CHAPTER 1 : INTRODUCTION

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION TO PROJECT:

Stock Price Prediction has been at focus for years since it can yield significant profits. Predicting the stock market is not a simple task, mainly as a consequence of the close to random-walk behavior of a stock time series. Fundamental and technical analyses were the first two methods used to forecast stock prices. Artificial Neural networks (ANNs) is the most commonly used technique. In most cases ANNs suffer from over-fitting problem due to the large number of parameters to fix, and the little prior user knowledge about the relevance of the inputs in the analyzed problem.

Also, Support vector machines (SVMs) had been developed as an alternative that avoids such limitations. Their practical successes can be attributed to solid theoretical foundations based on VC-theory. SVM compute globally optimal solutions, unlike those obtained with ANNs, which tend to fall into local minima. Least squares –support vector machines (LS-SVM) method was presented in, which was reformulated the traditional SVM algorithm. LS-SVM uses a regularized least squares function with equality constraints, leading to a linear system which meets the Karush-Kuhn-Tucker (KKT) conditions for obtaining an optimal solution. Although LS-SVM simplifies the SVM procedure, the regularization parameter and the kernel parameters play an important role in the regression system. Therefore, it is necessary to establish a methodology for properly selecting the LS-SVM free parameters, in such a way that the regression obtained by LS-SVM must be robust against noisy conditions, and it does not need priori user knowledge about the influence of the free parameters values in the problem studied.

The perceived advantages of evolutionary strategies as optimization methods motivated some researchers to consider such stochastic methods in the context of optimizing SVM. A survey and overview of evolutionary algorithms (EAs) found in. Particle swarm optimization (PSO) is one of the most used EAs. PSO is a recently proposed algorithm by James Kennedy and Russell Eberhart in 1995, motivated by social behavior of organisms such as bird flocking and fish schooling. The optimizer which is used in the particle swarm optimization algorithm, while making adjustment towards "local" and "global" best particles, is conceptually similar to the crossover operation used by genetic algorithms. As well particle swarm optimization includes fitness function, which measures the closeness of the corresponding solution to the optimum. The main difference of particle swarm optimization concept from the evolutionary computing is that flying potential solutions through hyperspace are accelerating toward "better" solutions, while in evolutionary computation schemes operate directly on potential solutions which are represented as locations in hyperspace. SVM was used in stock market forecasting in. Financial time series forecasting using SVM optimized by PSO was presented in. The Optimization of Share Price Prediction Model Based on Support Vector Machine is presented in. Financial time series forecasting based on a wavelet kernel support vector was presented in.

Computational Intelligence Approaches for Stock Price Forecasting was introduced in. A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting was presented in. An interval type-2 Fuzzy Logic based system for modeling generation and summarization of arbitrage opportunities in stock markets was presented in. Robust stock trading using fuzzy decision trees is presented in. Ensemble ANNs-PSO-GA Approach for Day-ahead Stock E-exchange Prices Forecasting was presented in.

Index prediction with Neuro - genetic hybrid network was proposed in. A hybrid fuzzy intelligent agent-based system for stock price prediction was introduced

in. Improved Stock Market Prediction by Combining Support Vector Machine and Empirical Mode Decomposition was presented in. Neural Network Ensemble Model Using PPR and LS-SVR for Stock Market Forecasting was proposed in. Computational Intelligence Techniques for Risk Management in Decision Making was introduced in. Stock market prediction algorithm using Hidden Markov Models was proposed presented in. Neural Networks and Wavelet De-Noising for Stock Trading and Prediction was introduced in. The aim of this paper is to develop a machine learning model that hybrids the PSO and LS-SVM model. The performance of LS-SVM is based on the selection of free parameters C (cost penalty), ϵ (insensitive-loss function) and γ (kernel parameter). PSO will be used to find the best parameter combination for LS-SVM.

1.2 STATEMENT OF THE PROBLEM:

Investors are familiar with the saying, “buy low, sell high” but this does not provide enough context to make proper investment decisions. Before an investor invests in any stock, he needs to be aware how the stock market behaves.Investing in a good stock but at a bad time can have disastrous results, while investment in a mediocre stock at the right time can bear profits.

Financial investors of today are facing this problem of trading as they do not properly understand as to which stocks to buy or which stocks to sell in order to get optimum profits. Predicting long term value of the stock is relatively easy than predicting on day-to-day basis as the stocks fluctuate rapidly every hour based on world events.

Broadly, stock market analysis is divided into two parts
- Fundamental Analysis
- Technical Analysis.

Fundamental Analysis involves analyzing the company’s future profitability on the basis of its current business environment and financial performance.

Technical Analysis, on the other hand, includes reading the charts and using statistical figures to identify the trends in the stock market.

As you might have guessed, our focus will be on the technical analysis part. We'll using a dataset from Yahoo Finance (you can find historical data for various stocks here) and for this particular project, we have used the data for "*Apple Incorporation [AAPL]*"

According to the Firm Foundation theory the market is defined from the reaction of the investors, which is triggered by information that is related to the "real value" of firms. The "real value" or else the intrinsic value is determined by careful analysis of present conditions and future prospects of a firm. On the other hand, according to the Castles in the Air theory the investors are triggered by information that is related to other investor's behavior. So, for this theory the only concern the investor should have is to buy today with the price of 20 and sell tomorrow with the price of 30, no matter what the intrinsic value of the firm he (or she) invests in is. Therefore, the Firm Foundation theory favor the view that the market is defined mostly by logic, while the Castles in the Air theory supports that the market is defined mostly by psychology. Still there are some aspects which affect the accuracy of the predictions as they have some serious impact on the Stock Prices like

1. Global Indices
2. Indian Indices
3. Currency Prices
4. Sector Behavior
5. Market Movers
6. News:
    a. Economic Policies
    b. Political News
    c. Natural Disasters

Trading in shares is big business in many economies including BSE. Currently, Stockbrokers who execute trades and advise clients, rely on their experience, technical analysis (price trends) or fundamental analysis (buy and hold) in picking their stocks. These current methods are subjective and are usually short sighted due to their limited capacity. With the value of trade money involved, improper investment could easily mean great losses to investors, especially if they keep making wrong decisions. Lack of guaranteed returns has also led to the reluctance by potential investors to participate in the market. It is therefore desirable to have a tool that can guide on the most likely next day prices (prediction) as a basis of making any investment decision. The use of fundamental and technical analysis methods are on basis of the predictions of future stock price movement. These tools show a trend on future movement and not the figure of the most likely trade price for any stock in future. It is there- fore desirable to have a tool that does not just point at a direction of price movement. Machine Learning methods that can actually analyze the stock prices over time and gain intelligence, then use this intelligence in prediction, can be used to model such a tool.

Financial analysts investing in stock market usually are not aware of the stock market behavior. They are facing the problem of trading as they do not properly understand which stocks to buy or which stocks to sell in order to get more profits. In today’s world, all the information pertaining to stock market is available. Analyzing all this information individually or manually is tremendously difficult. As such, automation of the process is required. This is where Data mining techniques help. Understanding that analysis of numerical time series gives close results, intelligent investors use machine learning techniques in predicting the stock market behavior. This will allow financial analysts to foresee the behavior of the stock that they are interested in and thus act accordingly. The input to our system will be historical data from Yahoo Finance. Appropriate data would be applied to find the stock price trends. Hence the prediction model will notify the up or down of the stock price movement for the next trading day and investors can act upon it so as to maximize their chances of gaining a profit. The entire system would be

implemented in Python/Java and R language using open source libraries. Hence it will effectively be a zero-cost system.

We’ll write a python program that predicts the price of stocks using two different machine learning algorithms, one is called a Support Vector Regression (SVR) and the other is Linear Regression. Also, we’ll use TensorFlow to write the neural network program to predict the stocks.

CHAPTER 2 : SYSTEM ANALYSIS

CHAPTER 2
SYSTEM ANALYSIS

2.1 EXISTING SYSTEM:

Recently, a lot of interesting work has been done in the area of applying Machine Learning Algorithms for analyzing price patterns and predicting stock price. Most stock traders nowadays depend on Intelligent Trading Systems which help them in predicting prices based on various situations and conditions. Recent researches uses input data from various sources and multiple forms. Some systems use historical stock data, some use financial news articles, some use expert reviews while some use a hybrid system which takes multiple inputs to predict the market.

Also, a wide range of machine learning algorithms are available that can be used to design the system. These systems have different approaches to solve the problem. Some systems perform mathematical analysis on historic data for prediction while some perform sentiment analysis on financial news articles and expert reviews for prediction. However, because of the volatility of the stock market, no system has a perfect or accurate prediction.

2.1.1 KNOWLEDGE DISCOVERY IN DATABASES (KDD)

The importance of the knowledge from the databases had been introduced in the workshop on KDD in 1989 (Frawley et. al., 1991). Knowledge discovery is the extraction of useful, unknown information from the data. This work shop on KDD started cultivation of many technologies in KDD. The data was being collected and stored for knowledge extraction. But the existing technology,

methods and tools were not state-of-the-art level for handling rapidly growing data.

The KDD is considered to be the intersection of databases, artificial intelligence, pattern recognition, information retrieval and expert systems (Fayyad et al., 1996).Data Mining had been evolved as a step in KDD. Data mining refers to the process of producing the useful patterns, by applying the data analysis methods and algorithms. These methods and algorithms take computational efficiencies into consideration (Parker et.al.,1998).

