Data miners of Wall Street

Prediction of Stock Market using Multiple Parameters

-Prem Amal

-Shreyas Sawant

-Pranay Chopra

-Jash Shah

Why prediction of Stock Market is important?

Data is the key in making improved decisions about stocks.

With the help of data we can make better and informed decisions about our Investments, Risks and Portfolio management

Stock Market data is random walk however we can mine the data to derive insights

Scratch the previous slide – Money is the key

- The basic motivation of this project and the audience here is to make some money out of the market and beat the market .
- However, beating the market is not an easy task.
- Our WebApp, will come to your rescue which will fulfill all the requirements of the previous slides and most importantly take you a step closer to getting rich.

Stock market is a complex form of math

- Before diving into the data mining concepts, let us look at what fundamental and technical indicators are!
- **Fundamental Indicators:**
- Book/SH-Book Value
- P/E-Price per Earning Ratio
- P/S-Price per Share
- ROE-Return on Equity
- ROI-Return on Investment
- EPS-Earnings per Share
- Debt/Eq-Debt per Equity

Technical Indicators:

Simple Moving Average

Average True Range

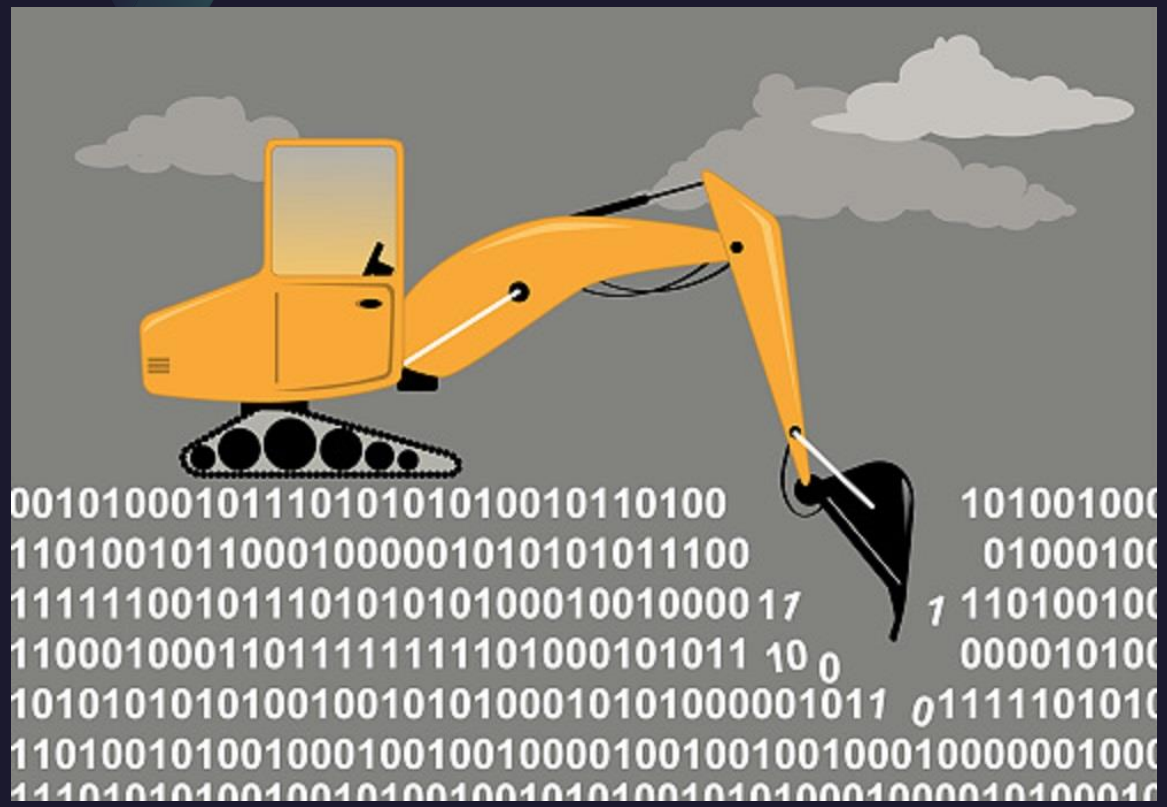
Average Directional Index

Stochastic Oscillators

Relative Strength Index

Bollinger Bands

