

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA**

**Department of Electrical and electronics engineering**

A Project Report

on

**PREDICTION OF STOCK MARKET USING MACHINE LEARNING ALONG WITH DATABASE SYSTEMS**

PRESENTED BY-

Devansh Baldwa

BTECH/10492/18

**Index**

i) Acknowledgement

1. **Introduction**
2. **Software Used**

**2.1 JetBrains PyCharm Community Edition**

**2.2 MySQL**

**2.3 Quandl.com**

1. **Literature Review**
2. **Working principle**
3. **Procedure**
4. **Code Implementation**
5. **Advantages and Disadvantages**
6. **Applications**
7. **Future Scope**

**ACKNOWLEDGEMENT**

The opportunity I had for creating this project was a great chance for learning and professional development. Therefore, I consider myself to be very lucky as I was provided with an opportunity to be a part of it.

I want to use this opportunity to express my deepest gratitude and special thanks to my seniors who in spite of being extraordinarily busy, took time out to hear, guide and keep me on the correct path and allowed me to carry out the project at this esteemed society.

**1. INTRODUCTION**

**PROBLEM**

Few people incur heavy losses in the Stock Market as they cannot decide when to buy or sell their stock and end up selling at a less profitable rate or even take losses. Also, the decision making cannot be very precise at times which could lead to the loss of the person, company or the firm.

**PRESENT SCENARIO**

Commonly, the Stock Market is considered to be risky and many people aren’t interested in it. Investments made by people are not systematic and generally influenced by the opinions of others, which are often wrong. Also, the accuracy and precision that can be provided by the computer cannot be provided by humans.

**SOLUTION**

We can use Machine Learning as a tool and can use it to benefit it us by predicting whether the price of a stock will rise or fall on the basis of its previous data. Also, we can link database systems to it which can help to maintain a record of the data and information of the stock we purchased, or we want to purchase. This also provides accuracy, more than the we can provide, thus saving us from the loss. Hence, this can be very beneficial for people investing in the Stock Market and also can help other people gain interest in the Stock Market.

**2. SOFTWARE USED**

**2.1 JETBRAINS PYCHARM COMMUNITY EDITION**

PyCharm is a Python IDE with complete set of tools for Python development. In addition, the IDE provides capabilities for professional Web development using the Django framework. Code faster and with more easily in a smart and configurable editor with code completion, snippets, code folding and split windows support.

**PyCharm Features**

* **Project Code Navigation -** Instantly navigate from one file to another, from method to its declaration or usages, and through classes hierarchy. Learn keyboard shortcuts to be even more productive
* **Code Analysis -** Take advantage of on-the-fly code syntax, error highlighting, intelligent inspections and one-click quick-fix suggestions to make code better
* **Python Refactoring -** Make project-wide code changes painlessly with rename, extract method/superclass, introduce field/variable/constant, move and pull up/push down refactorings
* **Web Development with Django -** Even more rapid Web development with Django framework backed up with excellent HTML, CSS and JavaScript editors. Also with CoffeeScript, Mako and Jinja2 support
* **Google App Engine Support -** Develop applications for Google App Engine and delegate routine deployment tasks to the IDE. Choose between Python 2.5 or 2.7 runtime
* **Version Control Integration -** Check-in, check-out, view diffs, merge — all in the unified VCS user interface for Mercurial, Subversion, Git, Perforce and other SCMs

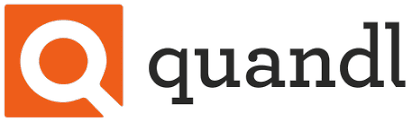
**2.2 MySQL**

MySQL, pronounced either "My S-Q-L" or "My Sequel," is an open source relational database management system. It is based on the structure query language (SQL), which is used for adding, removing, and modifying information in the database. Standard SQL commands, such as ADD, DROP, INSERT, and UPDATE can be used with MySQL.

MySQL can be used for a variety of applications, but is most commonly found on Web servers. A website that uses MySQL may include Web pages that access information from a database. These pages are often referred to as "dynamic," meaning the content of each page is generated from a database as the page loads. Websites that use dynamic Web pages are often referred to as database-driven websites.

Many database-driven websites that use MySQL also use a Web scripting language like PHP to access information from the database. MySQL commands can be incorporated into the PHP code, allowing part or all of a Web page to be generated from database information. Because both MySQL and PHP are both open source (meaning they are free to download and use), the PHP/MySQL combination has become a popular choice for database-driven websites.

**2.3 Quandl.com**

Quandl is a marketplace for financial, economic and alternative data delivered in modern formats for analysts, including Python, Excel, Matlab, R, etc.

**­**

**3. Literature review**

This project was made using PyCharm where the Python code for prediction of stock market using Machine Learning was written. This was also linked to MySQL where the code for the database management system was written. The data of the Stock Prices along with the dates was downloaded from Quandl.

The Machine Learning can be used by different algorithms which can depend on the accuracy; but here we don’t know the prediction is accurate or no, so we can use the algorithms which are more complex and depend upon better logics for prediction since that increases the accuracy of the prediction.

The Database Systems used for the project in MySQL is based upon creating tables since the data is in rows and columns. The rows contain different values of the stock at different times. The columns contain different values of the stock at a certain time. The rows and columns of the database system can be changed according to our wish and convenience at any time we want.