The various steps in the KDD process are shown in the Figure 2.1. In line with KDD, KDT (Knowledge Discovery in Text) showed its existence as well. According to the definition given by (Fayyad, 1996), KDT is the process of identifying the useful, novel and understandable patterns in the unstructured text data. The term KDT was also used by (Karanikas, 2002)

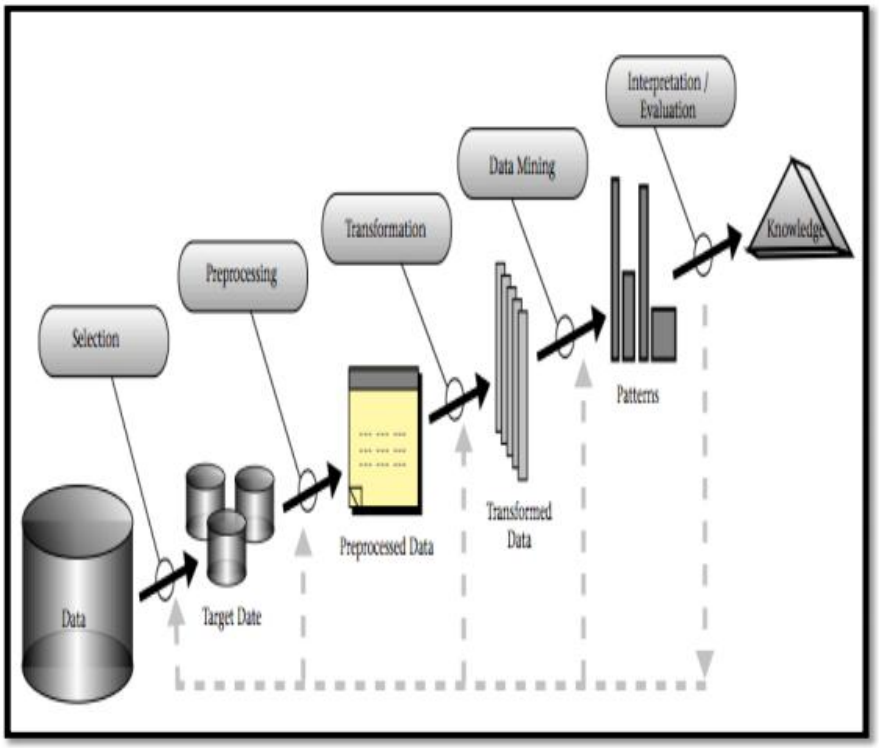


Fig 2.1 Overview of steps in KDD process

2.1.2 METHODS OF STOCK MARKET ANALYSIS

The investors use three main methods for analysing the stock market:
Technical, Fundamental and Sentiment analysis of news articles.

- **Fundamental Analysis**

Fundamental analysis(Abarbanell et.al.,1997) is a way to evaluate the stock for predicting the stock price movement. It uses a method called “financial analysis” to achieve the same. The information that has to be taken into consideration, for analysis, includes the annual financial statements and reports of the company, its balance sheet, its health, its future prospects, industry comparisons, market environment and changes in the government policies etc. Fundamental analysis examines the firm’s financial statements to decide upon its worth, to invest in the stock of the firm. Financial statements indicate cash flow, income as well as the balance sheets. This kind of information helps the investors to get some knowledge about the financial makeup of the company behind the stock.

The balance sheet shows the owner’s equity, the assets as well as the liabilities. Assets are the properties that the company owns which also have a potential to provide future value. It consists of the properties as well as the cash. Liabilities mostly comprise of the mortgages and the debts among a few others. Owner’s equity is termed as the amount of money raised by issuing the stocks to investors. Balance sheets show the investor, the way that company raises its money. Income statements do a similar job; they show the revenues and expenses of that company. These can be considered as the costs associated to run a business. The net income is calculated using the difference of the revenues and the expenses; this is essentially the earning of the company. The cash flow statements display how the company uses its cash for operations and making investments.

The stock market investor can use all these facts and figures to decide upon the feasibility to invest on that particular company. They can additionally use the ratios for further analysis like Price/Earning (P/E ratio), Price /Book value, Debt/Equity, Return on Equity, Current Ratio and Net Profit Margin.

The Price/Earnings Ratio or P/E ratio is a ratio used for valuing a company. The value is measured by its current share price, relative to its profit per share earnings. The Price/Earnings Ratio can be calculated as its share value in the market/ earning per share. It comes with a few limitations like the one when comparing the P/E ratios of different companies, another which lies in the calculation of the formula.

$$\frac{P}{B}Ratio = \frac{Stock\ Price}{Total\ Assets - Intangible\ Assets\ and\ Liabilities}$$

When a company's stock's market value is compared with its Price to book ratio it is known as Price to Book ratio. The division of its current closing price and the book value per share (of latest quarter) gives the P/B ratio.
To measure the financial leverage of a company one can use the debt ratio. The Debt

$$Debt - Equity\ Ratio = \frac{Total\ Liabilities}{Shareholders' Equity}$$

Equity ratio is the division of the total liabilities of a company and equity of its stock holders. It shows the debt, the company uses to finance its assets relative to shareholder's equity.

$$Return\ o\ \ Euity = \frac{Net\ Income}{Shareholders' Equity}$$

The return on equity measures the company’s profit by conceding the amount of profit it generates with the investment of the shareholders.

$$Current\ Ratio = \frac{Current\ Assets}{Current\ Liabilities}$$

The current ratio which is also known as the liquidity ratio is a measure of the company’s capability to pay long term and short term obligations. This capability is measured by considering company’s total assets relative to its liabilities.

$$Net\ Profit\ Margin\ = \frac{Total\ Revenue - Total\ Expences}{STotal\ Revenue}$$

Or,

$$Net\ Profit\ Margin\ = \frac{Net\ Profit}{Total\ Revenue}$$

Net profit margin is the revenue which remains after deducting all the interest, expenses, tax and dividends from the total revenue.

Technical Analysis

The technical analysis, on the other hand, is a research on the stock prices in the stock market with the intent of making profits and/or investment decisions (YingziZhu et al., 2009). The technical analysis, when applied to the stock market, predicts the direction of the future stock prices based on their historic data. With a close examination of the previous price movements of a stock, the investor can predict the future price movements of that particular stock. But again, this forecasting too may not be 100% accurate but just like the weather forecasting; it gives the investors an overall picture of what likely is supposed to happen to the price of the stock. Technical analysis has a solid influence on the investor's decision if he can or cannot, safely bet on the stock. The stock can be bought when it's running low and can be sold when it is at its peak. The technical analysis makes use of the price charts, studies the pricing patterns and makes use of certain formulae to bring in the future of the stock price. This technique is mainly useful for the investors who look for some short term investments (also known as short term trading). This is readily applicable to the stocks whose prices are affected by the pressure of supply and demand. The price being considered here could be
low, high, open or the closing price of that particular stock. The time frame to be considered can be intraday, daily, weekly, monthly or yearly. In case of intraday, it can be in every 10minutes, half hourly or hourly.Charles Dow served as one of the initial editors for the Wall Street Journal and his ideas formedthe basis for technical analysis. His theory named "Dow Theory" laid the basic foundation for the technical analysis. The main principles put forth by Dow through his theory were:

Market Price Discounts Everything

Technical analysts take into consideration only the current price of the stock while ignoring the fundamental analysis of the company. They believe that the current price puts light on all the other information required, right from the

summed up knowledge of the investors, analysts, traders, technical and fundamental analysts to various other people who could affect the company in some way. They nevertheless believe that all those factors which are fundamental, psychological and economic, are already included in the price and hence there isn't any need to analyse the company's fundamental information. This leads to the analysis of the price developments of a particular stock in the stock market, which can now be seen as a result of the supply and demand.

Prices Movements are not Totally Random/Price move in trends

The technical analysts believe that the price developments of a stock follow a particular trend. They say, if the stock price of a stock falls, it will continue falling and if the stock price of a stock rises, it'll continue to do so. This indicates that after the establishment of a major trend, the future trend is most likely to follow a similar direction. If the price is rising, the market is said to be "bullish" or "bearish", otherwise.

Historic trends usually repeat in the same patterns

History repeats, here in the context of the technical analysis of the stock market. Every time the price of a stock moves in a particular direction, the investors react consistently in the same manner to the price movements of the stock. The analysts of technical analysis use price charts and the technical indicators for future prediction. These form the mathematical derivatives of the volumes and prices of the stock. The analysts believe that the investors recursively go back to the past behavior in the same environment which forms the patterns on the chart, the technical analysts exploit these patterns and trends to foresee the direction of the stock market in future.

Tony Plummer paraphrased Oscar Wilde in his book titled, The Psychology of Technical Analysis, stating that, a technical analyst knows the price of everything but the value of nothing. The following two factors are of utmost

concern to the technical analysts: the present price,the price movements in the past. Of course, the main war is between the demand and supply of the stock of the company and the price is the result of that battle, which is also termed as the end product. The whole objective of the technical analysis is forecasting the future direction of the stock price of the company. Technical analysts focus only on what the price is, in contrast to the fundamental analysts who are concerned with the prices going up and down. As mentioned above, the technical analysis is based entirely on price and volume. The fields, which are of prime concern, are:

Open: This is the price at which the stock trades immediately after the opening of an exchange on a given day. An opening price of the stock is a very important indicator of trading activity for short-term traders such as day traders. It is considered to be the consensus price, after it is possible for the traders to “sleep on that”.

High: A high is the highest price that the stock traded, during the day. At this point, there are
fewer buyers and more sellers.

Low: A low is the lowest price that the stock traded during the day. At this point, there are
fewer sellers and more buyers.

Close: Close refers to the final price of particular trading stock for that particular day. This is the price used most often by the analysts. Most of the technical analysts use the relationship between the closing and opening price of a stock for its analysis and this kind of relationship is used with candlestick charts for analysis.

Volume: The volume could be the total number of shares traded during a day. The relation between the volume and prices often forms an important factor for analysis.

Technicians apply a variety of methods, involving various techniques and tools, one among which is using charts. The use of charts allows the technical analysts to identify the pricing patterns as well as the market trends of the financial markets and helps them try to exploit such patterns. Technicians make use of charts in order to search for certain archetypal price patterns, an example to which would be the well-known head and shoulders or double bottom/top reversal patterns, moving averages, study indicators etc.

Technical analysts use the various available market indicators like the mathematical transformation of the price which includes the up and down volume, decline/advance data etc. Such indicators help to qualify if an asset is trending, and if it is the probable direction of its continuation. The technical analysts also observe the relation, the price/volume indices and the market indicators have. For example, the moving average, the relative strength index and Moving Average Convergence and Divergence (MACD) etc. Following any one of the many available techniques which can do the technical analysis, adherence of various techniques such as candlestick charting, Elliott Wave theory or Dow Theory, may lead to ignorance of other approaches. Mostly, the traders trust the combination of more than one such technique for their analysis. While some other technical analysts take help of subjective judgment to decide which pattern, a particular instrument reflects at a given time and how to interpret that pattern. Some other technical analysts employ a carefully planned, systematic or a strictly mechanical approach in order to identify the pattern and use it for further interpretation.

- **Sentiment Analysis**

Earlier, when the internet did not come into existence, people used to take advice from experts and go through newspapers to monitor the stocks they want to invest in. But in the most recent times they use internet for these activities. It has been observed that there is extremely large growth in the

internet users. This is because of the availability of new devices, technologies which are affordable to most of the users.

