STOCK MARKET PREDICTION

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering Degree in Computer Science and Engineering

By

sanjith,
(Reg. No. 38110501)
Romal Fernando
(Reg. No. 38110466)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SCHOOL OF COMPUTING
SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEDMED TO BE UNIVERSITY)
Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHI SALAI,
CHENNAI – 600119, TAMILNADU

APRIL 2021

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade "A" by NAAC
(Established under Section 3 of UGC Act, 1956)
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI– 600119
www.sathyabamauniversity.ac.in

SCHOOL OF COMPUTING

BONAFIDE CERTIFICATE

This is to certify that thiSProject Report is the bonafide work **sanjith,(Reg. No. 38110501)Romal Fernand(Reg. No. 38110466)**who carried out the project entitled **"STOCK MARKET PREDICTION"** under our supervision from Nov 2019 to April 2020.

Internal Guide
Dr. SUBHASHINI M.E., Ph.D.,

Head of the Department
Dr. S. VIGNESHWARI, M.E., Ph.D.,

Submitted for Viva voce Examination held on_____

Internal Examiner   **External Examiner**

DECLARATION

I, **sanjith,(Reg. No. 38110501),** hereby declare that the Professional Training Report on "**STOCK MARKET PREDICTION"** done underthe guidance of **Dr. SUBHASHINI M.E., Ph.D.,** at Sathyabama Institute of Science and Technology is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering.

DATE:

PLACE: CHENNAI   **SIGNATURE OF THE CANDIDATE**

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to Board of Management of **Sathyabama Institute of Science and Technology** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T. SASIKALA, M.E., Ph.D., Dean, School of Computing and Dr. S. VIGNESHWARI, M.E., Ph.D., Head of the Department**, **Department of Computer Science and Engineering** for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. SUBHASHINI M.E., Ph.D.,** for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of **Computer Science and Engineering** who were helpful in many ways for the Completion of the project.

ABSTRACT

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. Stock Market Analysis of stocks using data mining will be useful for new investors to invest in stock market based on the various factors considered by the software. Stock market includes daily activities like Sensex calculation, exchange of shares. The exchange provides an efficient and transparent market for trading in equity, debt instruments and derivatives. Our aim is to create software that analyses previous stock data of certain companies, with help of certain parameters that affect stock value. We are going to implement these values in data mining algorithms and we will be able to decide which algorithm gives the best result. This will also help us to determine the values that particular stock will have in near future. We will determine the patterns in data with help of machine learning algorithms.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	Application Programming Interface
CMS	Content Management System
CSS	Cascading Style Sheets
CRUD	Create, Read, Update, Delete
GIF	Graphics Interchange Format
GUI	Graphical User Interface
HTML	Hypertext Markup Language
MVC	Model View Controller
OOAD	Object Oriented Analysis and Design
RAM	Random Access Memory
SQL	Structured Query Language
UX	User Experience

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

In recent times stock market predictions is gaining more attention, maybe due to the fact that if the trend of the market is successfully predicted the investors may be better guided. The profits gained by investing and trading in the stock market greatly depends on the predictability. If there is a system that can consistently predict the direction of the dynamic stock market will enable the users of the system to make informed decisions. More over the predicted trends of the market will help the regulators of the market in taking corrective measures.

1.2 AIM AND OBJECTIVE

The aim of the project is to examine a number of different forecasting techniques 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 by interpreting the seemingly chaotic market data.

1.3 STOCK MARKET

A stock market, equity market or share market is the aggregation of buyers and sellers (a loose network of economic transactions, not a physical facility or discrete entity) of stocks (also called shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange as well as those only traded privately. Examples of the latter include shares of private companies which are sold to investors through equity crowd funding platforms. Stock exchanges list shares of common equity as well as other security types, e.g. corporate bonds and convertible bonds.

Stock price prediction is one of the most widely studied problem, attracting researchers from many fields. The volatile nature of the stock market makes it really difficult to apply simple time-series or regression techniques. Financial institutions and active traders have created various proprietary models to beat the market for

themselves or their clients, but rarely did anyone achieve consistently higher than the average returns on investment. The challenge of stock market price forecasting is so appealing because an improvement of just a few points of percentage can increase the profit by millions of dollars. This paper discusses the application of Support Vector Machines and Linear Regression in detail along with the pros and cons of the given methods. The paper introduces the parameters and variables which can be used to recognize the patterns in stock prices which can be helpful in future stock prediction and how boosting can be integrated with various other machine learning algorithms to improve the accuracy of our prediction systems.

