**DATA MINING ASSIGNMENT – GROUP 4**

**STOCK PRICE PREDICTION**

**Introduction:**

The stock market broadly refers to the collection of exchanges and other venues where the buying, selling, and issuance of shares of publicly held companies take place. Such financial activities are conducted through institutionalized formal exchanges (whether physical or electronic) or via over the counter (OTC) marketplaces that operate under a defined set of regulations. A share price is the price of a single share of several saleable equity shares of a company.  
In layman's terms, the stock price is the highest amount someone is willing to pay for the stock, or the lowest amount that it can be bought for. Stock prices change every day by market forces. By this we mean that share prices change because of supply and demand. If more people want to buy a stock (demand) than sell it (supply), then the price moves up.

Originally incorporated as Bajaj Auto Finance Limited on March 25, 1987, as a non-banking financial company, primarily focused on providing two and three-wheeler finance. After 11 years in the auto finance market, Bajaj Auto Finance Ltd launched its initial public issue of equity share and was listed on the Bombay Stock Exchange and National Stock Exchange of India.

At the turn of the 20th century, the company ventured into the consumer durables finance sector and started offering small-size loans at zero interest rates. In the subsequent years, Bajaj Auto Finance diversified into business and property loans as well.

**Data Set:**

The data set used for the assignment is the change in stock price of the shares of Bajaj finance listed both on the national stock exchange and Bombay stock exchange. This three year daily stock price data was collected from the official Yahoo finance website.

The data consists of the following attributes:

* Date - consists of all the days from 30th of December 2018 to 30th of December 2021 the stock had traded on the exchanges
* Open - This column consists of the prices at which the stock opened
* High - This column displays the highest price the stock touched on the respective dates
* Low - This column displays the lowest price the stock touched on the respective dates
* Close - This column suggests the price at which the stock closed at the end of the day.
* Adj or adjusted close - The adjusted closing price amends a stock's closing price to reflect that stock's value after accounting for any corporate actions.
* Volume - Volume measures the number of shares traded in a stock or contracts traded in futures or options on the respective dates

**Body:**

Stock market prediction has become a skill that every investor wants to acquire. To have Machine learning and Predictive analysis back up that skill is an icing on the cake. Predictive analytics is driven by predictive modelling. It’s more of an approach than a process. Predictive analytics and [machine learning](https://www.sas.com/en_gb/insights/analytics/machine-learning.html) go hand-in-hand, as predictive models typically include a [machine learning](https://www.sas.com/en_gb/insights/analytics/machine-learning.html) algorithm. These models can be trained over time to respond to new data or values, delivering the results the business needs. Predictive modelling largely overlaps with the field of [machine learning](https://www.sas.com/en_gb/insights/analytics/machine-learning.html).

Different Machine Learning techniques can be used to predict daily stock prices. A few of them include:

* Regression
* Decision trees
* Neural networks

Objective - Predicting the future value of closing prices of the Stock using Decision Tree Regressor and Linear Regressor to find out which one is a better fit for continuous data.

Process followed: The dataset consists of 742 observations. As a part of the data pre-processing, the “Date” column was parsed and made the index column for easy access of the observations. The data has no missing values.

Graphical user interface, application

Description automatically generated

The target column taken is the “Close” column. Another variable “Predictions” is defined in which the closing price values are shifted up by the number of future days to accommodate the predicted values. The “Close” column is split into the training set(on which the model is trained) and testing set(on which the model is applied). Both the regression models are applied.

Table

Description automatically generated with low confidence

Chart, line chart

Description automatically generatedGraphical user interface, chart, line chart

Description automatically generated

The Coefficient of Determination of the both the models are found to see which model fits better for this data.

Graphical user interface, text, application

Description automatically generated

**Results:**The models predict the values of the Closing prices of the Bajaj Finance Stock for 20 days in the future. The accuracy of the predictions can be determined using the Coefficient of determination. The r2 score of the Linear regression model is around 99% , much higher that the r2 score of the Decision tree regressor, thus indicating that a Linear Regression model better predicts values for a continuous numerical data set.

**Conclusion:**

Predicting values based on statistical techniques has proven to be more accurate than baseless predictions. Investing is a very tricky field and speculating the prices to make a good buy and sell decision is crucial to every investor. Predictive analysis combined with machine learning helps in making decisions which have a solid backup. Regression can therefore be extremely beneficial to the following interest groups in relation to prediction of stock prices:   
  
**Fundamental Investors** – These investors judge a stock based on the company’s business. They look at their revenue and profits/losses for valuation purposes. If a company’s revenues and profits are growing, then their shares become more valuable.

**Technical Investors** – These investors judge a stock based on price action and chart patterns. Usually these are investors looking for short term trades and day trades. If a stock is having a hot week making new all-time highs, a day trader might jump in and ride the stock higher for a short period before selling and taking profits.

**Analysts** – Research and analyse certain stocks to predict future expected earnings and revenues. They issue analyst reports for their companies they work for like banks, hedge funds, mutual funds etc, to help management in deciding to purchase a stock or not.

**Online Investor** – someone who wants to trade from home and sets up an online brokerage account with anywhere from a few hundred dollars to many thousands of dollars. This is the self-managed investor who trusts himself or herself over an experienced financial adviser.

### Domestic AMCs - Asset Management Companies (AMCs) are essentially firms that pool funds from various clients and invest the capital raised in various financial market securities. Domestic AMCs are firms that are based out of India and include mutual fund houses such as ICICI Prudential Mutual Fund, HDFC Mutual Fund and the like.

### Retail (Indian and NRI and OCI) - Individuals like you who invest in the stock market are known as retail investors. It is however important to check the accuracy of the model that is going to be used in order to rely on the analysis.