Stock Analyzer

Using Technical, Fundamental, and Sentimental Analysis

Objective

Determine whether a stock is worth Buying or Shorting

Using Machine Learning Classification, Time Series, and NLP Sentiment Analysis

The Data Sources

Data Sources

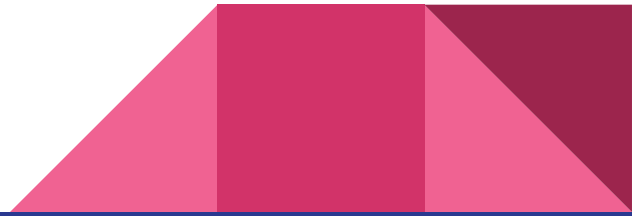
- Stockpup.com
  - Contains Quarterly Reports for 765 stocks dating back to 1990.
- Yahoo Finance
  - CSV files for the Historical Daily Closing Prices.
- Twitter
  - Tweets concerning selected stock.

STOCKPUP

yahoo!

Data Engineering

- Quarterly Reports:
  - ***Percent Changes*** between each report replaced previously listed values.
  - Engineered interaction terms and utilized **PCA**.
  - Determined most important features
  - Created class labels: *Buy, Sell, Hold*



Stock Analysis

How to Analyze a Stock

Fundamentals:

- Observe Financial statements such as Balance Sheets, Income statements, and Cash Flow.
- Basically studying the overall financial health of the company.

Technicals:

- Observe price movements and patterns from the past to determine the future price.
- Utilizing technical indicators such as Moving Averages and MACD.

Sentimental:

- General opinion on stock movements
- People’s feeling towards a specific stock.

Applying Machine Learning Models

Fundamentals:

Implemented Classification Models

- Buy
- Hold
- Sell

Technical:

Implemented Time Series Models

- Used Daily Historical Prices to forecast the next *n* days.

Sentimental:

Implemented NLP Sentimental Analysis:

- Used Tweets to gauge general opinion

Classification

Classification for Fundamentals

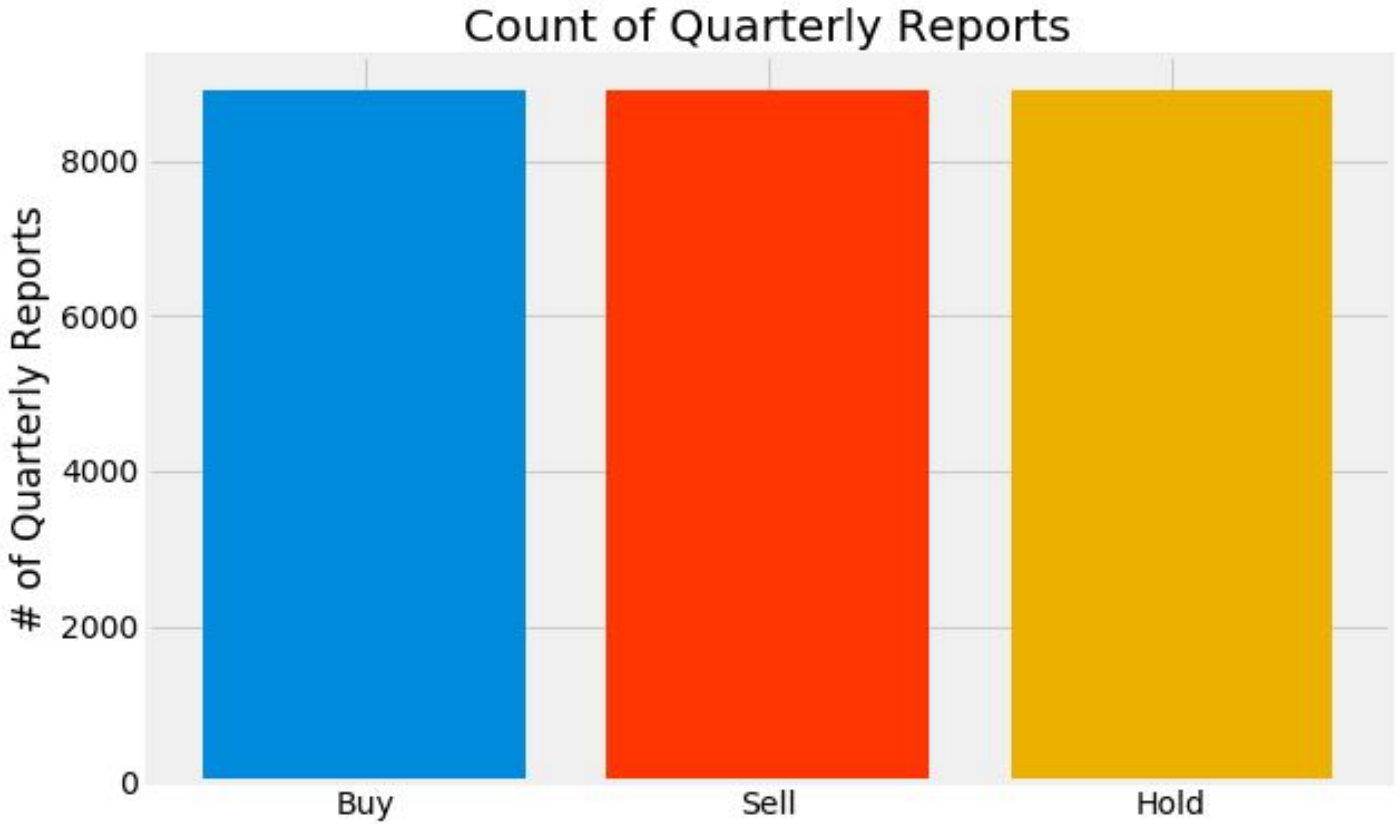
What determines a Buy, Hold, or Sell?