Quandl is free data providing site which is useful to collect data. This helps us to get the past data and performance of a company, which is used for prediction.

**4. WORKING PRINCIPLE**

Data of stock values of a stock (in a given interval of time) is downloaded or provided as an input to our program. The program uses the data and analyses it with the help of different packages in Python (the language in which our code is written) like Numpy, Pandas, Matplotlib, etc. Then, using Machine Learning algorithms, the program predicts the values of our stock. This prediction is plotted as a graph, which helps to visualize the prediction values of the stock. Then, we can store the present value, stock name, company name, etc. in a database system using MySQL. The prediction values of the stock can be stored in a separate column, thus enabling us to have complete info of a stock and making it easier for us to decide whether to buy or sell the stock at a given time.

**5. Procedure**

**Step 1:** Select a particular stock whose stock value you want to predict. Download the data of the stock from Quandl.com

**Step 2:** Input the data into the program.

**Step 3:** The program processes the data, and the predicted values are shown.

**Step 4:** This predicted value can be noted. If the value of stock is predicted to rise, then the value can be noted.

**Step 5:** The output from the data can be seen and entered in MySQL.

**Step 6**: This data of the and its predicted value can be referred anytime, according to our need.

**6. Code Implementation**

**PYTHON CODE**

import csv

import numpy as np

from sklearn import linear\_model

import matplotlib.pyplot as plt

from datetime import datetime

dates = []

prices = []

def get\_data(filename):

with open(filename, 'r') as csvfile:

csvFileReader = csv.reader(csvfile)

next(csvFileReader)# skipping column names

for row in csvFileReader:

dates.append(float(row[0]))

prices.append(float(row[1]))

return

def show\_plot(dates, prices):

linear\_mod = linear\_model.LinearRegression()

dates = np.reshape(dates, (len(dates), 1)) # converting to matrix of n X 1

prices = np.reshape(prices, (len(prices), 1))

linear\_mod.fit(dates, prices) # fitting the data points in the model

plt.scatter(dates, prices, color='yellow') # plotting the initial datapoints

plt.plot(dates, linear\_mod.predict(dates), color='blue', linewidth=3) # plotting the line made by linear regression

plt.show()

return

def predict\_price(dates, prices, x):

linear\_mod = linear\_model.LinearRegression() # defining the linear regression model

dates = np.reshape(dates, (len(dates), 1)) # converting to matrix of n X 1

prices = np.reshape(prices, (len(prices), 1))

linear\_mod.fit(dates, prices) # fitting the data points in the model

predicted\_price = linear\_mod.predict(x)

return predicted\_price[0][0], linear\_mod.coef\_[0][0], linear\_mod.intercept\_[0]

get\_data('SHARADAR-SEP.csv') # calling get\_data method by passing the csv file to it

print(dates)

print(prices)

print("\n")

show\_plot(dates, prices)

**MySQL CODE**

SELECT VERSION();

secure\_file\_priv=/var/lib/mysql-files/

SELECT @@GLOBAL.secure\_file\_priv;

CREATE TABLE database1 (

dates FLOAT PRIMARY KEY,

price FLOAT,

predicted FLOAT

);

LOAD DATA INFILE 'C:\Users\devan\Desktop\SHARADAR-SEP.csv'

INTO TABLE database1

FIELDS TERMINATED BY ','

ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

SELECT \* FROM database1;

**7. advantages and Disadvantages**

**ADVANTAGES-**

**SAVINGS**

It can help us save the cost of a broker or a consultancy firm.

**CONVENIENCE**

It can be used just by a laptop and can be used at home itself. It is convenient as reduces the effort of going to an office or contacting a broker, etc.

**COMFORT**

Since it can be used anywhere, it creates a comfortable atmosphere for the user and can benefit the user because of the suitable environment. S

**EASE TO USE**

It is very easy to use as we only have to give the program an input which can be downloaded easily.

**DISADVANTAGES-**

**ACCURACY**

The predicted values necessarily need not be accurate all the time. Sometimes the predicted value can be positive, but the value of stock falls, which can cause the users some loss.

**DATA**

If the data inputted is not proper or of a long period of time, then the values predicted can be misleading and could give inappropriate results.

**RELIABILITY**

It cannot be completely relied upon since there could be errors and a large amount being invested based upon this project can be risky.

**8. Applications**

1. It can be helpful for the common people with no much knowledge about the stock market, since it makes buying and selling of stocks easier for them. This can also help them increasing their profit margin without putting much effort.

1. It can also be used in FinTech companies (Financial Technology companies i.e. the technology and innovation that aims to compete with traditional financial methods in the delivery of financial services) in managing databases along with Stock Market prediction also.
2. Advanced versions of this project can be used by consultancy firms for them to recommend their customers what stocks to invest in and get a commission out of the profit made, since the prediction has a good accuracy at high levels.
3. It can also be used by the companies itself to analyse their performance and bring changes to improve.

**9. Future scope**

This project has lots of applications and is presently being used in different fields. Even then it has a bright future ahead and there is lots of scope of this project in the future.

It can be used with neural networks and Deep Learning to make a better and advanced version of it to make it much more useful and profitable for its users.

Also, it can bring a change or probably can revolutionize the way stock market works, because if everybody predicts right then there will be unevenness in the stock market.

It can also lead to the rise of new start ups which have their ideas based on this project.