As mentioned above, with the advent of internet and online trading in India, it has been observed that, there is an exponential increase in user generated data on the internet. The people started using various blogs, websites to express their opinions and emotions on various subjects, be it political, brands or movies etc. This huge amount of data on the Internet provided the researchers a chance to analyse this data to find patterns, interesting observations and other knowledge from the same. It was found that this knowledge could be used to take some decisions.

As a consequence of exponential growth of on-online trading, huge amounts of data pertaining to stocks are available. Stock related data that is available online can be thought of as two types: numerical data in the form of historical prices and textual information which contains the news articles available in the blogs or websites of news media. Stock prediction, in the earlier work, was based on historical data. Most of the researchers used fundamental analysis and technical predictors, based on the historical data, to predict the fall or raise in the future stock price. In the recent times, the online content in the form of news articles has become an important role player in the stock market prediction.

There’s always a quest in the investors regarding the news items and expert advices, so that they can predict the direction of a particular stock price, they are interested in.

This task needs analysis of information which is in the form of text. But manually going through the huge amount of data became a daunting task. The solution did lie in automatically digesting the gist of the news. The text analysis in the form of sentiment analysis which automatically extracts the gist i.e. whether the news article is positive, negative or neutral was required. The techniques used for text analysis can be classified into three main categories:

Bag of Words Approach, List of keywords approach and sentiment analysis. The bag of words is used in information retrieval and natural language processing. In the bag of words model, the news article is represented by a vector which consists of the weights of the words, the grammar and the order are not considered. Generally, it is used in document classification methods in which the classifier is trained by using the word count as a feature(Fung et al, 2002; Schumaker et al., 2009). In the second approach i.e., list of keywords approach, the researchers have a list of keywords or terms related to the news articles.(Satoru Takahashi,2006) analysed analysts' reports by using keyword information and confirmed that it plays an important role in predicting future direction of stock earnings. Sentiment Analysis is the third method which is used by numerous researchers. There is always an emotion, called sentiment, attached with these news items. For instance, an expert in a financial blog may predict that the stock price of a particular company goes up. Here the sentiment is considered as positive. (Liu, 2010) gaveavery convincing definition of sentiment analysis. It defines a model in which the unstructured texts are discussed as structured data. The model is a quintuple consisting of (oj, fjk, ooijkl, hi, tj). This means, for an object oj, we have a set of features fjk. The opinion orientations for this structure are ooijkl and these can be tied to the root object or to specific features. These opinions can also have a specific orientation and strength (joy, anger, positive, sad, negative, etc.). A few nonsubjective opinions can also be added as part of factual data if the structure is intended to completion (neutral opinion orientations). Hence, the model will contain a combination of factual data as well as opinions referred to an object along with its features. The opinion holder hi can be considered if a person or organisation claims this opinion at a particular time tj.

According to Efficient Markets Hypothesis (EMH) (Malkiel, Burton G, 2003), the stock price of a company reflects the whole information and that the new information causes a change in the stock price. The researchers of the stock market then started applying the sentiment analysis on the stock market data. The word sentiment showed its existence as the opinion of an investor(DeLong et al., 1990).As the research in financial market field progressed, it has shown

its evolution. The sentiment which was the response to the news article about the company, corporate announcements etc., was attributed to the stock price movements.

Sentiment analysis is a research area where the sentiment of a news article, or a review is defined as the opinion or sentiment expressed in that news article or the review(Turney,2002). Sentiment analysis is also referred to as opinion analysis (Lee et al,2008). The terminology, history and information regarding these terms were in their work. They concluded that both terms are similar. We use the term sentiment analysis through this thesis. For analysing sentiment of a news article there exists three methods(Lee et al.,2008), Machine learning, linguistic and lexicon methods(Liu, B.,2012) showed that the sentiment can be detected at three levels namely document, sentence and entity. As the name suggests the document level sentiment is the sentiment of overall document and sentence level sentiment is the sentiment expressed in each sentiment. We consider the news article as a document and the sentiment is the overall sentiment of the whole news article.

2.2 PROPOSED SYSTEM:

The existing systems have a huge overhead of extra manual ETL (Extraction, Transformation, Loading) job to do and then selectively run a single type of algorithm as to classify the dataset and to possibly gain insights out of it. Not to mention dealing with all kinds of minor data shift errors while working on the raw data list.

Our system address these problems in a unique approach i.e., we have created two different systems meant to deal with specific day stock price predictor which is based on Support Vector Machines (SVM) & Linear Regression (LR) and another meant to do prediction for a given period which is being achieved by creating Neural Networks using TensorFlow's Artificial Neural Network (TFANN).

2.2.1 SUPPORT VECTOR MACHINES (SVM)

Many researchers of sentiment analysis found Support Vector Machines to be very effective. When compared to Naïve Bayes Classifier in text categorization, SVM outperformed NBC (Joachims, 1998).In contrast to NBC and Maximum Entropy, SVM is large margin classifier, rather than probabilistic classifier. In case of two class label, the procedure for training is finding a hyperplane that is represented by the vector$\vec{\ }$,which separates, not only document classes of the two classes, but for the margin, as large as possible. The corresponding constrained optimization search problem is:

Let cjbe the class and cj∈ {1,-1} (1 for positive and -1 for negative).

Let djbe the document it belongs to the class cj.

The solution for the above search problem can be put as:

$\vec{\ }$: $-\sum\alpha jcj\ j,\ \alpha j>0$

where we can obtain the αj’s from the solutions of dual optimization problem.The documents j whose αj values are greater than 0 are nothing but support vectors.Theseare the document vectors that only contribute to the

vector αj. The classification of test instances could be made by considering the side of , hyperplane they fall on.

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well (look at the below snapshot).

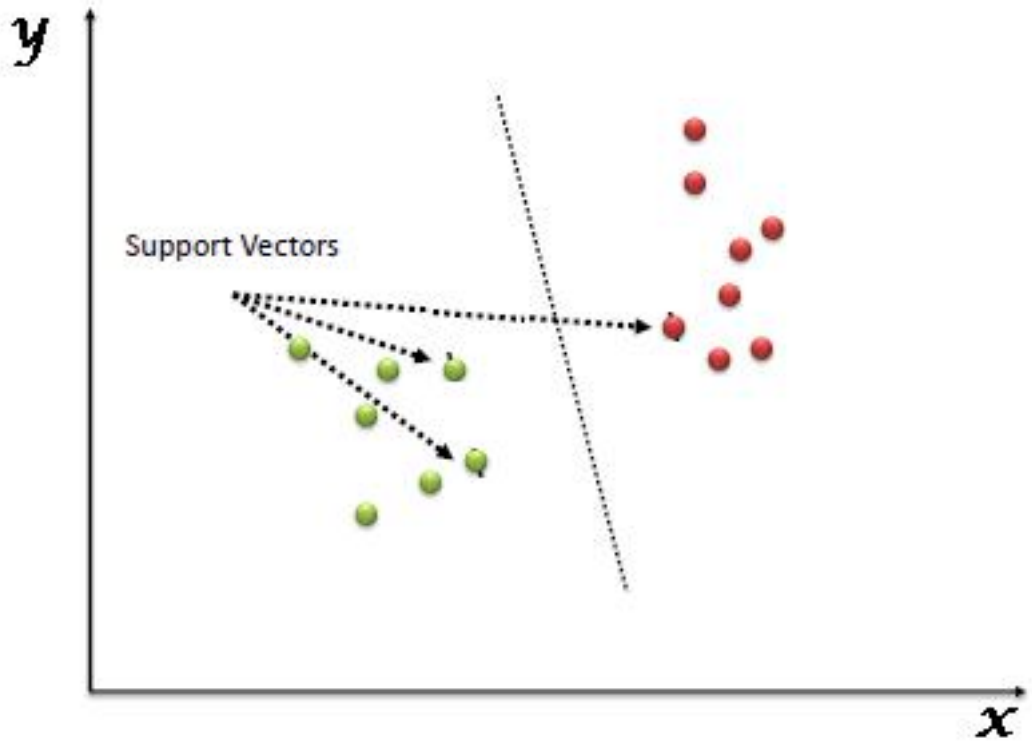


Fig 2.2 Support Vector Machine x-y plane

Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes (hyper-plane/ line).

2.2.2 LINEAR REGRESSION

In statistics, linear regression is a linear approach to modelling 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. For more than one explanatory variable, the process is called multiple linear regression. This term is distinct from multivariate linear regression, where multiple correlated dependent variables are predicted, rather than a single scalar variable.

In linear regression, the relationships are modelled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models. Most commonly, the conditional mean of the response given the values of the explanatory variables (or predictors) is assumed to be an affine function of those values; less commonly, the conditional median or some other quantile is used. Like all forms of regression analysis, linear regression focuses on the conditional probability distribution of the response given the values of the predictors, rather than on the joint probability distribution of all of these variables, which is the domain of multivariate analysis.

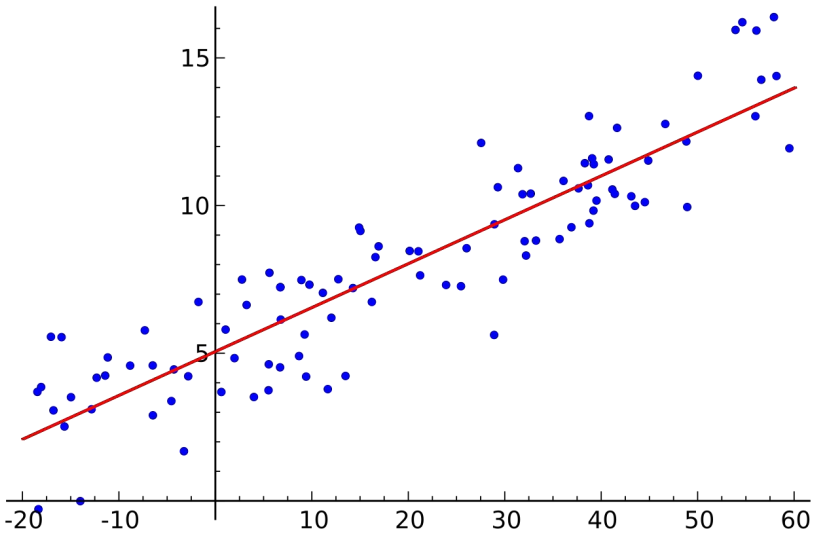


Fig 2.3 Linear Regression example

2.2.3 NEURAL NETWORKS

A neural network is a series of algorithms that endeavours to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading systems.

