**Artificial Intelligence-UCS411**

**LAB PROJECT SUBMISSION**

**Submitted To:** Ms Navpreet Kaur

**Submitted By:**

Deepanshi Sharma(102103376)

Kriti(102103388)

Ujjval Vashisht(102153029)

Kshitij Sharma(102103374)

**Problem statement:**

The aim of the project is to examine a number of different forecasting techniques to predict future stock returns based on past returns and numerical news indicators to construct a portfolio of multiple stocks in order to diversify the risk. We do this by applying supervised learning methods for stock price forecasting by interpreting the seemingly chaotic market data.

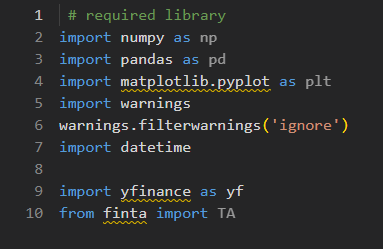
**Description of problem:**

The fluctuation of stock market is violent and there are many complicated financial indicators. However, the advancement in technology, provides an opportunity to gain steady fortune from stock market and also can help experts to find out the most informative indicators to make better prediction. The prediction of the market value is of paramount importance to help in maximizing the profit of stock option purchase while keeping the risk low.

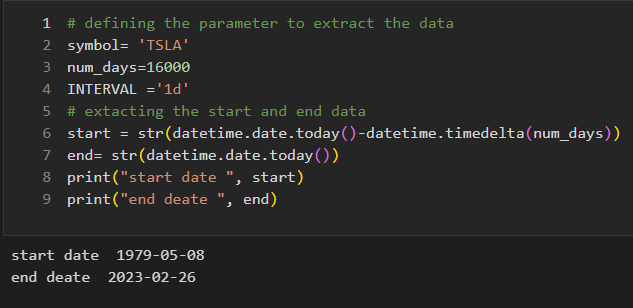
**Code and Explanation**

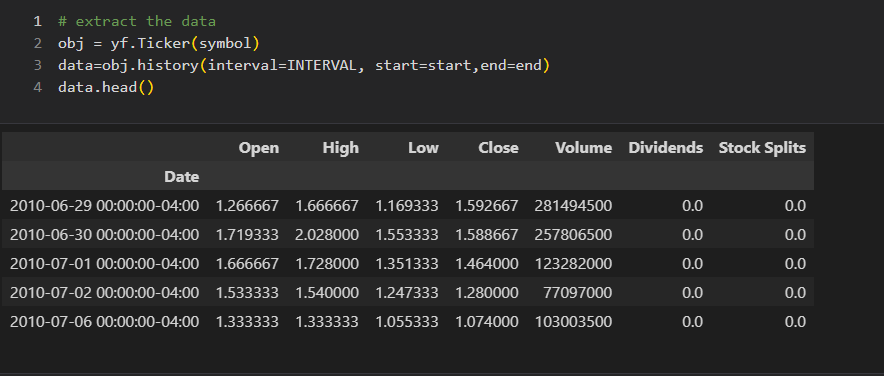
1. Importing Libraries

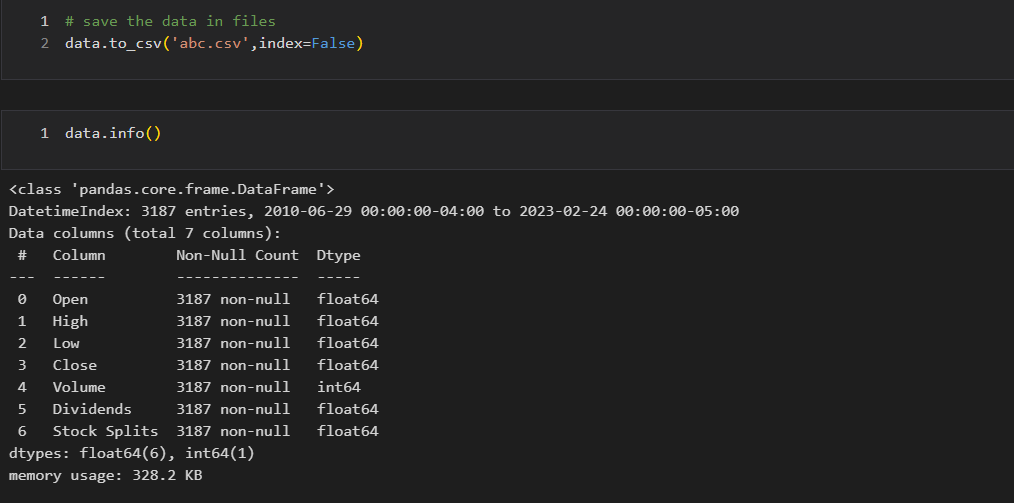
* NumPy contains a multi-dimensional array and matrix data structures. It can be utilised to perform a number of mathematical operations on arrays.
* Matplotlib is a multi-platform data visualization library built on NumPy arrays. It allows us visual access to huge amounts of data in easily digestible visuals. Matplotlib consists of several plots like line, bar, scatter, histogram etc.
* Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data.
* The datetime module supplies classes for manipulating dates and times. While date and time arithmetic is supported, the focus of the implementation is on efficient attribute extraction for output formatting and manipulation.
* The yfinance is one of the famous modules in Python, which is used to collect online data, and with it, we can collect the financial data of Yahoo. With the help of the yfinance module, we retrieve and collect the company's financial information (such as financial ratios, etc.)