1.4 MOTIVATION

Stock price prediction is a classic and important problem. With a successful model for stock prediction, we can gain insight about market behavior over time, spotting trends that would otherwise not have been noticed. With the increasingly computational power of the computer, machine learning will be an efficient method to solve this problem.

Thus, our motivation is to design a public service incorporating historical data and users predictions to make a stronger model that will benefit everyone.

CHAPTER 2

LITERATURE SURVEY

1. Survey of stock market prediction using machine learning approach
Authors: Ashish Sharma ; Dinesh Bhuriya ; Upendra Singh
2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)

Stock market is basically nonlinear in nature and the research on stock market is one of the most important issues in recent years. People invest in stock market based on some prediction. For predict, the stock market prices people search such methods and tools which will increase their profits, while minimize their risks. Prediction plays a very important role in stock market business which is very complicated and challenging process. Employing traditional methods like fundamental and technical analysis may not ensure the reliability of the prediction. To make predictions regression analysis is used mostly. In this paper we survey of well-known efficient regression approach to predict the stock market price from stock market data based. In future the results of multiple regression approach could be improved using more number of variables.

2. Short-term prediction for opening price of stock market based on self-adapting variant PSO-Elman neural network

Authors: Ze Zhang ; Yongjun Shen ; Guidong Zhang ; Yongqiang Song ; Yan Zhu, 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)

Stock price is one of intricate non-linear dynamic system. Typically, Elman neural network is a local recurrent neural network, having one context layer that memorizes the past states, which is quite fit for resolving time series issues. Given this, this paper takes Elman network to predict the opening price of stock market. Considering that Elman network is limited, this paper adopts self-adapting variant PSO algorithm to optimize the weights and thresholds of network. Afterwards, the optimized data, regarded as initial weight and threshold value, is given to Elman network for training, accordingly the prediction model for opening price of stock market based on self-

adapting variant PSO-Elman network is formed. Finally, this paper verifies that model by some stock prices, and compares with BP network and Elman network, so as to draw the result that shows the precision and stability of this predication model both are superior to the traditional neural network.

3. Combining of random forest estimates using LSboost for stock market index prediction

Authors: Nonita Sharma ; Akanksha Juneja,2017 2nd International Conference for Convergence in Technology (I2CT)

This research work emphases on the prediction of future stock market index values based on historical data. The experimental evaluation is based on historical data of 10 years of two indices, namely, CNX Nifty and S&P Bombay Stock Exchange (BSE) Sensex from Indian stock markets. The predictions are made for 1-10, 15, 30, and 40 days in advance. This work proposes to combine the predictions/estimates of the ensemble of trees in a Random Forest using LSboost (i.e. LS-RF). The prediction performance of the proposed model is compared with that of well-known Support Vector Regression. Technical indicators are selected as inputs to each of the prediction models. The closing value of the stock price is the predicted variable. Results show that the proposed scheme outperforms Support Vector Regression and can be applied successfully for building predictive models for stock prices prediction.

4. Using social media mining technology to assist in price prediction of stock market

Authors: Yaojun Wang ; Yaoqing Wang,2016 IEEE International Conference on Big Data Analysis (ICBDA)

Price prediction in stock market is considered to be one of the most difficult tasks, because of the price dynamic. Previous study found that stock price volatility in a short term is closely related to the market sentiment; especially for small-cap stocks. This paper used the social media mining technology to quantitative evaluation market segment, and in combination with other factors to predict the stock price trend in short term. Experiment results show that by using social media mining combined with other information, the stock prices prediction model can forecast

more accurate.

5. Stock market prediction using an improved training algorithm of neural network

Authors: Mustain Billah ; Sajjad Waheed ; Abu Hanifa,2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)

Predicting closing stock price accurately is an challenging task. Computer aided systems have been proved to be helpful tool for stock prediction such as Artificial Neural Net-work(ANN), Adaptive Neuro Fuzzy Inference System (ANFIS) etc. Latest research works prove that Adaptive Neuro Fuzzy Inference System shows better results than Neural Network for stock prediction. In this paper, an improved Levenberg Marquardt(LM) training algorithm of artificial neural network has been proposed. Improved Levenberg Marquardt algorithm of neural network can predict the possible day-end closing stock price with less memory and time needed, provided previous historical stock market data of Dhaka Stock Exchange such as opening price, highest price, lowest price, total share traded. Morever, improved LM algorithm can predict day-end stock price with 53% less error than ANFIS and traditional LM algorithm. It also requires 30% less time, 54% less memory than traditional LM and 47% less time, 59% less memory than ANFIS.