A neural network (NN), in the case of artificial neurons called artificial neural network (ANN) or simulated neural network (SNN), is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing based on a connectionistic approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network.

In more practical terms neural networks are non-linear statistical data modelling or decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

An artificial neural network involves a network of simple processing elements (artificial neurons) which can exhibit complex global behaviour, determined by the connections between the processing elements and element parameters.

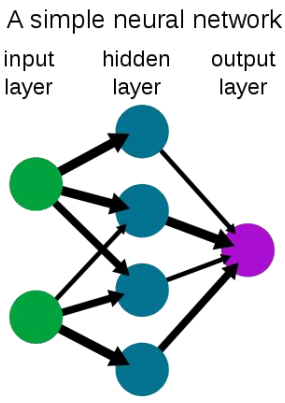


Fig. 2.4 Simple Neural Network

CHAPTER 3 : TOOLS AND TECHNOLOGIES

CHAPTER 3
TOOLS AND TECHNOLOGIES

3.1 HARDWARE REQUIREMENTS

Hardware requirements for implementing our specified system:

- Ubuntu machine
  - RAM: 16 GB
  - SSD: 120 GB (minimum)
  - GPU: Nvidia GTX 10 series or above
  - TPU: Provided by Google Colab
- High speed internet for best performance
- Functional peripheral I/O devices

3.2 SOFTWARE REQUIREMENTS

Software requirements for implementing our specified system:

- Packages:
  - SciKit
  - Pandas
  - NumPy
  - Matplot_lib
  - TensorFlow (compat v1)

- Google Chrome or Chromium browser (version 65 or above)

Services:

- **GOOGLE COLAB**

  - Colab is ideal for everything from improving your Python coding skills to working with deep learning libraries, like PyTorch, Keras, TensorFlow, and OpenCV.
  - You can create notebooks in Colab, upload notebooks, store notebooks, share notebooks, mount your Google Drive and use whatever you've got stored in there, import most of your favorite directories, upload your personal Jupyter Notebooks, upload notebooks directly from GitHub, upload Kaggle files, download your notebooks, and do just about everything else that you might want to be able to do.

- **YAHOO FINANCE**

  - Yahoo! Finance is a media property that is part of Yahoo!'s network. It provides financial news, data and commentary including stock quotes, press releases, financial reports, and original content. It also offers some online tools for personal finance management. In addition to posting partner content from a wide range of other web sites, it posts original stories by its team of staff journalists.

CHAPTER 4 : IMPLEMENTATION STEPS

CHAPTER 4
IMPLEMENTATION STEPS

4.1 COLLECTING HISTORICAL STOCK DATA

- Open up Google Chrome browser and go to website
  - https://in.finance.yahoo.com/

- Then search for "AAPL" or "Apple Incorporation"

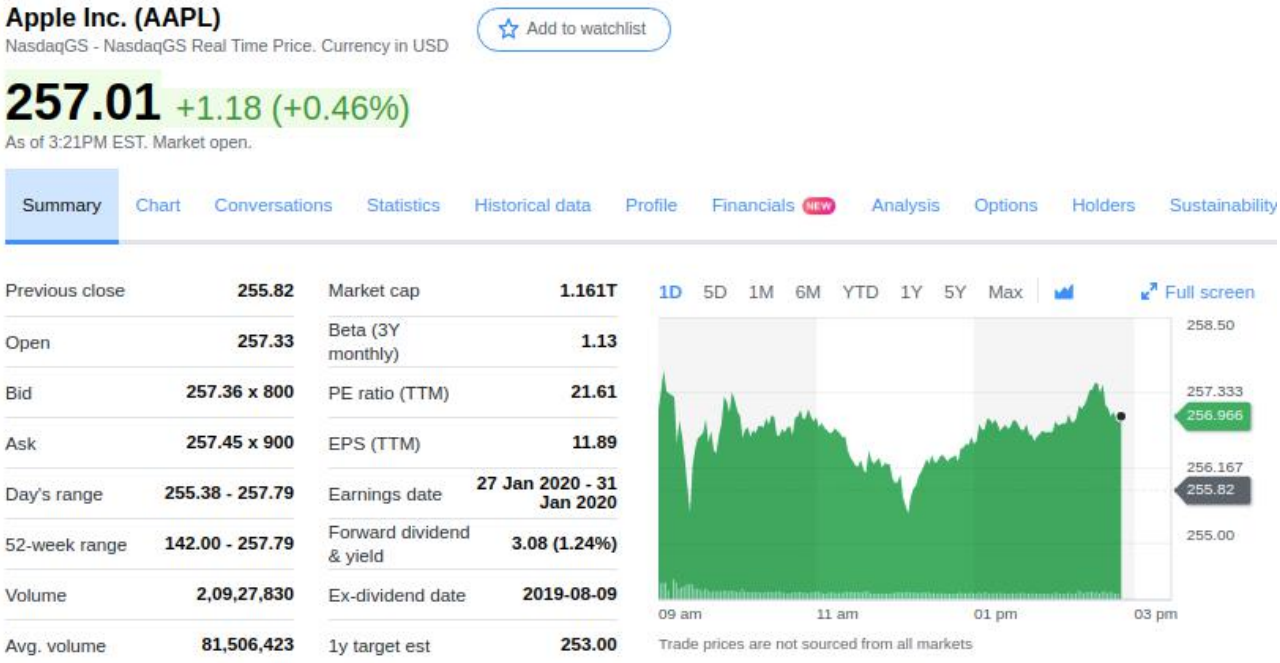


Fig 4.1 Yahoo Finance page displaying AAPL stock price in USD

- Then click on **Historical data.**

- Then we are needed to do 2 different stock dataset csv file download
  - First set : Time period = 1 month and download the data in .csv format.

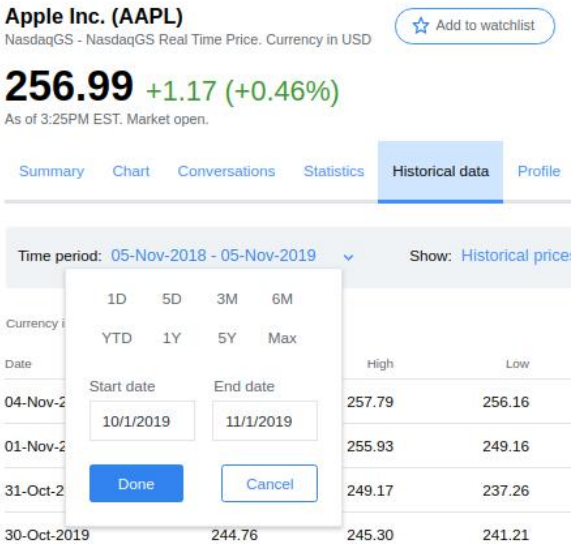


Fig 4.2 Selecting 1 month tenure of stock data for download

  - Second set : Time period = 5 years or more & download the data in .csv format.

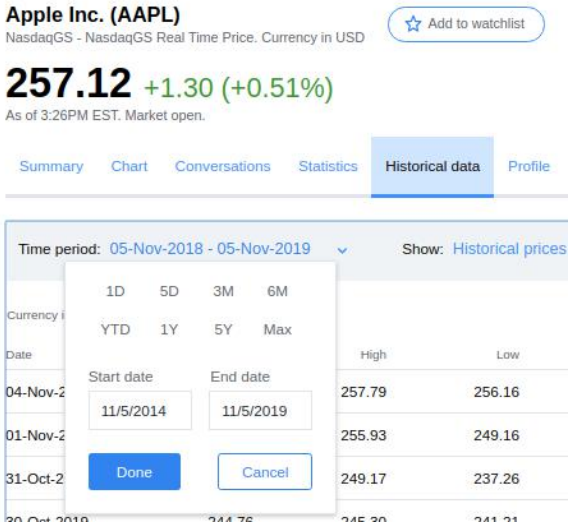


Fig 4.3 Selecting 5 years tenure of stock data for download

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Adj Clo	Volume
2	2019-09-30	220.9	224.58	220.79	223.97	223.97	3E+07
3	2019-10-01	225.07	228.22	224.2	224.59	224.59	3E+07
4	2019-10-02	223.06	223.58	217.93	218.96	218.96	3E+07
5	2019-10-03	218.43	220.96	215.13	220.82	220.82	3E+07
6	2019-10-04	225.64	227.49	223.89	227.01	227.01	3E+07
7	2019-10-07	226.27	229.93	225.84	227.06	227.06	3E+07
8	2019-10-08	225.82	228.06	224.33	224.4	224.4	3E+07
9	2019-10-09	227.03	227.79	225.64	227.03	227.03	2E+07
10	2019-10-10	227.93	230.44	227.3	230.09	230.09	3E+07
11	2019-10-11	232.95	237.64	232.31	236.21	236.21	4E+07
12	2019-10-14	234.9	238.13	234.67	235.87	235.87	2E+07
13	2019-10-15	236.39	237.65	234.88	235.32	235.32	2E+07
14	2019-10-16	233.37	235.24	233.2	234.37	234.37	2E+07
15	2019-10-17	235.09	236.15	233.52	235.28	235.28	2E+07
16	2019-10-18	234.59	237.58	234.29	236.41	236.41	2E+07
17	2019-10-21	237.52	240.99	237.32	240.51	240.51	2E+07
18	2019-10-22	241.16	242.2	239.62	239.96	239.96	2E+07
19	2019-10-23	242.1	243.24	241.22	243.18	243.18	2E+07
20	2019-10-24	244.51	244.8	241.81	243.58	243.58	2E+07
21	2019-10-25	243.16	246.73	242.88	246.58	246.58	2E+07
22	2019-10-28	247.42	249.25	246.72	249.05	249.05	2E+07
23	2019-10-29	248.97	249.75	242.57	243.29	243.29	4E+07
24	2019-10-30	244.76	245.3	241.21	243.26	243.26	3E+07
25							

Fig 4.4 AAPL stock data of a month in csv format

4.2 MAKING DATA MULTI REGION AVAILABLE

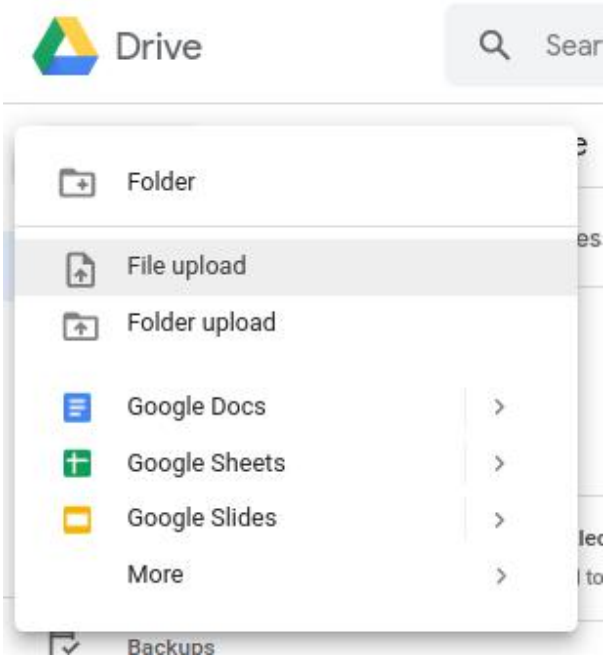


Fig 4.5 Uploading your downloaded stock data to Google Drive

- Now after getting the stockaapl.csv data, we are going to upload it to Google Drive for multi-region availability.
- Which basically allows us to fetch the .csv file from anywhere and doesn't constraint us to local storage only.
- Thus, we are required to upload "stockaapl.csv" to Google drive storage.

