market to make profitable investment decisions whether to buy/sell/hold a share. With the results obtained model can compete reasonably well with emerging forecasting techniques in short-term prediction. Here we examined and applied multilayer perceptron model by using the Neural Networks Predict tool. The results from analysis shows that Neural Networks 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 have the ability to forecast the stock price 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.