# C:\Users\Indrajeet Kaur\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\28C89F58.tmp

**Medi-Caps University, Indore**

**Department of Computer Science & Engineering**

**August, 2019**

**SYNOPSIS**

**ON**

Stock Market Analysis and Predication

using Python

**Project Guide:**  **Submitted By:**

Pritesh Saklecha Yuvraj Yadav(EN16CS301300)

(Assistant Professor) Mangalam Tripathi(EN17CS3L1002)

Utkarsh Jain(EN16CS301280)

**Submitted To:**

Sachin Solanki

(Project Coordinator)

**Department of Computer Science and Engineering**

**August, 2019**

**STOCK MARKET ANALYSIS AND PREDICTION USING PYTHON**

Predicting the Stock Market has been the bane and goal of investors since its existence. Everyday billions of dollars are traded on the exchange, and behind each dollar is an investor hoping to profit in one way or another. Entire companies rise and fall daily based on the behaviour of the market. Should an investor be able to accurately predict market movements, it offers a tantalizing promises of wealth and influence. Our aim is to create a Application that analyzes 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 data mining algorithms. The Application will have functionality of choosing the date intervals, showing graphs and generating a Analysis and predicgion reports in form of Jupyter Notebook.

The application will be developed using Python 3.\* and Jupyter Notebook for providing the analysis report.

**INTRODUCTION**

**Stock market analysis and 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. Others disagree and those with this viewpoint possess myriad methods and technologies which purportedly allow them to gain future price information.

The **efficient market hypothesis** posits that stock prices are a function of information and rational expectations, and that newly revealed information about a company's prospects is almost immediately reflected in the current stock price. This would imply that all publicly known information about a company, which obviously includes its price history, would already be reflected in the current price of the stock. Accordingly, changes in the stock price reflect release of new information, changes in the market generally, or random movements around the value that reflects the existing information set.

Some of the stock exchanges available are Newyork stock exchange, Shanghai stock exchange, etc,.

**1.1 Stocks and Stock Market**

**1.1.1 Stock:** The stock (also capital stock) of a corporation is all of the shares into which ownership of the corporation is divided. In American English, the shares are commonly known as "stocks". A single share of the stock represents fractional ownership of the corporation in proportion to the total number of shares.

**1.1.2 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, which represent ownership claims on businesses. These may include *securities* listed on a public stock exchange, as well as stock that is only traded privately.

**1.2 Artificial Intelligence and Applications**

Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction. Particular applications of AI include expert systems, speech recognition, machine vision.

**1.2.1 Machine Learning:** Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

**1.2.2 Data Mining:** Data mining is an important branch of Artificial intelligence (AI) that searches for valuable information in a large database commonly known as knowledge discovery. Due to significant strides in information processing technology there exists several data mining techniques like clustering, classification, prediction, association, characterization, rule generation, and pattern recognition, etc. for a variety of databases that include object oriented, relational, temporal, transactional types. The object of study in the present research is the time series database which is an interesting and important area of research as it belongs to many real world problems occurring in finance, industry, energy markets, medical diagnostics, Geo-technical engineering, etc.

**1.2.3 Artificial Neural Networks:** Artificial neural networks (ANN) or connectionist systems are computing systems that are inspired by, but not identical to, biological neural networks that constitute animal brains. Such systems "learn" to perform tasks by considering examples, generally without being programmed with task-specific rules.

**1.3 Database and APIs**

Database is a systematic collection of data. Databases support storage and manipulation of data. Databases make data management easy. Let's discuss few examples. An online telephone directory would definitely use database to store data pertaining to people, phone numbers, other contact details, etc. Your electricity service provider is obviously using a database to manage billing , client related issues, to handle fault data, etc.

An application programming interface (API) is an interface or communication protocol between a client and a server intended to simplify the building of client-side software. It has been described as a “contract” between the client and the server, such that if the client makes a request in a specific format, it will always get a response in a specific format or initiate a defined action.

**PROBLEM DOMAIN**

Stock Exchange classification is important parts in the fields of Stock exchange marketplace analysis. Among them, method for stock exchange classification based on payload signature promises high classification accuracy rate. But there is disadvantages in payload signature based analysis that it requires high processing load comparing to other analysis method. Therefore, as the exchange data volume is increased the application stock data exchange classification using payload signature suffers from the burden in memory, processing speed, storage space. The complete analysis observes that it need to implement exclusive grouping mechanism to handle large data set and generate more accurate traffic analysis output.

**2.1 Automatic database update**

Since the stock price changes on minute basis. The data we hold will not be sufficient to predict the future values and there for updating of database is necessary. In this project we will be using a Kaggle dataset of stock prices that includes date, open, high, low, close and volume of the traded stock units. A quality dataset is must for stock market as our whole process and project outcome will be depending on it.