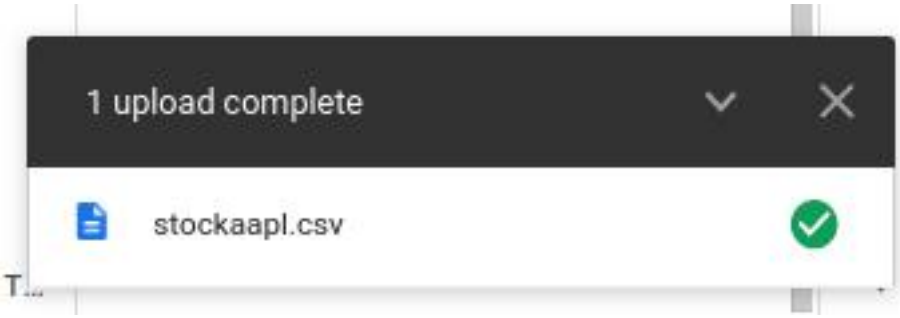


Fig 4.6 "stockaapl.csv" file uploaded

- Create new Python3 notebook

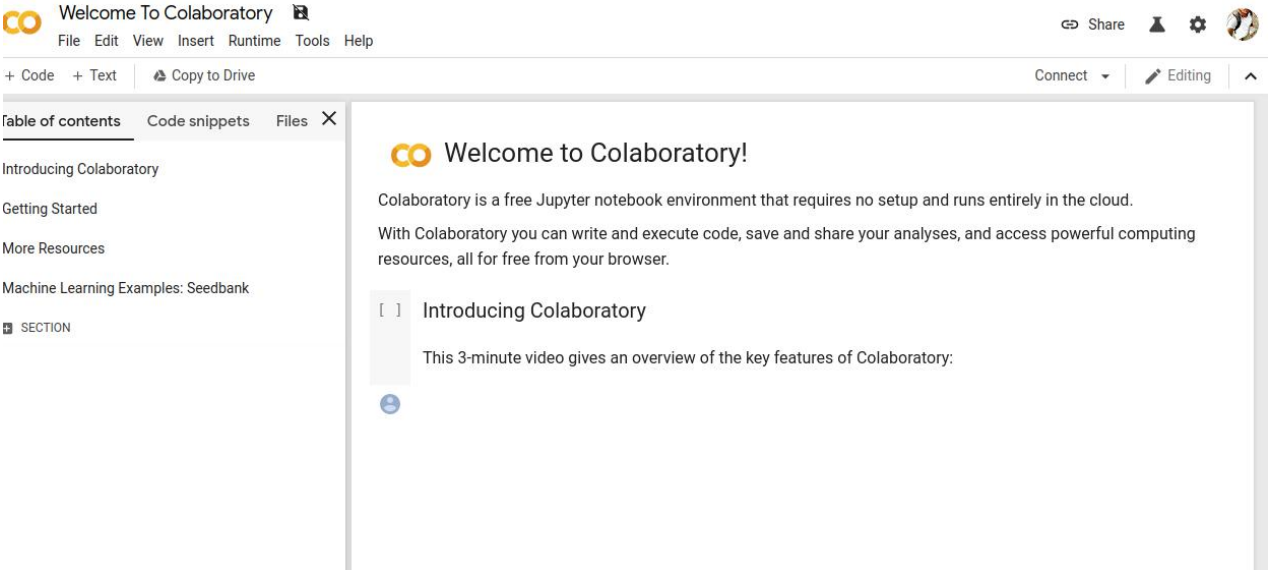


Fig 4.7 Google Colaboratory Homepage

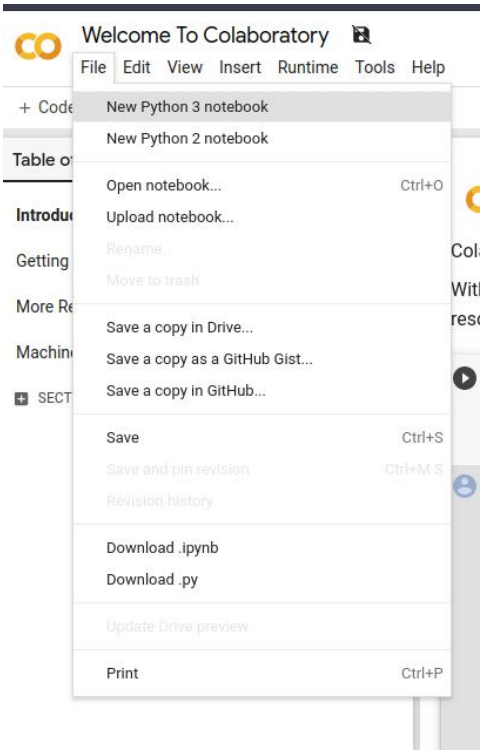


Fig 4.8 Create new notebook

4.3 WORKING ON GOOGLE COLABORATORY

- Mount the data from Google Drive (your aaplstocks.csv are available in it)

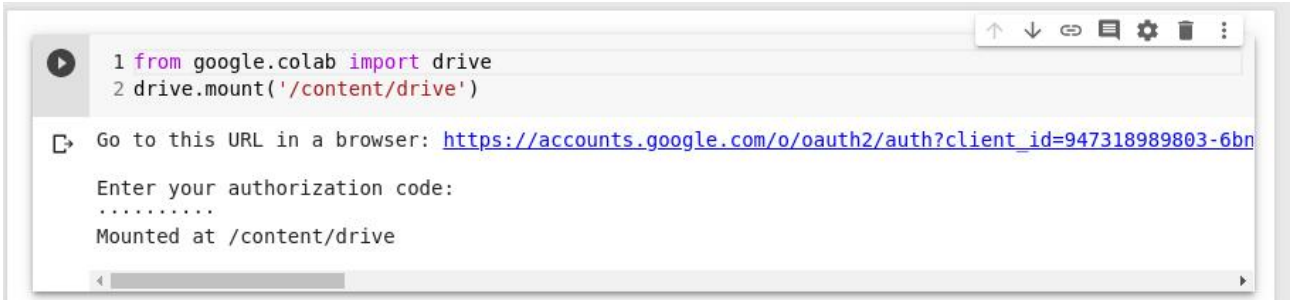
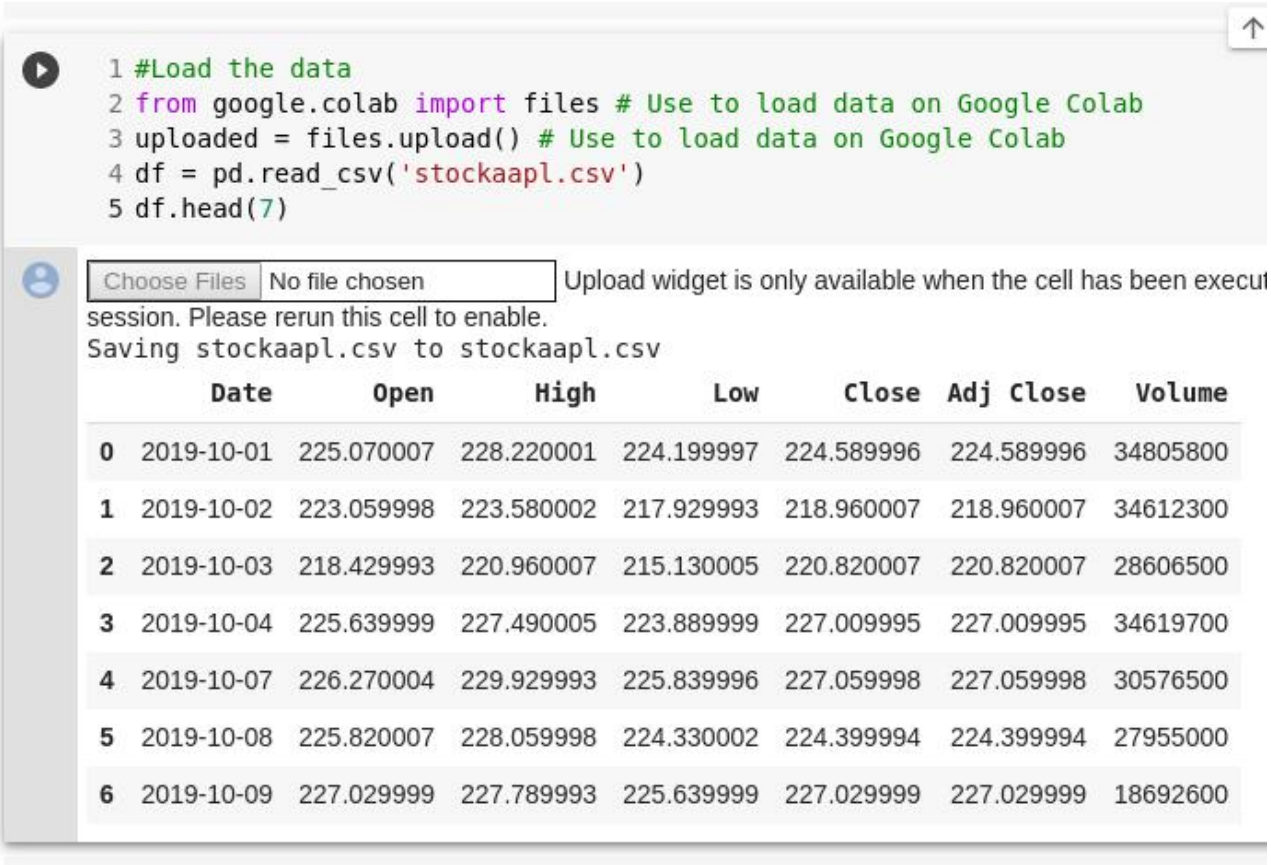


Fig 4.9 Data mounting

- Alternatively, you can upload your file locally too.