6. Efficacy of News Sentiment for Stock Market Prediction

Authors: Sneh Kalra ; Jay Shankar Prasad,2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)

Stock Market trend prediction will always remain a challenging task due to stochastic nature. The enormous amount of data generated by the news, blogs, reviews, financial reports and social media are considered a treasure of knowledge for researchers and investors. The present work focuses to observe fluctuations in stock prices with respect to the relevant news articles of a company. In this paper, a daily prediction model is proposed using historical data and news articles to predict the Indian stock market movements. Classifier Naïve Bayes is used to categorize the

news text having negative or positive sentiment. The count of the positive and negative sentiment of news articles for each day and variance of adjacent days close price along with historical data is used for prediction purpose and an accuracy ranging from 65.30 to 91.2 % achieved with various machine learning techniques.

7. Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools

Authors: Muhammad Firdaus ; Swelandiah Endah Pratiwi ; Dionysia Kowanda ; Anacostia Kowanda

This literature review is aiming to explore the use Artificial Neural Network (ANN) techniques in the field of stock market prediction. Design: Content analysis research technique. Data sources: Information retrieved from ProQuest electronic databases. Review methods: Utilizing key terms and phrases associated with Artificial Neural Network Stock Market Prediction from 2013-2018. Out of the 129 scholarly journal reviewed, there are 4 stock market studies met the inclusion criteria. The analysis and the evaluation includes 6 ANN derivatives techniques used to predict. Results: Findings from the reviewed studies revealed that all studies shows consistency that the accuracy rate of ANN stock market prediction is high. 2 Studies shows accuracy above 90%, 2 studies shows accuracy above 50%. Conclusion: This study reveals that the ability of ANN shows consistency of an accuracy rate of stock market prediction. Four method in predicting stock market had an accuracy above 95%. The highest accuracy achieved by using Signal Processing/Gaussian Zero-Phase Filter (GZ-Filter) with 98.7% prediction accuracy.
2018 Third International Conference on Informatics and Computing (ICIC)

8. Stock Market Movement Prediction using LDA-Online Learning Model

Authors:Tanapon Tantisripreecha ; Nuanwan Soonthomphisaj, 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)

In this paper, an online learning method namely LDA-Online algorithm is proposed to predict the stock movement. The feature set which are the opening price, the closing price, the highest price and the lowest price are applied to fit the Linear Discriminant

Analysis (LDA). Experiments on the four well known NASDAQ stocks (APPLE, FACBOOK GOOGLE, and AMAZON) show that our model provide the best performance in stock prediction. We compare LDA-online to ANN, KNN and Decision Tree in both Batch and Online learning scheme. We found that LDA-Online provided the best performance. The highest performances measured on GOOGLE, AMAZON, APPLE FACEBOOK stocks are 97.81%, 97.64%, 95.58% and 95.18% respectively.

9. Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment

Authors: Zhaoxia Wang ; Seng-Beng Ho ; Zhiping Lin, 2018 IEEE International Conference on Data Mining Workshops (ICDMW)

The price of the stocks is an important indicator for a company and many factors can affect their values. Different events may affect public sentiments and emotions differently, which may have an effect on the trend of stock market prices. Because of dependency on various factors, the stock prices are not static, but are instead dynamic, highly noisy and nonlinear time series data. Due to its great learning capability for solving the nonlinear time series prediction problems, machine learning has been applied to this research area. Learning-based methods for stock price prediction are very popular and a lot of enhanced strategies have been used to improve the performance of the learning based predictors. However, performing successful stock market prediction is still a challenge. News articles and social media data are also very useful and important in financial prediction, but currently no good method exists that can take these social media into consideration to provide better analysis of the financial market. This paper aims to successfully predict stock price through analyzing the relationship between the stock price and the news sentiments. A novel enhanced learning-based method for stock price prediction is proposed that considers the effect of news sentiments. Compared with existing learning-based methods, the effectiveness of this new enhanced learning-based method is demonstrated by using the real stock price data set with an improvement of performance in terms of reducing the Mean Square Error (MSE). The research work and findings

of this paper not only demonstrate the merits of the proposed method, but also points out the correct direction for future work in this area.

