STOCK PRICE PREDICTION USING APPLIED DATA SCIENCE

INTRODUCTION:

The stock market is a dynamic and complex system influenced by a multitude of factors, including economic indicators, news events, investor sentiment, and market trends. Predicting its movements has been a subject of great interest for investors, traders, and researchers alike. Applied Data Science provides a powerful toolkit to analyze historical data, extract meaningful patterns, and build models to make informed predictions about future market behavior

OBJECTIVES:

Project Objectives:

1.Stock Price Demographics Analysis:

-Identify and analyze key characteristics of stocks

2.Analyze Market trends:

-Analysis the market trends of each stock values

3.Optimize portfolio allocations:

-Investigate any correlations between stock trends

Analysis Approach:

1.Data Collection and Preprocessing:

-Acquire historical stock market data, including price, volume, and relevant financial indicators, from reputable sources.

-Clean and preprocess the data, handling missing values, and ensuring consistency.

2.Data Splitting:

- Split the dataset into training and testing sets.Typically 80% for training and 20% for testing

3.Insights Derivation:

- Derive actionable insights from the analyzed data.

4.Model Selection:

-Choose appropriate machine learning or statistical models based on the project's objective.

-Linear Regression.

-Random Forests.

-Decision Tree.

-Support Vector Machine.

5.Training and Evaluation:

-Split the data into training and testing sets.

-Train the selected models and evaluate their performance using metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

6.Fine-tuning and Hyperparameter Optimization:

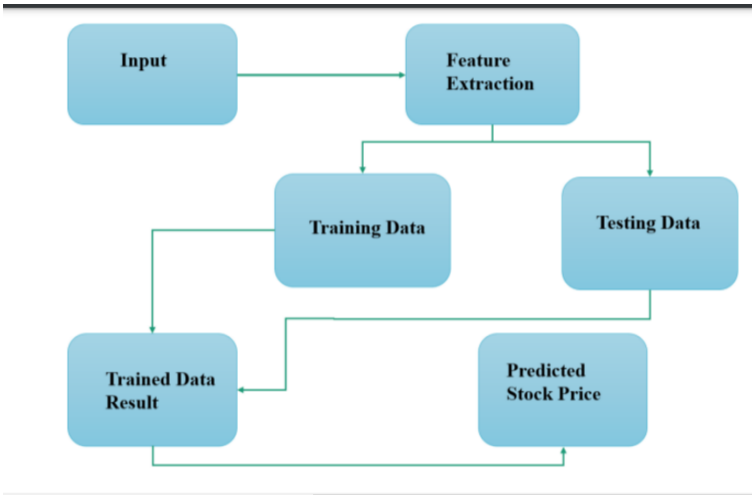
-Experiment with different hyperparameters and conduct cross-validation to improve model performance

7.Model Interpretation:

-Understand how the model arrives at its predictions.

-Utilize techniques like feature importance analysis or SHAP values for interpretability

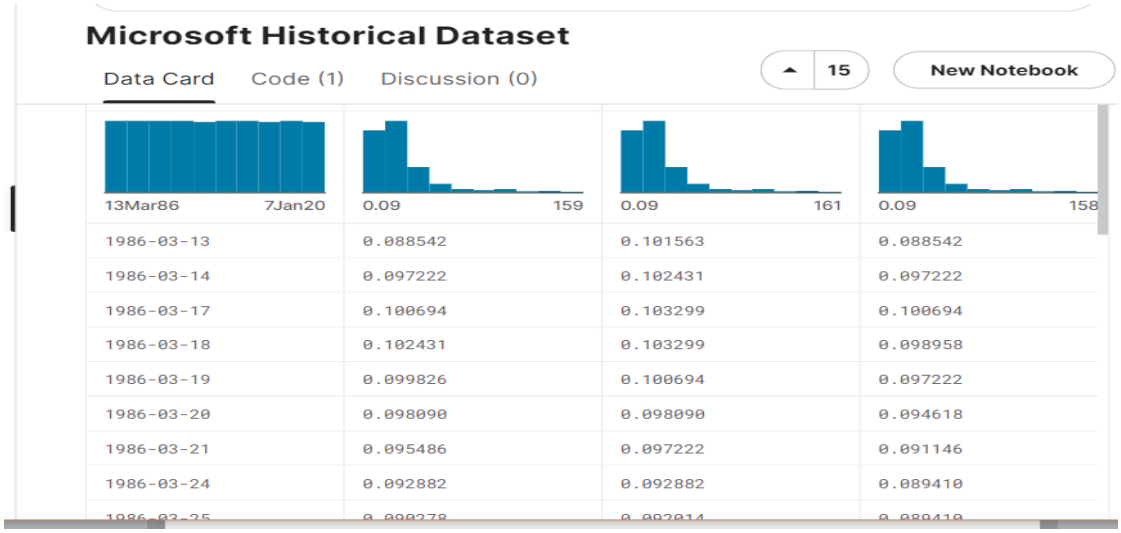
FLOWCHART:



DATA COLLECTION:

Collect data from the resource provided (https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset)

Here is a snippet of the provided dataset,



FEATURE SELECTION:

Utilize machine learning models such as decision trees or random forests to assess feature importance. Features that contribute the most to model accuracy are considered important.

CORRELATION ANALYSIS

Calculate correlations between each feature and the target variable (stock prices or returns). Features with high correlation are likely to be important. However, be cautious of multicollinearity, where two or more features are highly correlated with each other.