	Date	Open	High	Low	Close	Adj Close	Volume
0	2019-10-01	225.070007	228.220001	224.199997	224.589996	224.589996	34805800
1	2019-10-02	223.059998	223.580002	217.929993	218.960007	218.960007	34612300
2	2019-10-03	218.429993	220.960007	215.130005	220.820007	220.820007	28606500
3	2019-10-04	225.639999	227.490005	223.889999	227.009995	227.009995	34619700
4	2019-10-07	226.270004	229.929993	225.839996	227.059998	227.059998	30576500
5	2019-10-08	225.820007	228.059998	224.330002	224.399994	224.399994	27955000
6	2019-10-09	227.029999	227.789993	225.639999	227.029999	227.029999	18692600

4.4 USING SUPPORT VECTOR MACHINES & LINEAR REGRESSION

```
#Import the libraries
import pandas as pd
import numpy as np
from sklearn.svm import SVR
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt

#Load the data
from google.colab import files # Use to load data on Google Colab
uploaded = files.upload() # Use to load data on Google Colab
df = pd.read_csv('stockaapl.csv')
df.head(7)

#Create the lists / X and Y data sets
dates = []
prices = []

#Get the number of rows and columns in the data set
df.shape

#Get the last row of data (this will be the data that we test on)
df.tail(1)

#Get all of the data except for the last row
df = df.head(len(df)-1)
df

#The new shape of the data
df.shape

#Get all of the rows from the Date Column
df_dates = df.loc[:, 'Date']
#Get all of the rows from the Open Column
df_open = df.loc[:, 'Open']
```

```
#See what days were recorded
print(dates)

def predict_prices(dates, prices, x):

  #Create the 3 Support Vector Regression models
  svr_lin = SVR(kernel='linear', C= 1e3)
  svr_poly= SVR(kernel='poly', C=1e3, degree=2)
  svr_rbf = SVR(kernel='rbf', C=1e3, gamma=0.1)

  #Train the SVR models
  svr_lin.fit(dates,prices)
  svr_poly.fit(dates,prices)
  svr_rbf.fit(dates,prices)

  #Create the Linear Regression model
  lin_reg = LinearRegression()
  #Train the Linear Regression model
  lin_reg.fit(dates,prices)

  #Plot the models on a graph to see which has the best fit
  plt.scatter(dates, prices, color='black', label='Data')
  plt.plot(dates, svr_rbf.predict(dates), color='red', label='SVR RBF')
  plt.plot(dates, svr_poly.predict(dates), color='blue', label='SVR Poly')
  plt.plot(dates, svr_lin.predict(dates), color='green', label='SVR Linear')
  plt.plot(dates, lin_reg.predict(dates), color='orange', label='Linear Reg')
  plt.xlabel('Days')
  plt.ylabel('Price')
  plt.title('Regression')
  plt.legend()
  plt.show()
```

4.5 USING NEURAL NETWORKS

```
!pip install TFANN
import numpy as np
import matplotlib.pyplot as mpl
import tensorflow.compat.v1 as tf
tf.disable_v2_behavior()
from sklearn.preprocessing import scale
from TFANN import ANNR
from google.colab import files


files.upload()
!ls


#reads data from the file and creates a matrix with only dates and the prices
stock_data = np.loadtxt('aapl_5y.csv',delimiter=",",skiprows=1,usecols=(1, 4))
#scales the data to smaller values
stock_data = scale(stock_data)
#gets the price and dates from the matrix
prices = stock_data[:, 1].reshape(-1, 1)
dates = stock_data[:, 0].reshape(-1, 1)
#creates a plot of the data and then displays it
mpl.plot(dates[:, 0], prices[:, 0])
mpl.show()


#Number of neurons in the input, output, and the hidden layers
input = 1
output = 1
hidden = 50
#array of layers, 3 hidden and 1 output, along with the tanh activation function
layers = [('F', hidden), ('AF', 'tanh'), ('F', hidden), ('AF', 'tanh'), ('F', hidden), ('AF', 'tanh'), ('F',
output)]
#construct the model and dictate params
mlpr = ANNR([input], layers, batchSize = 256, maxIter = 20000, tol = 0.2, reg = 1e-
4, verbose = True)


#number of days for the hold-out period used to access progress
holdDays = 5
totalDays = len(dates)
#fit the model to the data "Learning"
mlpr.fit(dates[0:(totalDays-holdDays)], prices[0:(totalDays-holdDays)])
```

```
#Predict the stock price using the model
pricePredict = mlpr.predict(dates)
#Display the predicted reuslts agains the actual data
mpl.plot(dates, prices)
mpl.plot(dates, pricePredict, c='#5aa9ab')
mpl.show()

#Number of neurons in the input, output, and hidden layers
input2 = 1
output2 = 1
hidden2 = 50
#array of layers, 3 hidden and 1 output, along with the tanh activation function
layers = [('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', out
put2)]
#construct the model and dictate params
mlpr2 = ANNR([input2], layers, batchSize = 256, maxIter = 10000, tol = 0.1, reg = 1e-4, verbose = True)

holdDays = 5
totalDays = len(dates)
mlpr2.fit(dates[0:(totalDays-holdDays)], prices[0:(totalDays-holdDays)])

#Number of neurons in the input, output, and hidden layers
input2 = 1
output2 = 1
hidden2 = 50
#array of layers, 3 hidden and 1 output, along with the tanh activation function
layers = [('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', out
put2)]
#construct the model and dictate params
mlpr2 = ANNR([input2], layers, batchSize = 256, maxIter = 10000, tol = 0.1, reg = 1e-4, verbose = True)

holdDays = 5
totalDays = len(dates)
mlpr2.fit(dates[0:(totalDays-holdDays)], prices[0:(totalDays-holdDays)])

pricePredict2 = mlpr2.predict(dates)
mpl.plot(dates, prices)
mpl.plot(dates, pricePredict, c='#5aa9ab')
mpl.plot(dates, pricePredict2, c='#8B008B')
mpl.show()
```

CHAPTER 6 : SNAPSHOTS

6.1 STOCK PREDICTION FOR SPECIFIC DAY

acnologia.ipynb ☆

File Edit View Insert Runtime Tools Help

+ Code + Text

```
#from google.colab import drive
#drive.mount('/content/drive')
```

```
1
```

```
# PROJECT ACNOLOGIA
```

```
# Description: This program predicts the price of AAPL stock for a specific day
#              using the Machine Learning algorithm called Support Vector Regression (SVR)
#              & Linear Regression.
```

```
#Import the libraries
import pandas as pd
import numpy as np
from sklearn.svm import SVR
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
```

```
#Load the data
from google.colab import files # Use to load data on Google Colab
uploaded = files.upload() # Use to load data on Google Colab
df = pd.read_csv('stockaapl.csv')
df.head(7)
```

Choose Files | No file chosen — Upload widget is only available when the cell has been executed in the current browse

Saving stockaapl.csv to stockaapl.csv

	Date	Open	High	Low	Close	Adj Close	Volume
0	2019-10-01	225.070007	228.220001	224.199997	224.589996	224.589996	34805800
1	2019-10-02	223.059998	223.580002	217.929993	218.960007	218.960007	34612300
2	2019-10-03	218.429993	220.960007	215.130005	220.820007	220.820007	28606500
3	2019-10-04	225.639999	227.490005	223.889999	227.009995	227.009995	34619700
4	2019-10-07	226.270004	229.929993	225.839996	227.059998	227.059998	30576500
5	2019-10-08	225.820007	228.059998	224.330002	224.399994	224.399994	27955000
6	2019-10-09	227.029999	227.789993	225.639999	227.029999	227.029999	18692600

```
#Create the lists / X and Y data sets
dates = []
prices = []
```

```
#Get the number of rows and columns in the data set
df.shape
```

```
(22, 7)
```

```
#Get the last row of data (this will be the data that we test on)
df.tail(1)
```

	Date	Open	High	Low	Close	Adj Close	Volume
21	2019-10-30	244.759995	245.300003	241.210007	243.259995	243.259995	30950600

```
#Get all of the data except for the last row
df = df.head(len(df)-1)
df
```