10. Stock Price Prediction Using News Sentiment Analysis
Authors:Vijayvergia ; David C. Anastasiu,2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)

Predicting stock market prices has been a topic of interest among both analysts and researchers for a long time. Stock prices are hard to predict because of their high volatile nature which depends on diverse political and economic factors, change of leadership, investor sentiment, and many other factors. Predicting stock prices based on either historical data or textual information alone has proven to be insufficient. Existing studies in sentiment analysis have found that there is a strong correlation between the movement of stock prices and the publication of news articles. Several sentiment analysis studies have been attempted at various levels using algorithms such as support vector machines, naive Bayes regression, and deep learning. The accuracy of deep learning algorithms depends upon the amount of training data provided. However, the amount of textual data collected and analyzed during the past studies has been insufficient and thus has resulted in predictions with low accuracy. In our paper, we improve the accuracy of stock price predictions by gathering a large amount of time series data and analyzing it in relation to related news articles, using deep learning models. The dataset we have gathered includes daily stock prices for S&P500 companies for five years, along with more than 265,000 financial news articles related to these companies. Given the large size of the dataset, we use cloud computing as an invaluable resource for training prediction models and performing inference for a given stock in real time. Index Terms-stock market prediction, cloud, big data, machine learning, regression.

CHAPTER 3

SYSTEM ANALYSIS

3.1 OBJECTIVES

The aims of this project are as to identify factors affecting share market, To generate the pattern from large set of data of stock market for prediction of NEPSE and to predict an approximate value of share price to provide analysis for users through web application

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.

3.2 EXISTING SYSTEM

Nowadays, as the connections between worldwide economies are tightened by globalization, external perturbations to the financial markets are no longer domestic. With evolving capital markets, more and more data is being created daily.

The intrinsic value of a company's stock is the value determined by estimating the expected future cash flows of a stock and discounting them to the present, which is known as the book value. This is distinct from the market value of the stock, that is determined by the company's stock price. This market value of a stock can deviate from the intrinsic value due to reasons unrelated to the company's fundamental operations, such as market sentiment.

The fluctuation of stock market is violent and there are many complicated financial indicators. Only few people with extensive experience and knowledge can understand the meaning of the indicators and use them to make good prediction to get fortune. Most people have to rely solely on luck to earn money from stock trading. However, the advancement in technology, provides an opportunity to gain steady fortune from stock market and also can help experts to find out the most informative indicators to make better prediction. The prediction of the market value is

of paramount importance to help in maximizing the profit of stock option purchase while keeping the risk low.

3.3 PROPOSED SYSTEM

Linear Regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). The case of one explanatory variable is called simple linear regression.

Advantages

➢ Space complexity is very low it just needs to save the weights at the end of training. Hence, it's a high latency algorithm.

➢ Its very simple to understand

➢ Good interpretability

Feature importance is generated at the time model building. With the help of hyperparameter lamba, you can handle features selection hence we can achieve dimensionality reduction

CHAPTER 4

SYSTEM IMPLEMENTATION

4.1 INTRODUCTION

Design is a multi- step that focuses on data structure software architecture, procedural details, procedure etc… and interface among modules. The design procedure also decodes the requirements into presentation of software that can be accessed for excellence before coding begins. Computer software design change continuously as novel methods; improved analysis and border understanding evolved. Software proposal is at relatively primary stage in its revolution.

Therefore, software design methodology lacks the depth, flexibility and quantitative nature that are usually associated with more conventional engineering disciplines. However, methods for software designs do exit, criteria for design qualities are existing and design notation can be applied.

