* **Stock Market Price Prediction :-**

1. **Understanding the problem:**

Our goal is to forecast whether the stock prices of Microsoft (MSFT) will rise or fall. By analyzing historical data and using predictive models, we aim to provide insights into the direction of closing values and future stock movements.

1. **Gathering Relevant Data:**

Microsoft Stock Data was extracted using YFinance library in python in csv file

Have 9498 records.

It starts from 1982-01-01 and ends at 2023-11-17

Has 6 attributes:

**Attributes :**

1. Open: price at which at stock started trading when the market opened on a particular day

2. High: highest price at which a stock traded during the period

3. Low: lowest price of the period

4. Close: price of an individual stock when the stock exchange closed market for the day. It represents the last buy sell order executed between 2 traders

5. Adj Close: adjusted closing price is a calculation adjustment made to the stock’s closing price. it is more complex and accurate than the closing price. The adjustment made to the closing price depicts the true price of the stock because the outside factors could have altered the true price.

6. Volume: total amount of activity during a period of time.

* **Milestones: 30 days to complete the Project**

|  |  |  |
| --- | --- | --- |
| **Milestone** | **Duration** | **Task start - End Date** |
| Stock Prices Dataset Collection | 4 Days | 18 / 11 / 2023 to  20 / 11 / 2023 |
| EDA | 1 Week | 21 / 11 / 2023 to  27 / 11 / 2023 |
| Model Building | 1 Week |  |
| Model Evaluation | 1 Week |  |
| Feedback | 1 week |  |
| Deployment |
| Final presentation | 1 Day |  |

1. **Data Preparation and Exploratory Data Analysis (EDA): Understand the Data:**

Begin by understanding the structure of your dataset. Identify the columns and their meanings.

**Check for Missing Values:**

Examine if there are any missing values in your dataset. Decide on an appropriate strategy for handling them if they exist.

**Convert Dates:**

If your dataset has a time component, convert the date columns to a datetime format.

**Time Series Plot:**

Create a time series plot to visualize how the stock prices have evolved over time. This can provide an initial look at trends and patterns.

**Seasonal Decomposition:**

Use seasonal decomposition techniques (like STL decomposition) to break down the time series into its trend, seasonality, and residual components.

**Statistical Summary:**

Compute basic statistical measures such as mean, median, standard deviation, etc. for your numerical features.

**Histograms and Density Plots:**

Plot histograms or density plots for numerical variables to understand their distributions.

**Correlation Matrix:**

Create a correlation matrix to examine relationships between numerical variables. This is particularly important in understanding potential features for modeling.

* **Preliminary Visualizations:**

**Box Plots:**

Use box plots to visualize the distribution of stock prices across different time periods or categories (e.g., months, days of the week).

**Violin Plots:**

Similar to box plots, violin plots provide insights into the distribution of data, especially if there are multiple categories.

Autocorrelation Function (ACF) and Partial Autocorrelation

**Function (PACF) Plots:**

Create ACF and PACF plots to identify potential autocorrelation patterns in the time series.

**Rolling Statistics:**

Plot rolling statistics, such as rolling mean and rolling standard deviation, to identify trends or shifts in volatility.

**Volatility Clustering:**

Investigate if there are periods of high and low volatility. Visualize volatility clustering using plots.

**Return Distribution:**

Plot the distribution of daily returns to understand the risk and volatility in the stock prices.

**Heatmap:**

Create a heatmap to visualize correlations between different variables.

**Trading Volume Analysis:**

Analyze the trading volume over time to identify potential correlations with stock price movements.

1. **Model Building:**

1. Arima

2. Sarima

3. fbprophet

4.lstm

5.linear

**6.Evaluation:**

**MAE, RMSE AND AIC**

**7.Final Model :**

After evaluation we will deciding the final model based on the model giving best result

**8. Deployment:**

We will use streamlit for our deployment.

First, we will deploy offline and then we will try to deploy the project online using cloud like GitHub.