Fig 6.1 Snapshot #1

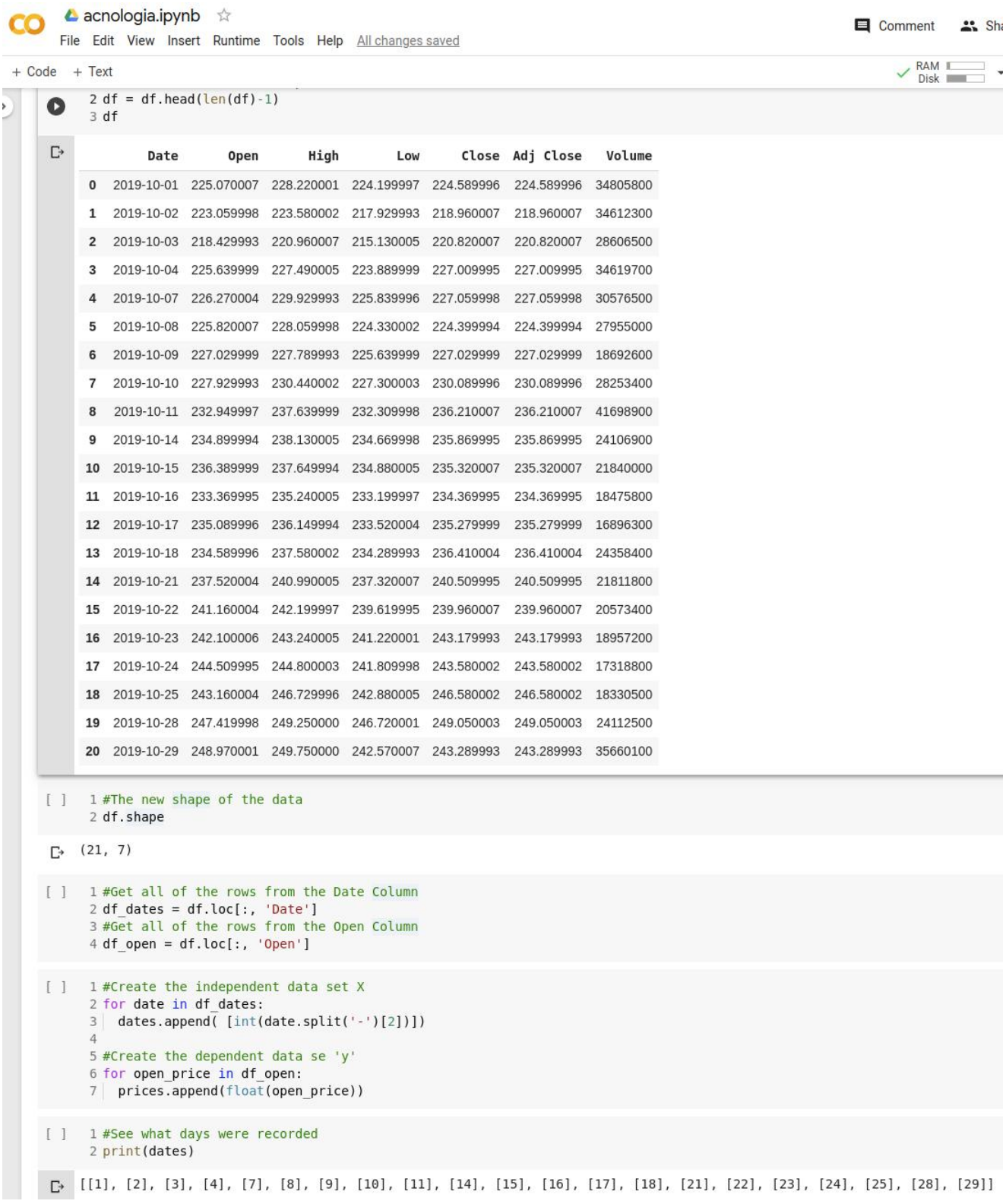


Fig 6.2 Snapshot #2

acnologia.ipynb ☆

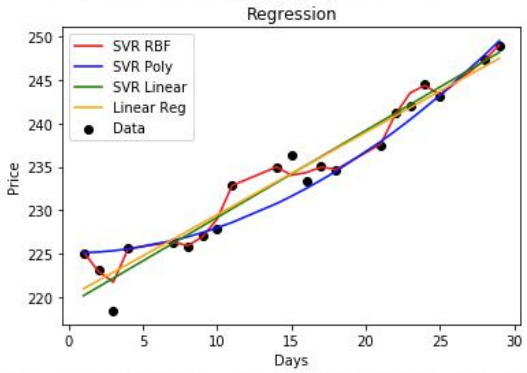
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[ ]  1 def predict_prices(dates, prices, x):
     2
     3   #Create the 3 Support Vector Regression models
     4   svr_lin = SVR(kernel='linear', C= 1e3)
     5   svr_poly= SVR(kernel='poly', C=1e3, degree=2)
     6   svr_rbf = SVR(kernel='rbf', C=1e3, gamma=0.1)
     7
     8   #Train the SVR models
     9   svr_lin.fit(dates,prices)
    10   svr_poly.fit(dates,prices)
    11   svr_rbf.fit(dates,prices)
    12
    13   #Create the Linear Regression model
    14   lin_reg = LinearRegression()
    15   #Train the Linear Regression model
    16   lin_reg.fit(dates,prices)
    17
    18   #Plot the models on a graph to see which has the best fit
    19   plt.scatter(dates, prices, color='black', label='Data')
    20   plt.plot(dates, svr_rbf.predict(dates), color='red', label='SVR RBF')
    21   plt.plot(dates, svr_poly.predict(dates), color='blue', label='SVR Poly')
    22   plt.plot(dates, svr_lin.predict(dates), color='green', label='SVR Linear')
    23   plt.plot(dates, lin_reg.predict(dates), color='orange', label='Linear Reg')
    24   plt.xlabel('Days')
    25   plt.ylabel('Price')
    26   plt.title('Regression')
    27   plt.legend()
    28   plt.show()
    29
    30   return svr_rbf.predict(x)[0], svr_lin.predict(x)[0],svr_poly.predict(x)[0],lin_reg.predict(x)[0]
```

```
1 #Predict the price of AAPL on day 30
2 predicted_price = predict_prices(dates, prices, [[30]])
3 print(predicted_price)
```

```
/usr/local/lib/python3.6/dist-packages/sklearn/svm/base.py:193: FutureWarning: The default value of ga
  "avoid this warning.", FutureWarning)
```




```
(246.2747129116121, 249.21313187487402, 251.26164839092428, 248.47301194121502)
```

Fig 6.3 Snapshot #3

6.2 STOCK PREDICTION FOR CERTAIN TIME PERIOD

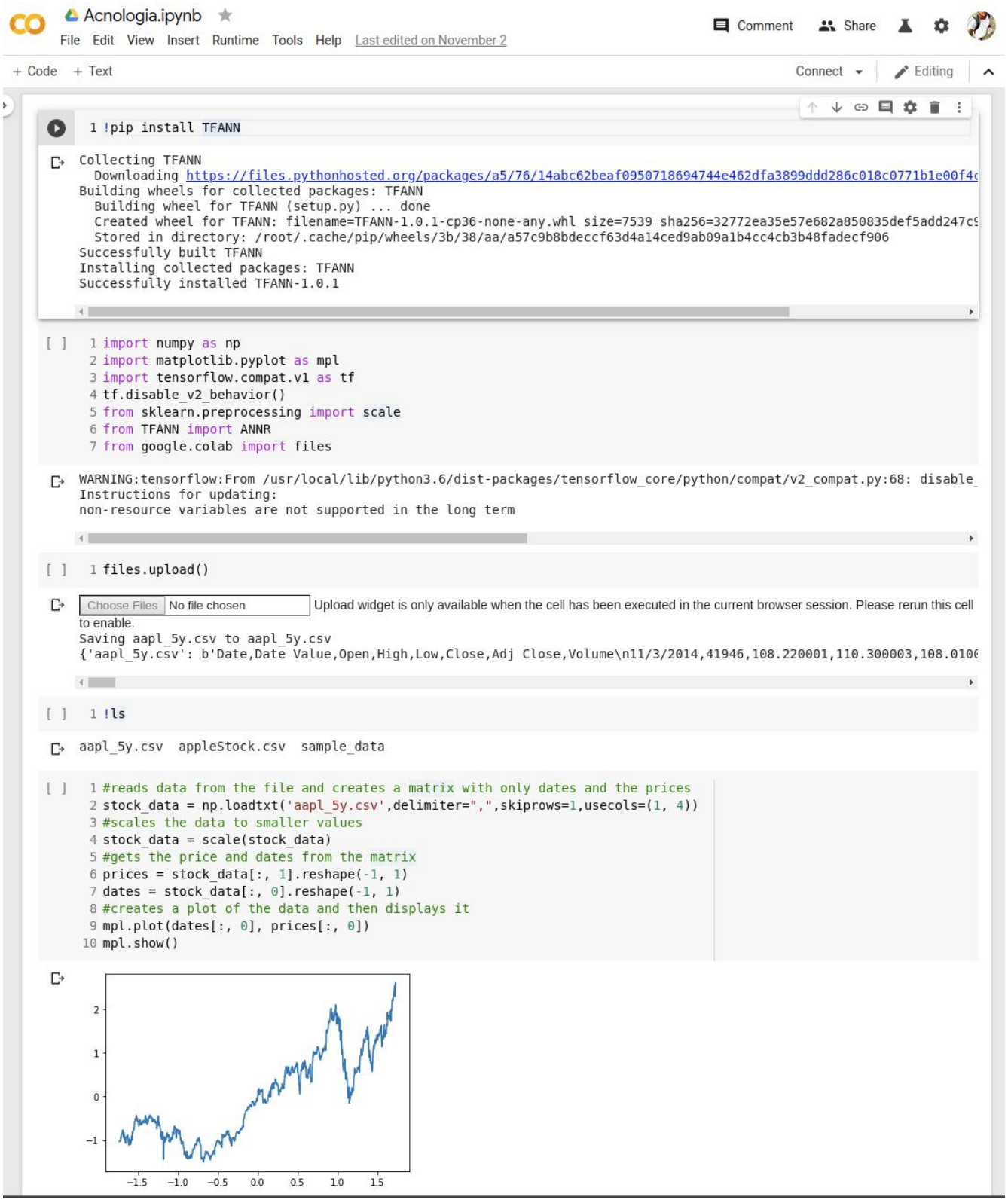


Fig 6.4 Snapshot #4

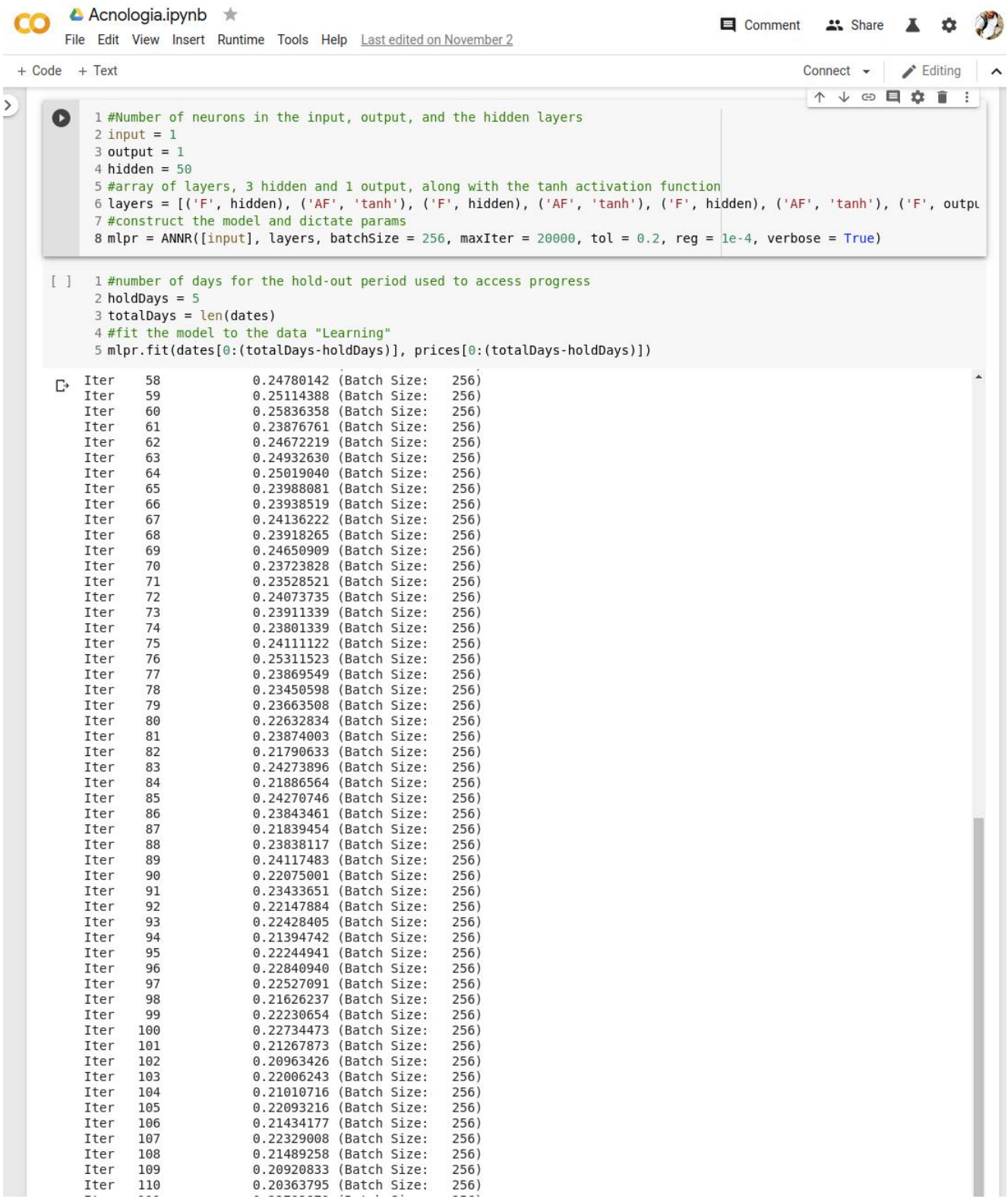


Fig 6.5 Snapshot #5

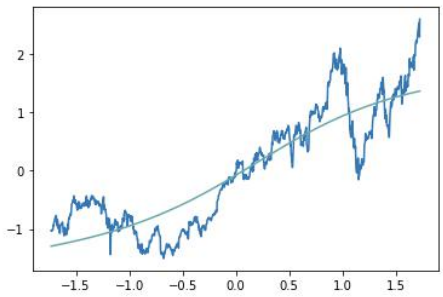
Acnologia.ipynb

File Edit View Insert Runtime Tools Help Last edited on November 2

Comment Share Connect Editing

+ Code + Text