4.2 SYSTEM REQUIREMENTS

4.2.1 Hardware Requirements

➢ PROCESSOR: PENTIUM IV

➢ RAM: 8 GB

➢ PROCESSOR: 2.4 GHZ

➢ MAIN MEMORY: 8GB RAM

➢ PROCESSING SPEED: 600 MHZ

➢ HARD DISK DRIVE: 1TB

➢ KEYBOARD :104 KEYS

4.2.2 Software Requirements

➢ FRONT END: PYTHON

➢ IDE: ANACONDA

➢ OPERATING SYSTEM: WINDOWS 10
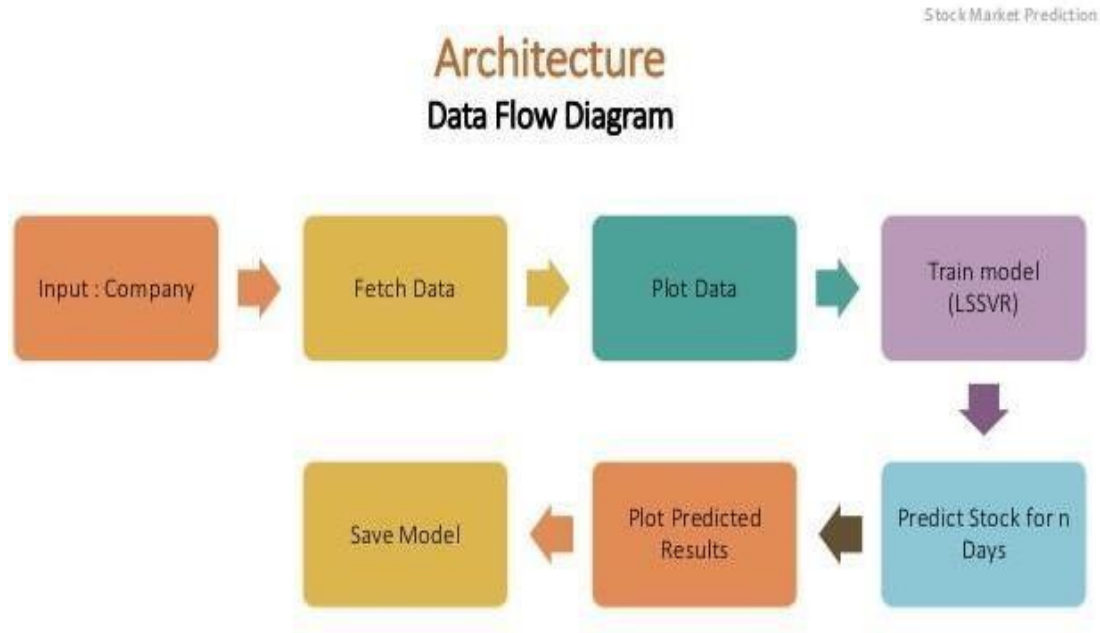
4.3 ARCHITECTURE
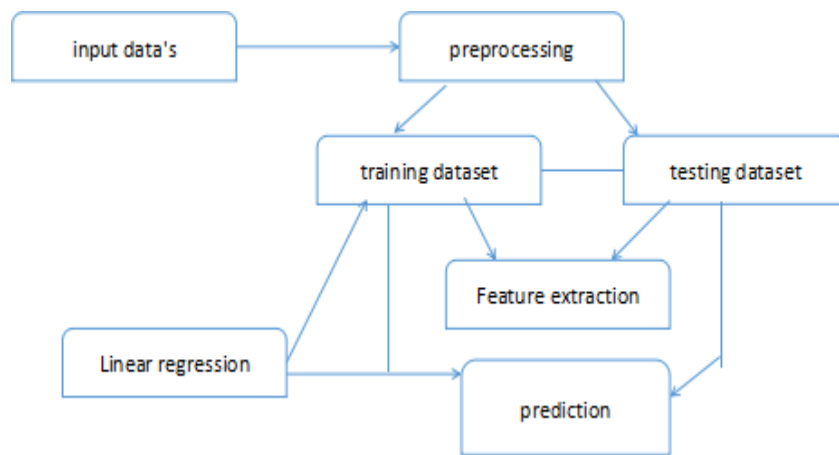

Fig 4.1 Data Flow Diagram

Fig 4.2 Architecture Design

4.4 MODULE DECRIPTION

The implementation of this project is divided into following steps

1. Data Preprocessing

2. Feature selection

3. Building and Traning Model

4.4.1 Data Preprocessing:

The entries are present in the dataset. The null values are removed using df = df.dropna() where df is the data frame. The categorical attributes (Date,High,Low,Close,Adj value) are converted into numeric using Label Encoder. The date attribute is splitted into new attributes like total which can be used as feature for the model.

