***Introduction:***

We have analysed a dataset spanning from January 2, 2019, to December 29, 2023, encompassing stock tickers such as Apple, Google, Microsoft, Intel, and Nvidia.

Our approach involved developing classifiers to predict the directional movement of these stock tickers—whether they will go up or down. Additionally, we implemented regression models to forecast the 1-day future percentage change in stock prices.

We utilized more than five different models in our analysis and subsequently integrated them into an ensemble model using a voting regressor for enhanced accuracy. To complement our quantitative analysis, we also incorporated sentiment analysis to evaluate its impact on the percentage change of the stock prices.

***Data Fetching, Feature Engineering, Exploration, and Preprocessing***

We utilized Yahoo Finance to download historical Adj\_Close stock data for Apple, Google, Microsoft, Intel, and Nvidia. To enhance our analysis, we introduced various technical indicators to capture the progression of the stock's percentage change. These indicators include Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), Commodity Channel Index (CCI), and others.

The data underwent thorough cleaning to address missing values and outliers, ensuring the integrity of our analysis. For sentiment analysis, we gathered stock news data from reputable sources such as Nasdaq News, CNBC, Benzinga, and several other news outlets. The news articles were converted into sentiment scores, ranging from -1 (negative sentiment) to 1 (positive sentiment), with 0 representing neutral sentiment.

We stacked all technical indicators for each stock to analyse the percentage change comprehensively.

***Methodology:***

We have first trained our data on 5 types of different model classifier with hyperparameters using grid search and then sent all the models to an ensemble model to check how better they can perform on the precision and recall of the classification report. Then simultaneously we have performed the regression style percentage prediction and again fed to the ensemble method to check the voting of a good ensemble method.

For sentimental analysis

***Results and Graphs:***

We are checking best performing model for each stock with minimum and maximum RMSE of the different regression models and we are performing cross validation scores to prevent the models from overfitting. We are also checking on the graphs of the residual analysis to analyse the actual values of the percentage change vs the predicted percentage change and helps in identifying the patterns.

SHAP analysis has been performed for the output of machine learning models, providing insights into feature importance and model behaviour.

Sentimental analysis

ATTACH GRPAHS, RESULTS YOU HAVE GOT

***Conclusion:***

We have checked the output results with classification, regression, and sentimental analysis.

Each stock has their own lowest and highest RMSE of different models and you can also check the price of a specific date in the plotted graph.