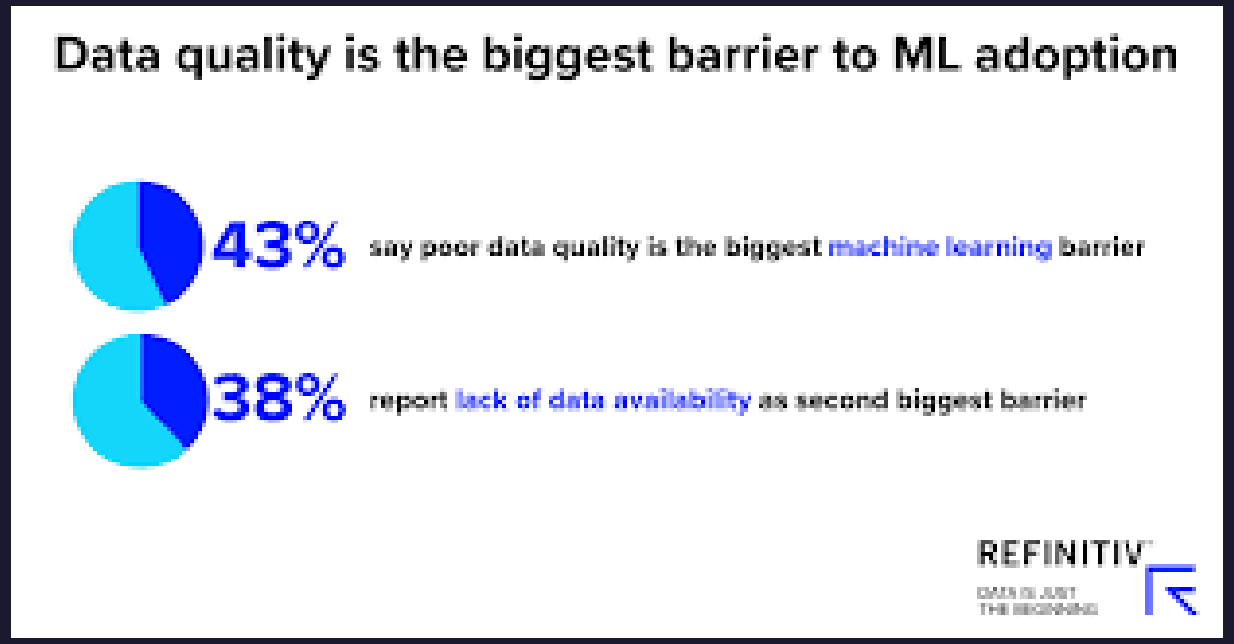
Data Mining starts here!



What's Next?

1) Data Gathering
2) Clustering Stocks
3) Predicting the market
4) Analyzing Sentiments about the market
5) Investment Decision

Shopping cart of our Data



- We have collected data from various API's for multiple purposes.
- For the fundamental indicators we have scraped the **Finviz Website.**
- For the historical stock data we have used **YahooFinance API.**
- For scraping the news article we have used **Finviz Website containing news.**
- For Tweets data we have used an open source library called as **Sns Scrape** to get tweets.

How Unsupervised Learning helped us filter Good Stocks

- K means helped to make clusters based on stocks having similar fundamental characteristics.
- To select the best cluster we have calculated the Sharpe Ratio of each cluster.
- We will select the cluster of stocks having the best Sharpe Ratio.
- The Sharpe ratio is a risk-adjusted measure that evaluates an investment's return in relation to its volatility.

After K means what?
- Generate more data

- Yes, if a strategy involves in predicting the actual stock price, there is a very high chance that it fails.
- Our approach is to predict a trend, but again with the close and open price we are one step closer to failure.
- Hence, extra features have been created taking into account the short term and long term trends.