4.4.2 Feature selection:

Features selection is done which can be used to build the model. The attributes used for feature selection are Date,Price,Adj close,Forecast X coordinate , Y coordinate, Latitude , Longitude, Hour and month,

4.4.3 Building and Training Model:

After feature selection location and month attribute are used for training. The dataset is divided into pair of xtrain ,ytrain and xtest, y test. The algorithms model is imported form skleran. Building model is done using model. Fit (xtrain, ytrain). This phase would involve supervised classification methods like linear regression, Ensemble classifiers (like Adaboost, Random Forest Classifiers), etc.

4.5 PYTHON TECHNOLOGY

Python is an interpreted, object- oriented programming language similar to PERL, that has gained popularity because of its clear syntax and readability. Python is said to be relatively easy to learn and portable, meaning its statements can be interpreted in a number of operating systems, including UNIX- based systems, Mac OS, MS- DOS, OS/2, and various versions of Microsoft Windows 98. Python was created by Guido van Rossum, a former resident of the Netherlands, whose favourite comedy group at the time was Monty Python's Flying Circus. The source code is freely available and open for modification and reuse. Python has a significant number of users.

A notable feature of Python is its indenting of source statements to make the code easier to read. Python offers dynamic data type, ready- made class, and interfaces to many system calls and libraries. It can be extended, using the C or C++language.

Python can be used as the script in Microsoft's Active Server Page (ASP) technology. The scoreboard system for the Melbourne (Australia) Cricket Ground is written in Python. Z Object Publishing Environment, a popular Web application server, is also written in the Python language's

4.5.1 Python Platform

Apart from Windows, Linux and MacOS, CPython implementation runs on 21 different **platforms**. IronPython is a .NET framework based **Python** implementation and it is cabable of running in both Windows, Linux and in other environments where .NET framework is available.

4.5.2 Python Library

Machine Learning, as the name suggests, is the science of programming a computer by which they are able to learn from different kinds of data. A more general definition given by Arthur Samuel is –"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed." They are typically used to solve various types of life problems.

In the older days, people used to perform Machine Learning tasks by manually coding all the algorithms and mathematical and statistical formula. This made the process time consuming, tedious and inefficient. But in the modern days, it is become very much easy and efficient compared to the olden days by various python libraries, frameworks, and modules. Today, Python is one of the most popular programming languages for this task and it has replaced many languages in the industry, one of the reason is its vast collection of libraries. Python libraries that used in Machine Learning are:

➢ Numpy

➢ Scipy

➢ Scikit- learn

- ➢ Theano

- ➢ TensorFlow

- ➢ Keras

- ➢ PyTorch

- ➢ Pandas

- ➢ Matplotlib

4.5.2.1 NumPy

NumPy is a very popular python library for large multi- dimensional array and matrix processing, with the help of a large collection of high- level mathematical functions. It is very useful for fundamental scientific computations in Machine Learning. It is particularly useful for linear algebra, Fourier transform, and random number capabilities. High- end libraries like TensorFlow uses NumPy internally for manipulation of Tensors.

4.5.2.2 SciPy:

SciPy is a very popular library among Machine Learning enthusiasts as it contains different modules for optimization, linear algebra, integration and statistics. There is a difference between the SciPy library and the SciPy stack. The SciPy is one of the core packages that make up the SciPy stack. SciPy is also very useful for image manipulation.

4.5.2.3 Skikit:

Skikit- learn is one of the most popular ML libraries for classical ML algorithms. It is built on top of two basic Python libraries, viz., NumPy and SciPy. Scikit- learn supports most of the supervised and unsupervised learning algorithms. Scikit- learn can also be used for data- mining and data- analysis, which makes it a great tool who is starting out with ML.

4.5.2.4 Theano:

We all know that Machine Learning is basically mathematics and statistics. Theano is a popular python library that is used to define, evaluate and optimize mathematical

expressions involving multi- dimensional arrays in an efficient manner. It is achieved by optimizing the utilization of CPU and GPU. It is extensively used for unit- testing and self- verification to detect and diagnose different types of errors. Theano is a very powerful library that has been used in large- scale computationally intensive scientific projects for a long time but is simple and approachable enough to be used by individuals for their own projects.