**2.2 Stock analysis**

Stock analysis is the evaluation of a particular trading instrument, an investment sector, or the market as a whole. Stock analysts attempt to determine the future activity of an instrument, sector, or market. The two basic types of stock analysis techniques are:

**2.2.1** **Fundamental Analysis:** Fundamental analysis concentrates on data from sources, including financial records, economic reports, company assets, and market share. To conduct fundamental analysis on a public company or sector, investors and analysts typically analyze the metrics on a company’s financial statements – balance sheet, income statement, cash flow statement, and footnotes.

**2.2.2** **Technical Analysis:** Technical analysis focuses on the study of past market action to predict future price movement. Technical analysts analyze the financial market as a whole and are primarily concerned with price and volume, as well as the demand and supply factors that move the market.

**2.3 Financial Modelling:**

Financial modeling is the process of creating a summary of a company's expenses and earnings in the form of a spreadsheet that can be used to calculate the impact of a future event or decision. A financial model has many uses for company executives. Financial analysts most often use it to analyze and anticipate how a company's stock performance might be affected by future events or executive decisions.

**2.4 Stock price prediction**

Predicting how the stock market will perform is one of the most difficult things to do. There are so many factors involved in the prediction – physical factors vs. physhological, rational and irrational behaviour, etc. All these aspects combine to make share prices volatile and very difficult to predict with a high degree of accuracy. Actual price of stock cannot be predicted but movement of price can be predicted. Prediction of movement of stock price can be done in following steps :

* Fundamental Analysis
* Volume Breakout
* Moving Averages
* Derivatives Segment
* Candlestick Patterns
* Human Intelligence

  
Illustration 1: A candlestick shows an asset’s price movement over a set amount of time

**2.5 Application**

A Application will be build to provide better GUI making user easier to access, use and interact with our work.

**2.6 Related Work**

**2.6.1 Stock Prediction Using Convolutional Neural Network:**

Machine learning is a very good way to make prediction in financial area. It has used a SVM wayto forecast financial time series data, and the experiment result of which has indicated that the use of support vector machine (SVM) to make prediction can be a good idea. Deep learning methods to deal problem has shown its superiority when compared with previous machine learning methods. The generally used model in deep learning includes two parts: Convolutional neural network and Recurrent neural network. Accuracy of the model 73.2%.

**2.6.2 Predicting the Direction of Stock Market Price Using Tree Based Classifiers:**

The relevant literature survey revealed that SVM has been used most of the time in stock prediction research. The sensitivity of stock prices to external conditions have been considered by Li et al. It was found that logistic regression turned out to be the best model with a success rate of 55.65%. In the paper by Dai and Zhang (2013), the data used for the analysis were stock (closing) prices of the company 3M. A Decision Tree classifiers is used in this one resulted in an accuracy of 68-75%.

**2.6.3 Other related works:**

* In the paper by Dai and Zhang (2013), the data used for the analysis were stock (closing) prices of the company 3M. It was found that logistic regression turned out to be the best model with a success rate of 55.65%.
* An extremely randomizedtree algorithm, as described by Geurts and Louppe (2014), was implemented for the selection of the most relevant features. These features were then fed to an SVM with RBF kernel for training. Devi et al. (2015) has proposed a model which uses a hybrid cuckoo search with support vector machine. Success rate :
* Boonpeng and Jeatrakul (2015) implemented a one-against-all (OAA-NN) and one-against-one neural network (OAO-NN) to classify buy, hold, or sell data and compared their performance with a traditional neural network. Historical data of Stock Exchange of Thailand for seven years was selected. It was found that OAA-NN performed better than OAO-NN and traditional NN models, producing an average accuracy of 72.50%.

**SOLUTION DOMAIN**

The data mining is a technique of software application that is used to analyse the huge amount of data. This approach becomes more crucial when data source found too large. A clustering based traffic analysis is proposed instead of signature based classification to observe the packet data set.

For building a machine learning model we will be diving into Deep Learning methods that is will be using Artificial neural networks in it. LSTM and ARIMA are to of the mostly used ones in stock market prediction. LSTM shows the best result as per the research.

|  |  |  |
| --- | --- | --- |
| **Problem Title** | **Problem Investigation** | **Proposed Solution** |
| 1. Automatic database update | Datasets are in CSV file format. They are in separate files and therefore can be updated when called. | Using Alpha Vantage’s US Stock market API. Datasets can be updated by timeseries daily trading object. |
| 2. Stock analysis | Correlations and dependencies between different stocks, stock performance and real world entities do exist. Analyzing them is a key part. | Building graphs and charts using matplotlib and plotline liberaries in python, correlations and movement can be observed. |
| 3. Financial Modeling | A companies expanses, earnings and future events plays major role in stock price manipulation. Competition between companies do effect stock prices. | Using Microsoft Excel’s inbuilt function for simple financial modeling, we can drive few insights. |
| 4. Stock price prediction | Stock prices shows some relation with Moving average, trade quantity, previous stock performance and company’s total revenue. | ARIMA and LSTM are one of the known architectures and models in stock market prediction. This ANNs work best as per research. |
| 5. Application | A application so that user can interact with our report and findings and use it. Features like selecting date intervals, daily, monthly, yearly analysis,etc,. Are some of them. | Python Django and Tkinter will be used for build the application. |