Classifier Model

Current targeted Stocks include:

Apple

Amazon

Google

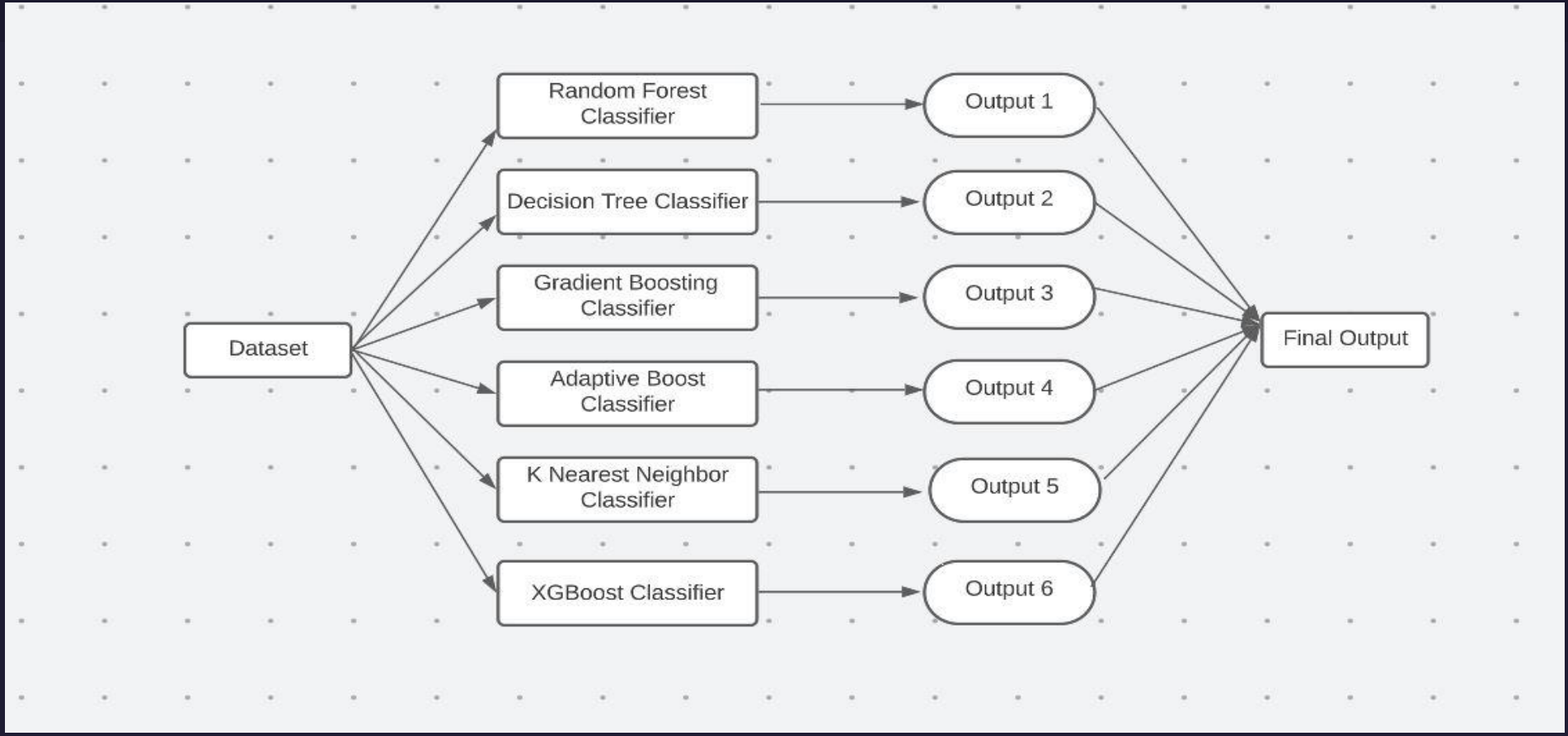
The training data is divided as 95 percent of years when company started stock trading and 5 percent as testing data (Capturing trends in recent time).