```
[ ]  1 #Predict the stock price using the model
     2 pricePredict = mlpr.predict(dates)
     3 #Display the predicted reuslts agains the actual data
     4 mpl.plot(dates, prices)
     5 mpl.plot(dates, pricePredict, c='#5aa9ab')
     6 mpl.show()
```




```
[ ]  1 #Number of neurons in the input, output, and hidden layers
     2 input2 = 1
     3 output2 = 1
     4 hidden2 = 50
     5 #array of layers, 3 hidden and 1 output, along with the tanh activation function
     6 layers = [('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', hidden2), ('AF', 'tanh'), ('F', ou
     7 #construct the model and dictate params
     8 mlpr2 = ANNR([input2], layers, batchSize = 256, maxIter = 10000, tol = 0.1, reg = 1e-4, verbose = True)
```

```
[ ]  1 holdDays = 5
     2 totalDays = len(dates)
     3 mlpr2.fit(dates[0:(totalDays-holdDays)], prices[0:(totalDays-holdDays)])
```

```
Iter    22            0.28383180 (Batch Size:    256)
Iter    23            0.27258667 (Batch Size:    256)
Iter    24            0.26614072 (Batch Size:    256)
Iter    25            0.27006004 (Batch Size:    256)
Iter    26            0.28239530 (Batch Size:    256)
Iter    27            0.26800821 (Batch Size:    256)
Iter    28            0.26202746 (Batch Size:    256)
Iter    29            0.26596263 (Batch Size:    256)
Iter    30            0.26879123 (Batch Size:    256)
Iter    31            0.27277590 (Batch Size:    256)
Iter    32            0.26032876 (Batch Size:    256)
Iter    33            0.25818819 (Batch Size:    256)
Iter    34            0.25649613 (Batch Size:    256)
Iter    35            0.25652603 (Batch Size:    256)
Iter    36            0.26299198 (Batch Size:    256)
Iter    37            0.25814207 (Batch Size:    256)
Iter    38            0.26628540 (Batch Size:    256)
Iter    39            0.26742219 (Batch Size:    256)
Iter    40            0.26246611 (Batch Size:    256)
Iter    41            0.25620936 (Batch Size:    256)
Iter    42            0.26091596 (Batch Size:    256)
Iter    43            0.25554886 (Batch Size:    256)
Iter    44            0.25337067 (Batch Size:    256)
Iter    45            0.26105212 (Batch Size:    256)
Iter    46            0.24778356 (Batch Size:    256)
Iter    47            0.24981624 (Batch Size:    256)
Iter    48            0.26290720 (Batch Size:    256)
Iter    49            0.26303132 (Batch Size:    256)
Iter    50            0.25399984 (Batch Size:    256)
Iter    51            0.25369590 (Batch Size:    256)
Iter    52            0.25313923 (Batch Size:    256)
Iter    53            0.25025823 (Batch Size:    256)
```

Fig 6.6 Snapshot #6

Acnologia.ipynb

File Edit View Insert Runtime Tools Help Last edited on November 2

Comment Share

+ Code + Text

Connect Editing

```
Iter   541           0.11776957 (Batch Size:   256)
Iter   542           0.11043593 (Batch Size:   256)
Iter   543           0.10884264 (Batch Size:   256)
Iter   544           0.12405100 (Batch Size:   256)
Iter   545           0.10814308 (Batch Size:   256)
Iter   546           0.11574722 (Batch Size:   256)
Iter   547           0.13014690 (Batch Size:   256)
Iter   548           0.11850892 (Batch Size:   256)
Iter   549           0.11092670 (Batch Size:   256)
Iter   550           0.10897603 (Batch Size:   256)
Iter   551           0.11281530 (Batch Size:   256)
Iter   552           0.11198710 (Batch Size:   256)
Iter   553           0.11216939 (Batch Size:   256)
Iter   554           0.11686799 (Batch Size:   256)
Iter   555           0.10601985 (Batch Size:   256)
Iter   556           0.11940147 (Batch Size:   256)
Iter   557           0.11258409 (Batch Size:   256)
Iter   558           0.11725652 (Batch Size:   256)
Iter   559           0.11498643 (Batch Size:   256)
Iter   560           0.11678091 (Batch Size:   256)
Iter   561           0.11511273 (Batch Size:   256)
Iter   562           0.12368028 (Batch Size:   256)
Iter   563           0.11892106 (Batch Size:   256)
Iter   564           0.12451975 (Batch Size:   256)
Iter   565           0.11265973 (Batch Size:   256)
Iter   566           0.12253586 (Batch Size:   256)
Iter   567           0.11722316 (Batch Size:   256)
Iter   568           0.11071625 (Batch Size:   256)
Iter   569           0.11263711 (Batch Size:   256)
Iter   570           0.11526777 (Batch Size:   256)
Iter   571           0.10797244 (Batch Size:   256)
Iter   572           0.11996076 (Batch Size:   256)
Iter   573           0.11808547 (Batch Size:   256)
Iter   574           0.11290061 (Batch Size:   256)
Iter   575           0.11325453 (Batch Size:   256)
Iter   576           0.11569783 (Batch Size:   256)
Iter   577           0.09820368 (Batch Size:   256)
```

```
1 pricePredict2 = mlpr2.predict(dates)
2 mpl.plot(dates, prices)
3 mpl.plot(dates, pricePredict, c='#5aa9ab')
4 mpl.plot(dates, pricePredict2, c='#8B008B')
5 mpl.show()
```

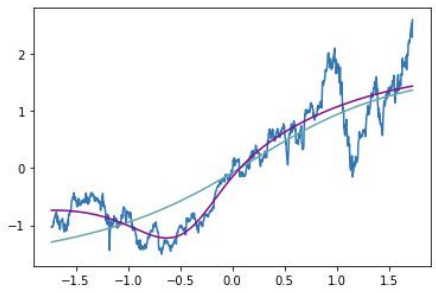


Fig 6.7 Snapshot #7

CHAPTER 6 : CONCLUSION

CHAPTER 6

CONCLUSION

This project was done with an aim to predict the future stock price(s) of any given company (here we have used Apple Inc. stock data). Even though our implementation focused on two different scenarios i.e., scenario 1 as predicting the stock price for a company on any given future date & scenario 2 as predicting the stock price for a company for a given tenure with optimal accuracy in it. Since Data Science is a huge domain of intercorrelated fields like stocks, housing, medicine prices and other financial dataset based analysis. All used algorithms in our project like Support Vector Machine(s), Linear Regression and even Neural Networks can be moulded and be implemented on other kind of datasets on accordance to the Data Scientist requirements.

At the end of the day what matters the most is the amount of valuable insights data scientist(s) can obtain in to provide required and better service to the companies with better deep insights on the company's internal & external factors and can give them a better business strategies to improve their product or service for better customer experience, fidelity, and ultimately helping the company to grow more and improve their yearly revenues and turn-overs. Also these insights plays a big role and helps the stock investors, angel investors and venture capitalists.

BIBLIOGRAPHY

WEBSITES :

- http://cs229.stanford.edu/
- https://ai.google/research/
- https://www.researchgate.net/
- http://www.ijircce.com/
- https://in.finance.yahoo.com/
- https://nseindia.com/
- https://www.analyticsvidhya.com/
- https://towardsdatascience.com/
- https://ieeexplore.ieee.org/
- https://acadpubl.eu/