4.5.2.5 TensorFlow:

TensorFlow is a very popular open- source library for high performance numerical computation developed by the Google Brain team in Google. As the name suggests, Tensorflow is a framework that involves defining and running computations involving tensors. It can train and run deep neural networks that can be used to develop several AI applications. TensorFlow is widely used in the field of deep learning research and application.

4.5.2.6 Keras:

Keras is a very popular Machine Learning library for Python. It is a high- level neural networks API capable of running on top of TensorFlow, CNTK, or Theano. It can run seamlessly on both CPU and GPU. Keras makes it really for ML beginners to build and design a Neural Network. One of the best thing about Keras is that it allows for easy and fast prototyping.

4.5.2.7 PyTorch:

PyTorch is a popular open- source Machine Learning library for Python based on Torch, which is an open- source Machine Learning library which is implemented in C with a wrapper in Lua. It has an extensive choice of tools and libraries that supports on Computer Vision, Natural Language Processing(NLP) and many more ML programs. It allows developers to perform computations on Tensors with GPU acceleration and also helps in creating computational graphs.

4.5.2.8 Pandas:

Pandas is a popular Python library for data analysis. It is not directly related to Machine Learning. As we know that the dataset must be prepared before training. In

this case, Pandas comes handy as it was developed specifically for data extraction and preparation. It provides high- level data structures and wide variety tools for data analysis. It provides many inbuilt methods for groping, combining and filtering data.

4.5.2.9 Matpoltlib:

Matpoltlib is a very popular Python library for data visualization. Like Pandas, it is not directly related to Machine Learning. It particularly comes in handy when a programmer wants to visualize the patterns in the data. It is a 2D plotting library used for creating 2D graphs and plots. A module named pyplot makes it easy for programmers for plotting as it provides features to control line styles, font properties, formatting axes, etc. It provides various kinds of graphs and plots for data visualization, viz., histogram, error charts, bar chats, etc,

CHAPTER 5

RESULTS AND DISCUSSION

Screenshot

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

By measuring the accuracy of the Linear Regression algorithms, we found that the most suitable algorithm for predicting the market price of a stock based on various data points from the historical data. The algorithm will be a great asset for brokers and investors for investing money in the stock market since it is trained on a huge collection of historical data and has been chosen after being tested on a sample data. The project demonstrates the machine learning model to predict the stock value with more accuracy as compared to previously implemented machine learning models.

6.2 FUTURE WORK

Future scope of this project will involve adding more parameters and factors like the financial ratios, multiple instances, etc. The more the parameters are taken into account more will be the accuracy. The algorithms can also be applied for analyzing the contents of public comments and thus determine patterns/relationships between the customer and the corporate employee.

REFERENCES

1. Survey of stock market prediction using machine learning approach, Ashish Sharma ; Dinesh Bhuriya ; Upendra Singh 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)

2. Short-term prediction for opening price of stock market based on self-adapting variant PSO-Elman neural network,Ze Zhang ; Yongjun Shen ; Guidong Zhang ; Yongqiang Song ; Yan Zhu, 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)

3. Combining of random forest estimates using LSboost for stock market index prediction,Nonita Sharma ; Akanksha Juneja,2017 2nd International Conference for Convergence in Technology (I2CT)

4. Using social media mining technology to assist in price prediction of stock market,Yaojun Wang ; Yaoqing Wang,2016 IEEE International Conference on Big Data Analysis (ICBDA)

5. Stock market prediction using an improved training algorithm of neural network,Mustain Billah ; Sajjad Waheed ; Abu Hanifa,2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)

6. Efficacy of News Sentiment for Stock Market Prediction,Sneh Kalra ; Jay Shankar Prasad,2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)

7. Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools,Muhammad Firdaus ; Swelandiah Endah Pratiwi ; Dionysia Kowanda ; Anacostia Kowanda

8. Stock Market Movement Prediction using LDA-Online Learning Model,Tanapon Tantisripreecha ; Nuanwan Soonthomphisaj,2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking

and Parallel/Distributed Computing (SNPD) 9.Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment, Zhaoxia Wang ; Seng-Beng Ho ; Zhiping Lin 2018 IEEE International Conference on Data Mining Workshops (ICDMW)

10.Stock Price Prediction Using News Sentiment Analysis Saloni Mohan ; Sahitya Mullapudi ; Sudheer Sammeta ; Parag, Vijayvergia ; David C. Anastasiu,2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)