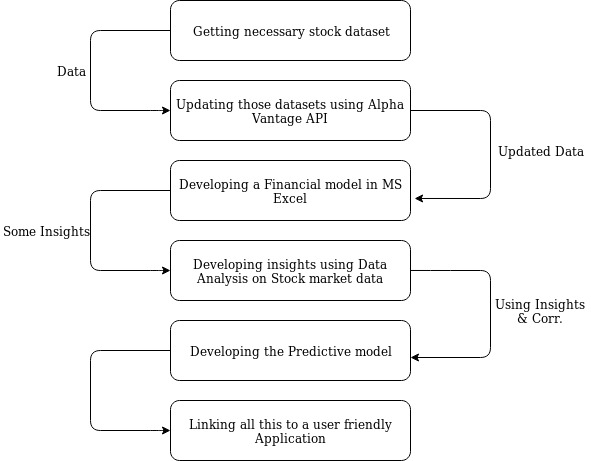
**3.1 Requirements**

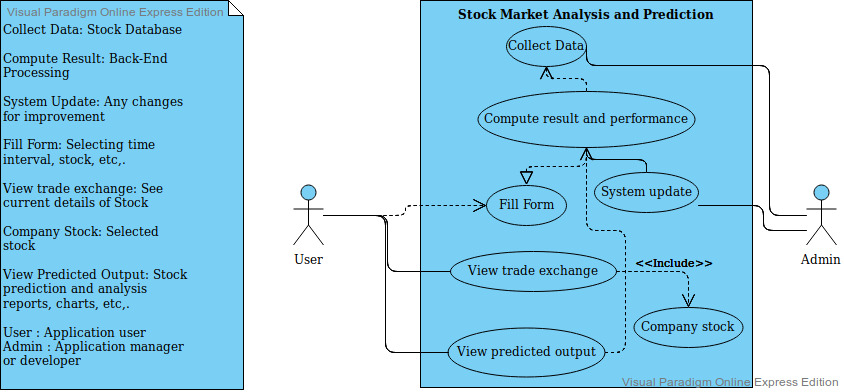
Requirements plays a major role in building a software project. Requirement of knowledge and expertise in the developing team.

* Machine learning algorithms
* Artificial neural networks
* Python, HTML, LaTex, Microsoft excel.
* Basic about Stock Market.

**3.2 Software Planning**

We will be following waterfall model in developing this entire project on Stock Market Prediction and Analysis. Steps in developing phase taken as per the flowchart given below.



**3.3 Architecture**

**3.4 Algorithms and Techniques**

**3.4.1 k-Nearest Neighbors:** the *k*-nearest neighbors algorithm (*k*-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space. *k*-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification.

**3.4.2 ARIMA:** ARIMA is a very popular statistical method for time series forecasting. ARIMA stands for Auto-Regressive Integrated Moving Averages. ARIMA has three components – AR (autoregressive term), I (differencing term) and MA (moving average term).

**3.4.3 LSTM:** Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data.

**SYSTEM REQUIREMENTS**

**4.1 Software Requirements**

1. Python3 (\*with Libraries like matplotlib, pandas, plotnine, etc,. installed)
2. Jupyter Notebook
3. Web Browser (firefox, chrome, etc,.)

**4.2 System Hardware Requirements**

1. Processors: Intel Atom® processor or Intel® Core™ i3 processor
2. Disk space: 2 GB
3. Operating systems: Windows\* 7 or later, macOS, and Linux
4. Monitor (21-inch recommended)
5. RAM : 1 GB
6. Internet Connection

**APPLICATION DOMAIN AND ENHANCEMENTS**

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. Though this version of this application can’t be exactly used for investing in stock markets because prediction can go wrong a lot of time and risk is too high. But this application can help its user to get some insights and with his/her sense, invest can be done.

**5.1 Applications**

* For getting insights on any stock
* Noting performance of a stock
* Generate a quality stock analysis report
* Generate a quality stock prediction report

**5.2 Enhancements**

Enhancements that can be done in this applications are:

* Comparative stocks analysis can be added.
* Multiple stocks can simultaneously be shown on a single graph or chart.
* A PDF version of reports can be made downloadable.
* Stock market data from more than one stock exchange can be compared.