TECHNOCOLABS DATA ANALYSIS INTERNSHIP

**PROJECT REPORT**

**TITLE: Predicting Stock price of Microsoft**



**AIM:**

In this project, we aim to predict whether the price in the next minute will go up or down, using the time series data of stock price, technical-analysis indicators, and trading volume.

# ABSTRACT:

Stock market prediction is an attempt of determining the future value of a stock traded on a stock exchange. This project focuses on classification problem, predicting the next-second price movement and acting upon the insights generated from our models. We implemented multiple machine learning algorithms including: support vector machines (SVM) linear and polynomial , to determine the price in the next minute. Using the predicted results from our models to generate the portfolio value over time, support vector machine with polynomial kernel performs the best among all of our models and gives best accuracy that all other models.

# INTRODUCTION:

The ability to precisely predict the price movement of stocks is the key to profitability in trading. Many investors spend time actively trading stocks in hope of outperforming the market, colloquially referred to as a passive investment. In light of the increasing availability of financial data, prediction of price movement in the financial market with machine learning has become a

topic of interests for both investors and researchers alike. Insights about price movements from the models could help investors make more educated decisions. In this project, we aim to focus on making short term price movements prediction using the timeseries data of stock price, commonly used technical-analysis indicators, and trading volume. Such predictions will then be used to generate short-term trading strategies to capitalize on small price movements in highly liquid stocks.

# Exploratory Data Analysis Report

# Historical Dataset with 1 day time interval

The Eda team downloaded the dataset of Microsoft with 1 day interval

NIFTY DATASET

NIFTY Samples: 22806

Start Date: 2021-01-01

Start Time: 09:16:00

End Date: 2021-03-31

End Time: 15:31:00

**Type,Date,Time,open,high,low,close,**

NIFTY,20210101,09:16,13997.9,14020.85,13991.35,14013.15,

NIFTY,20210101,09:17,14014.85,14018.55,14008.15,14009.05,

NIFTY,20210101,09:18,14008.05,14013.1,14005.05,14012.7,

NIFTY,20210101,09:19,14013.65,14019.1,14013.65,14016.2,

NIFTY,20210101,09:20,14015.45,14017.8,14011.95,14015.45,

# Microsoft Dataset 1 min interval

The Eda team downloaded the dataset of Microsoft with 1 minute interval

samples: 2726

Start Date: 2021-04-23

Start Time: 09:30:00-4:00:00

End Date: 2021-04-30

End Time: 15:55:00-04:00:00

**Date,Open,High,Low,Close,Adj Close,Volume**

1986-01-13 ,0.08854199945926666, 0.1015629991889, 0.08854199945926666, 0.09722200036048889, 0.061750609427690506,1031788800

1986-03-14, 0.09722200036048889, 0.10243099927902222, 0.09722200036048889 ,0.10069400072097778,0.0639558807015419,308160000

1986-03-17, 0.10069400072097778, 0.10329899936914444, 0.10069400072097778, 0.10243099927902222,0.0650591105222702,133171200

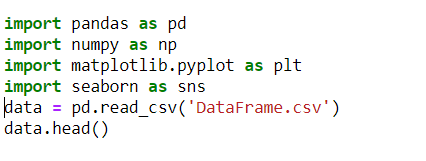
1986-03-18, 0.10243099927902222, 0.10329899936914444, 0.09895800054073334, 0.09982600063085556,0.06340455263853073,67766400

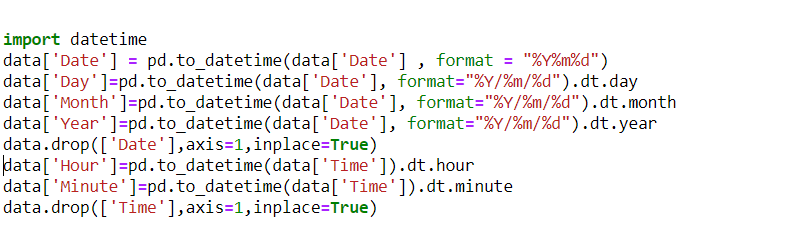
1986-03-19, 0.09982600063085556, 0.10069400072097778, 0.09722200036048889, 0.09809000045061111,0.062301911413669586,47894400

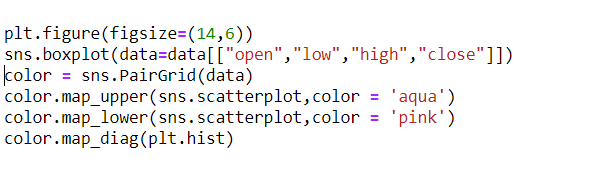
Model we are going to build for stock price prediction are as follows:

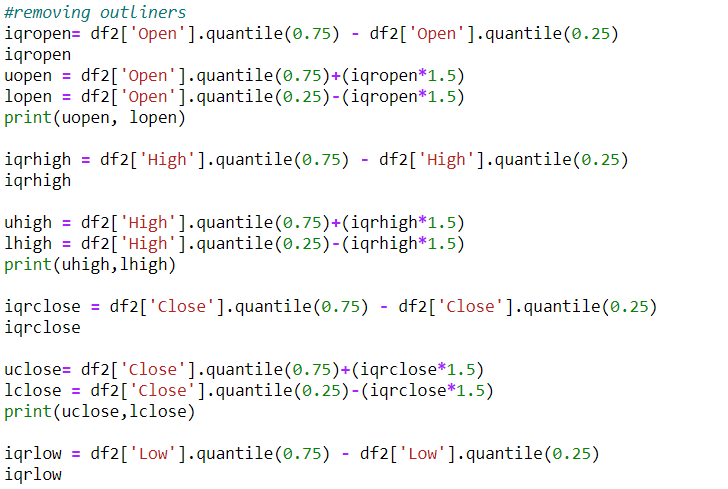
1. SVM linear
2. SVM polynomial

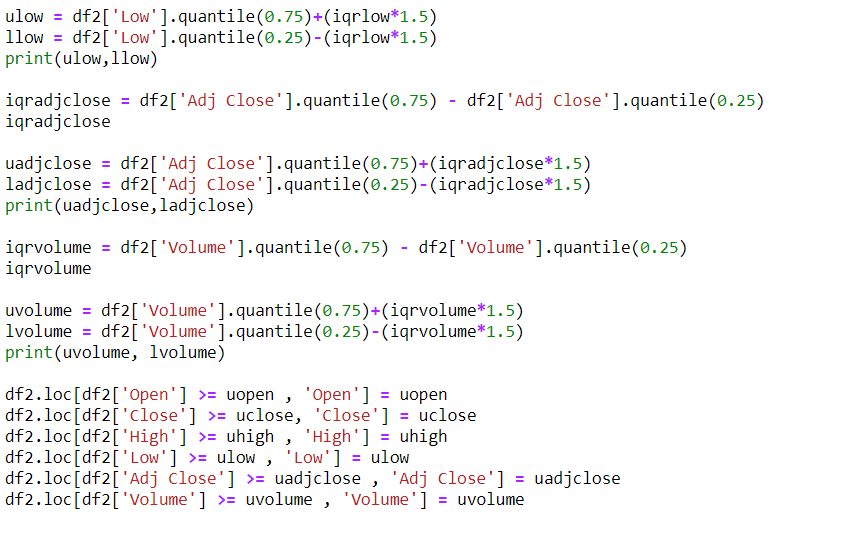
**Code:**

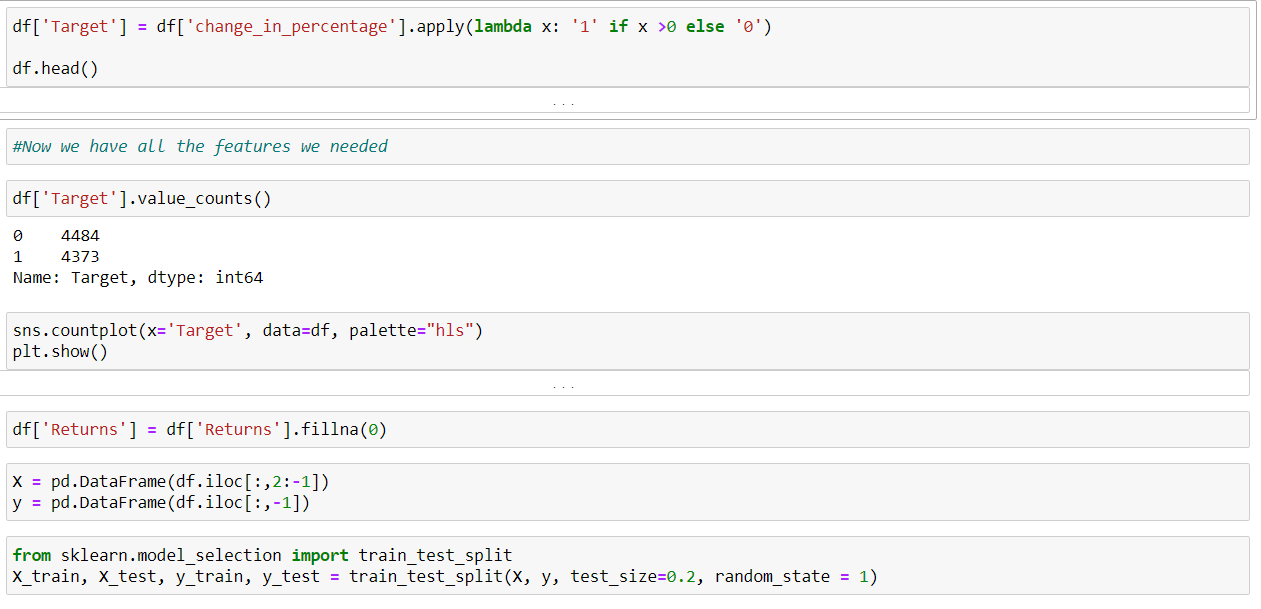


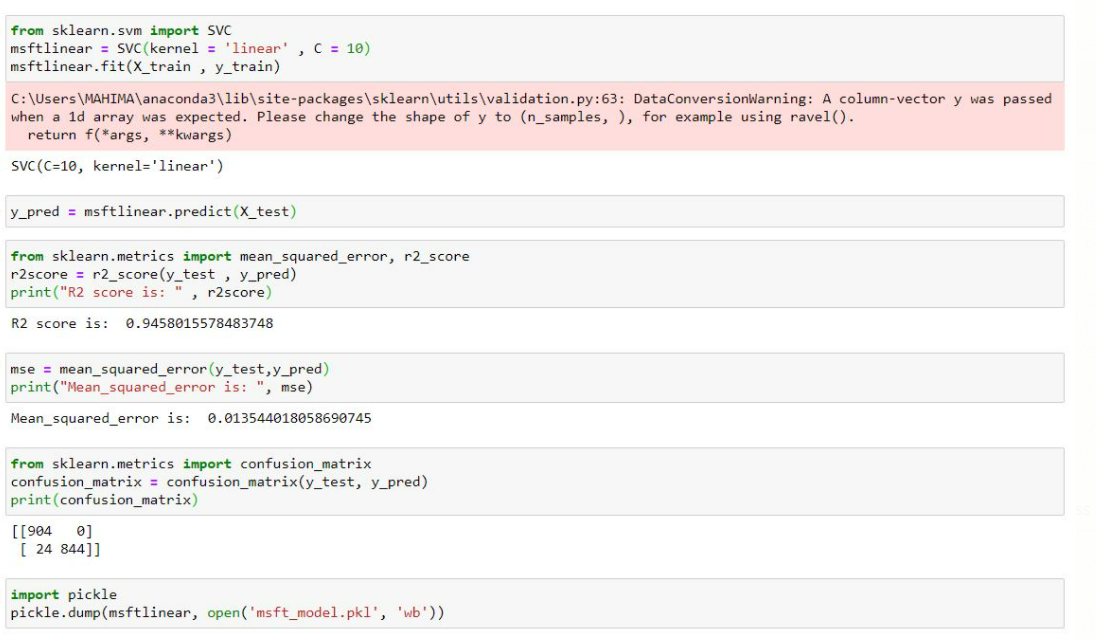


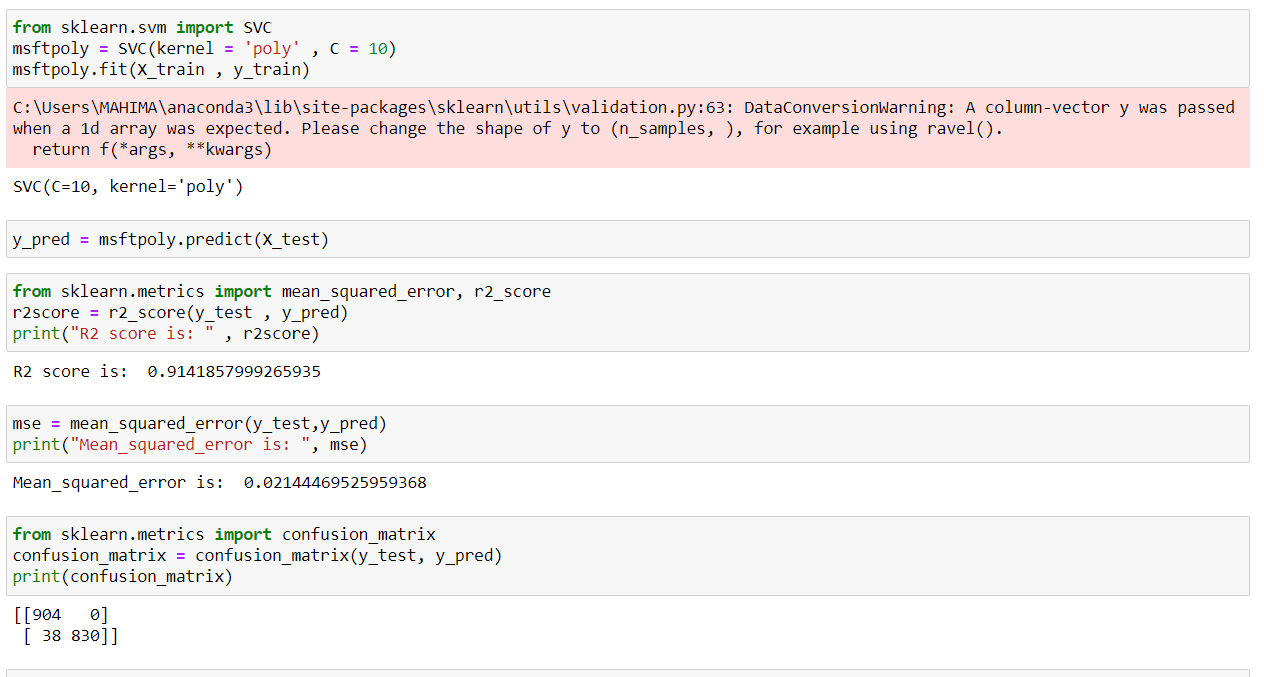












## SVM:

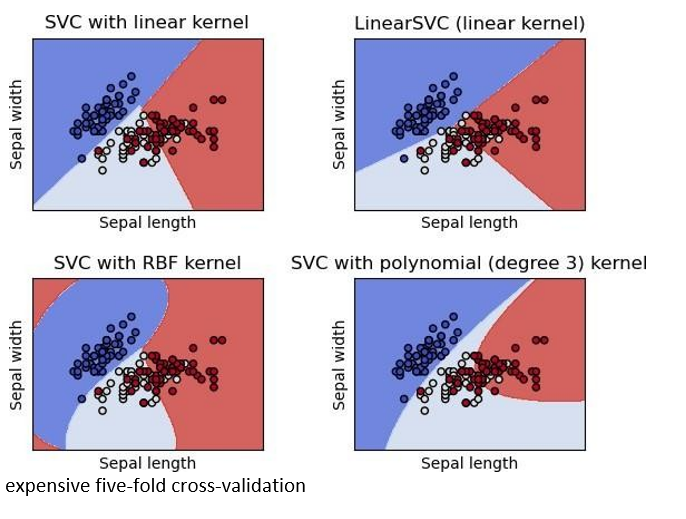
Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outlier’s detection.

The advantages of support vector machines are:

* Effective in high dimensional spaces.
* Still effective in cases where number of dimensions is greater than the number of samples.
* Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
* Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

The disadvantages of support vector machines include:

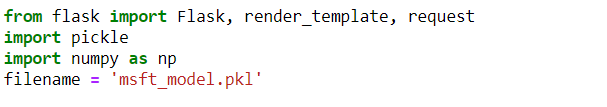
* If the number of features is much greater than the number of samples, avoid over-fitting in choosing Kernel functions

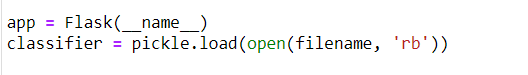


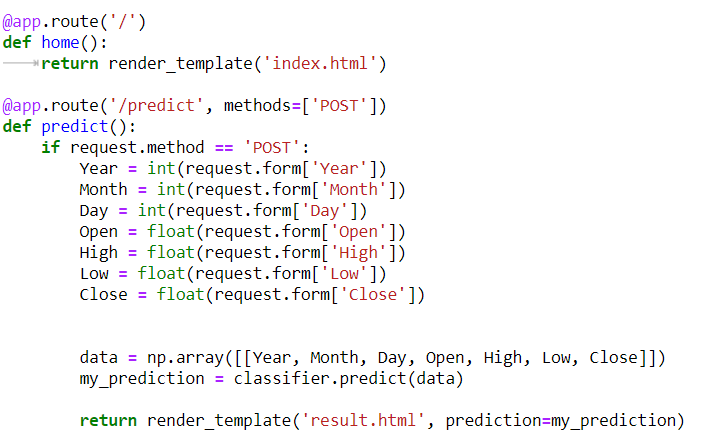
**Model Deployment:**

Deployment Link: <https://msft-stock-prediction.herokuapp.com/>

Part A:







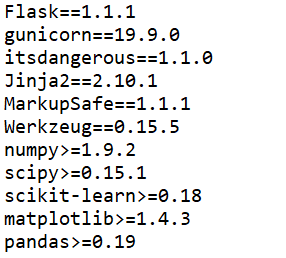


**Part B:**

**Creating the HTML template**



**Requirements.txt**



Part C:

Deployment Link: <https://msft-stock-prediction.herokuapp.com/>