The technical indicators were lagged by a window of 2 days for short term trends and another window of 5 days to manage the long term trends.

The target variable is calculated as Today's close price - yesterday's open price.

If the value is positive we tag the data as 1, else we tag the data as 0.

Architecture of Classifier Model



Classifier Model -Continued

- After the generation of target variable, we have created a pipeline of different classification models with their optimal parameters searched using Grid Search Cv.
- After running different models on the training data we have used the Voting Classifier to predict the majority vote.
- Precision Score Metric is used to measure the goodness of classifier models.
- For the different stocks we have gained precision as follows:

  Amazon - 0.68, Apple - 0.63, Google - 0.65

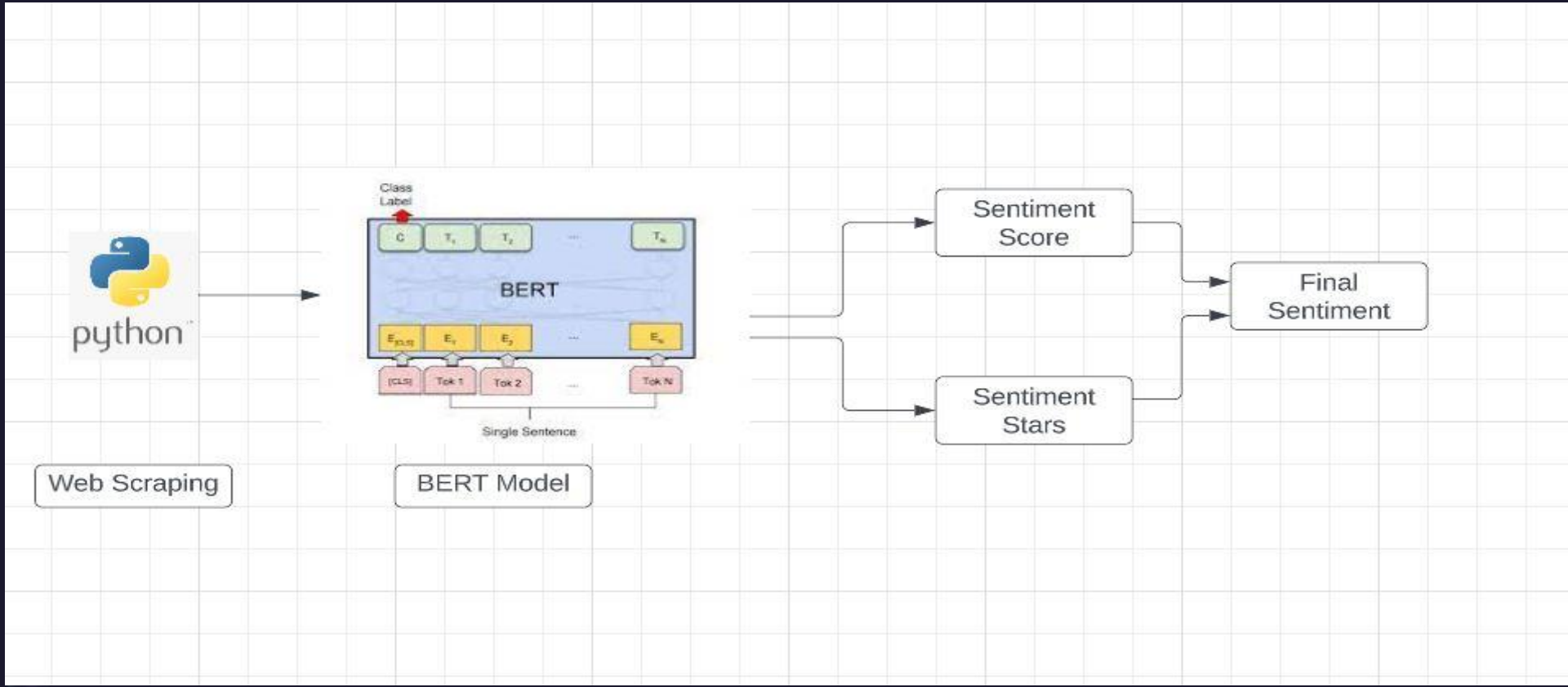
Question to the audience?