1. Extracting Data

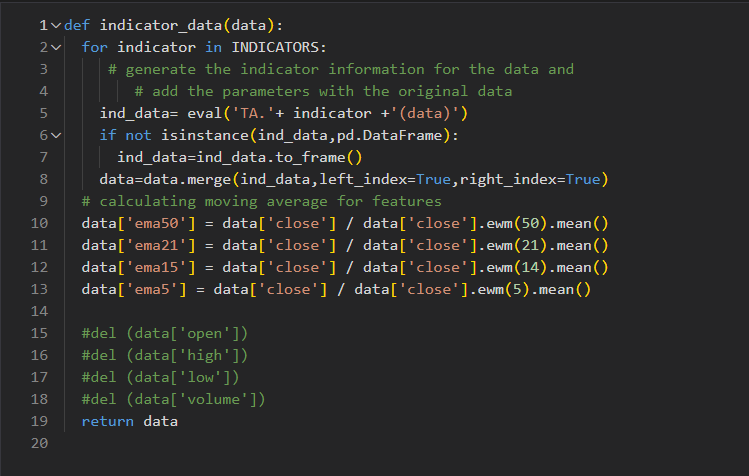






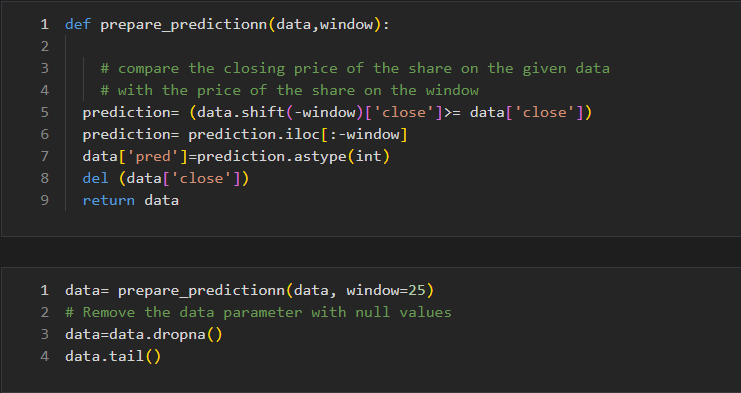
1. **Feature engineering**

* **RSI (Relative Strength Index):** It is a momentum indicator that compares the magnitude of recent gains to recent losses in an attempt to determine overbought and oversold conditions of an asset.
* **MACD (Moving Average Convergence Divergence):** It is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price.
* **STOCH (Stochastic Oscillator):** A momentum indicator that compares a stock's closing price to its price range over a given period of time. It is used to identify overbought or oversold conditions.
* **ADL (Accumulation/Distribution Line):** A momentum indicator that uses volume flow to assess the buying and selling pressure of a security.
* **ATR (Average True Range):** A volatility indicator that measures the range between the high and low prices of a security over a given period of time.
* **MOM (Momentum Indicator):** A momentum indicator that compares the current closing price of a security to its closing price "n" periods ago.
* **MFI (Money Flow Index):** A momentum indicator that uses both price and volume to measure buying and selling pressure.
* **ROC (Rate of Change):** A momentum indicator that measures the percentage change in price over a specified period of time.
* **OBV (On-Balance Volume):** A momentum indicator that uses volume to assess the buying and selling pressure of a security.
* **CCI (Commodity Channel Index):** A momentum indicator that measures the deviation of an asset's price from its statistical mean.
* **EMV (Ease of Movement):** A volume-based indicator that is used to assess the relationship between price and volume.



