BLUEPRINT

STOCK MARKET PREDICTION

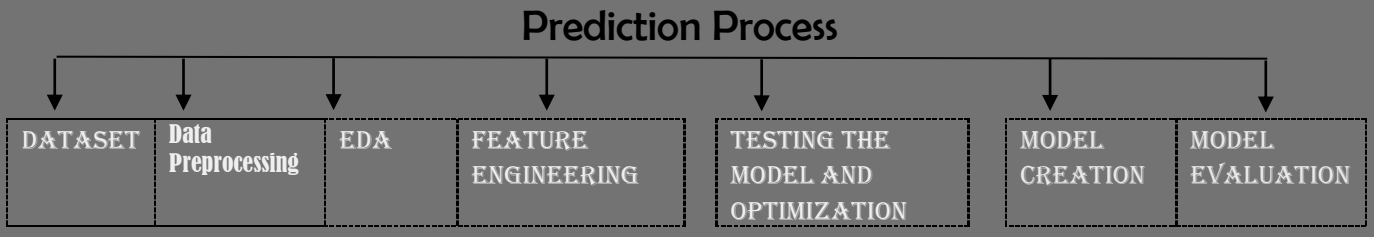
Team 10

Team Members:

Siddhi Bansal

Shubhi Jain

Geetanjali Bakshi



DATASET

- Stock_data.csv (19586 rows, 8 columns)
- MSFT.csv(8857 rows, 7 columns)
- File Format: csv

DATA PREPROCESSING

- For dataset 1: Neither missing values nor duplicate rows present.
- For dataset 2: Neither missing values nor duplicate rows present.

EDA

- Detect Outliers , anomalies and extract important variables.
- Grouping data [Basic grouping with group by].
- Checking correlation between variables of dataset.
- Show datatypes ,shape and other basic information of variables.
- Data visualization (plotting the graphs like line plot, bar plot, heatmap etc.

Feature Engineering

- Scaling: It is done to get the features on the same scale.
- Transformation: It is done to normalize the data(feature) by a function.

Testing the model and optimization

- **Split the data into train and test(train test split) .**
- **Optimization**

Model Creation

- **It is an iterative phase where we continuously train and testmachine learning models to decide the best one of them.**
- **Linear Regression, Logistic Regression and some other algorithms are used to predict the Stock market data.**

Model Evaluation

- **Check for accuracy by calculating error metrices: RMSE , MSE , MAE depending on model and select the one with largest accuracy score.**
- **Performance Measures.**
- **Check Functionality.**