- So, I claim I have a model which can make money 65 percent of the times what would be your response?
- Option A) Invest in Us
- Option B) Need more confidence?

News Sentiment Analyzer

- As we need more parameters to take a better decision, we will also apply Sentiment Analysis on the News Articles for a particular stock.
- The sentiment analyzer used is a pre-trained Bidirectional Encoders from transformers model and the output is in the form of stars.
- Single Star with a compound score closer to 1 signifies a strong Negative sentiment.
- Five Stars with a compound score closer to 1 signifies a strong Positive sentiment.
- So the output of the news classifier is the aggregation of all the sentiment score for all news articles on the particular day.

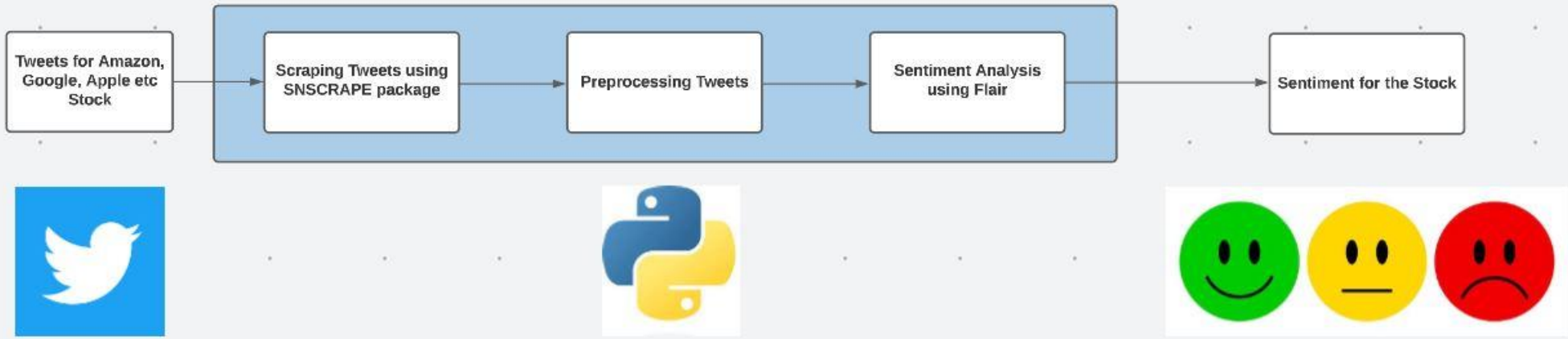
Architecture of Sentiment Analyzer



Explanation of Twitter Sentiment Analyzer

- Similar to the previous slide we have used the keyword related to stock ticker as the input to our scraping website and have generated all the tweets for a particular day.
- Identifying sentiments over twitter helps in concreting the predictions.
- Thus a three fold parameter would auger to a better decision for investments.

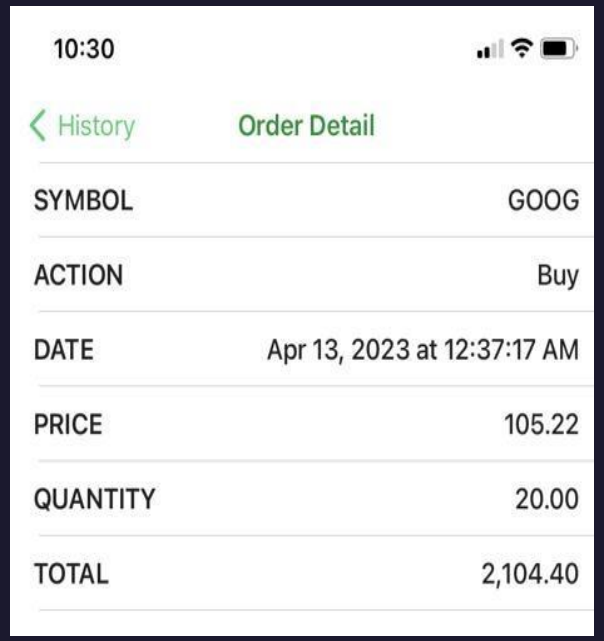
Architecture of Twitter Sentiment Analyzer