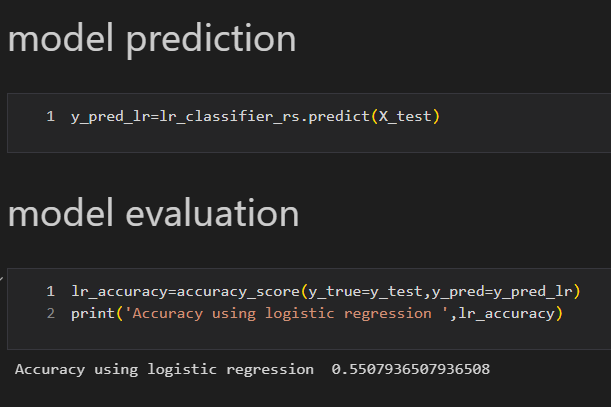
1. Prepare the prediction parameter

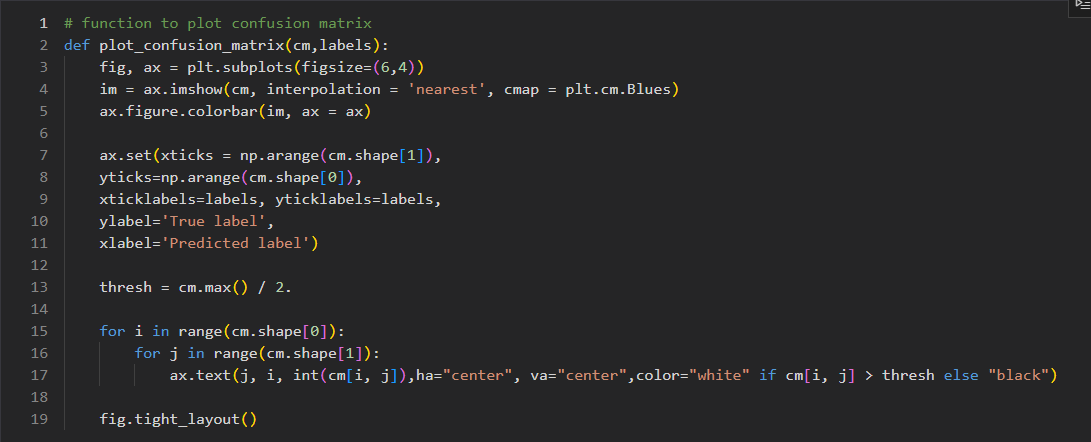
It is based on the idea that we can first learn or provide a distributed representation for each input feature and then learn (with another neural network called the parameter prediction network) how to map a feature's distributed representation to the vector of parameters specific to that feature in the classifier neural network (the weights which link the value of the feature to each of the hidden units).

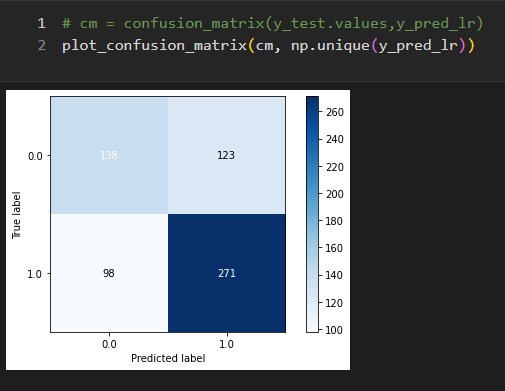


1. Logistic Regression with hyperparameter tunning and cross validation

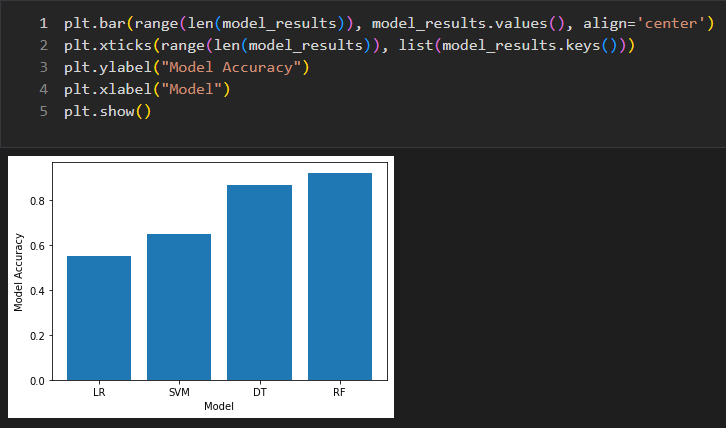
Logistic regression is a classification algorithm. It is used to predict a binary outcome based on a set of independent variables.  A **binary outcome** is one where there are only two possible scenarios—either the event happens (1) or it does not happen (0). **Independent variables** are those variables or factors which may influence the outcome (or dependent variable).







1. Compare the model performance



THANK YOU!