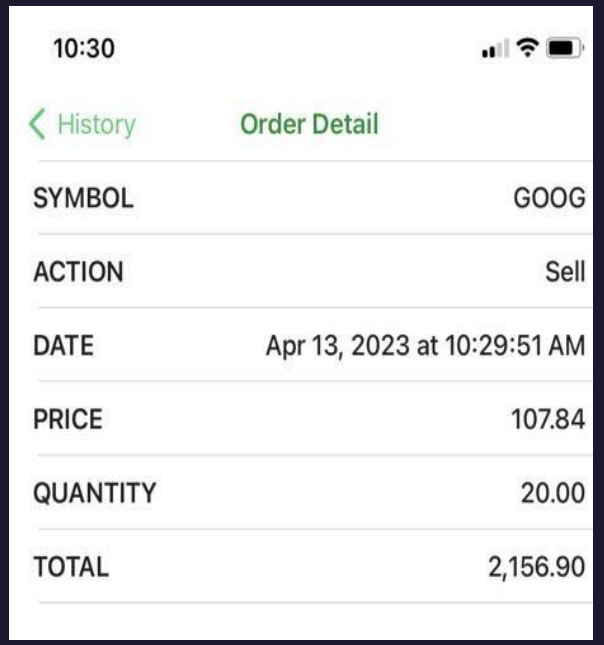


How to use our WebApp?

- The main goal is to focus on the closing price for today and run our WebApp to check whether the stock would be in an upward trend or downward trend.
- For a risk-averse trader if our application shows a Buy Signal the trader can place an order for the stock first thing when market opens.
- Based on our model the trader can wait for the end of the day and sell the price to make some profits.
- We have first hand tested the strategy and on a stock simulator earned some money. Fake money indeed.

Practical results on Google Stock



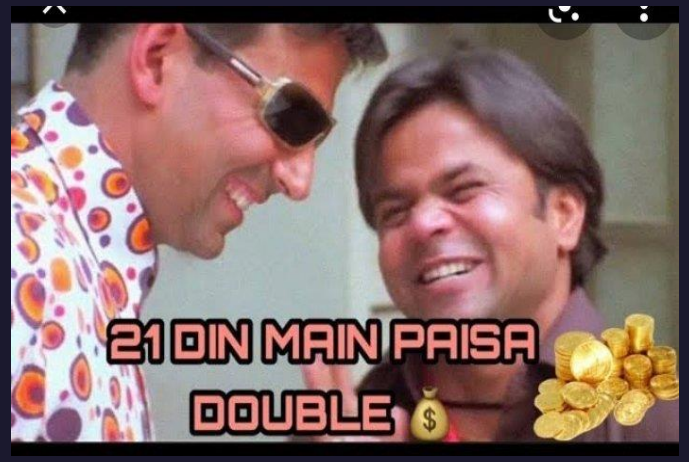


- On 13th April our model predicted a buy signal for Google Stock, i.e. both the News Sentiment Analyzer as well as the Classifier Model, and these are the results of our trade:
- We bought the stock and sold the next day to get a 2 percent profit on our capital.
- Likewise for the entire week our classifier predicted to short Amazon stock and it worked accurately.

Future Scope

- Since the world is a dynamic place and any event can shake the stock market, our next target would be to set up an event-based notification system that can detect News and Tweets in real-time using Paid Financial API's and to stream the data using real time services like Kafka.
- Our model is trained for three stocks, however for a trader according to the needs, we can trade the model on multiple stocks from different sectors and draw our attention away from blue chip stocks.
- Another key aspect of financial long-term investment is to scrape the financial reports of Companies and get hidden insights about the health of the companies and mine the enormous PDFs into something meaningful.
- Looking forward to your questions!

Q/A's?



Thank you!