Buys - if next quarter's highest high and lowest low both *increase* by 5% or more.

Sells - if next quarter's lowest low and highest high both *decrease* by 5% or more.

Holds - if neither happens.

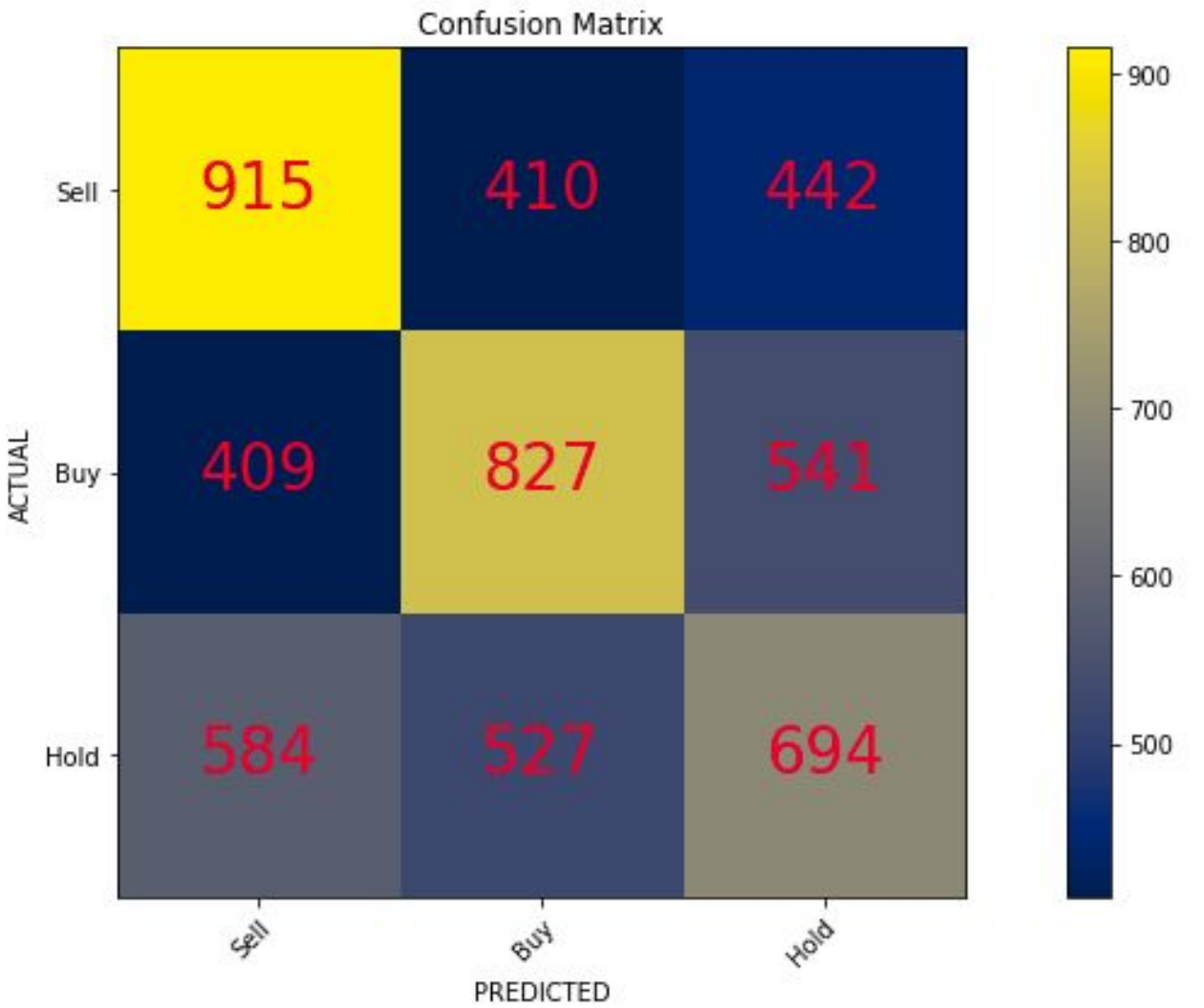
Class Balance



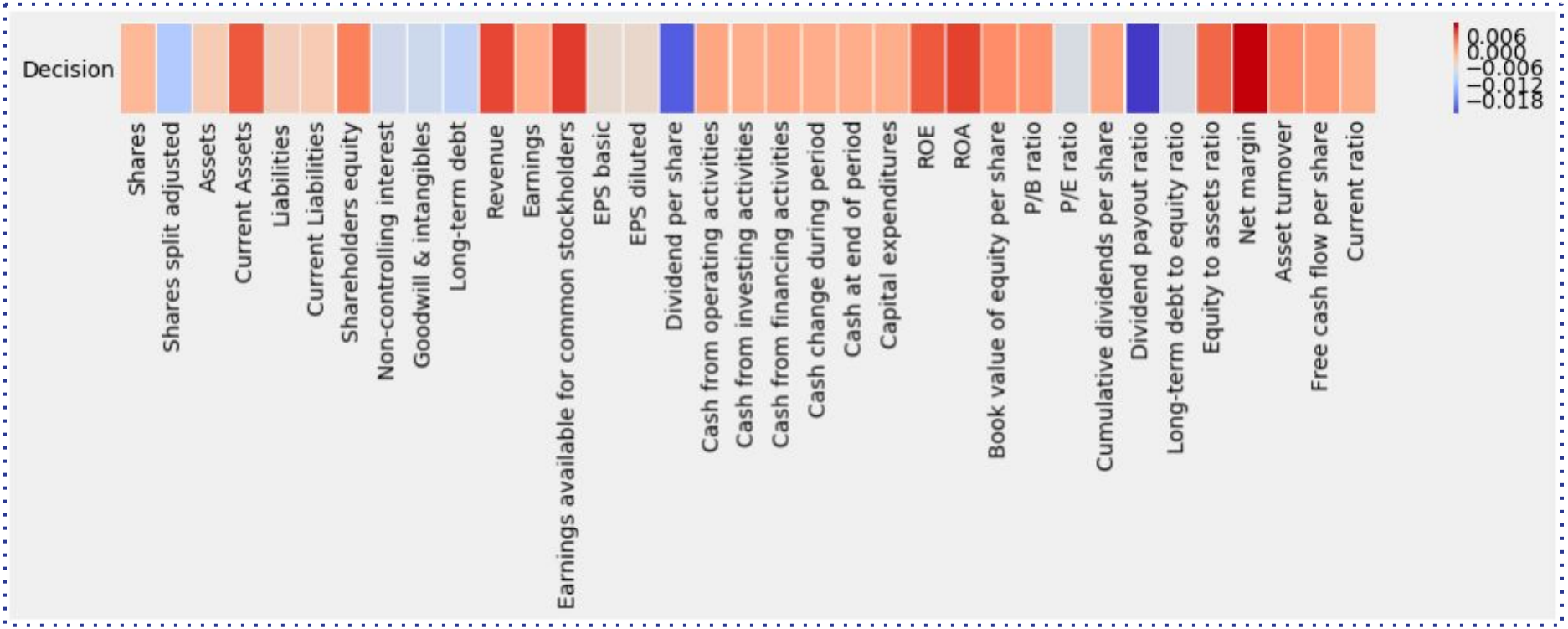
XGBoost

Accuracy Scores:

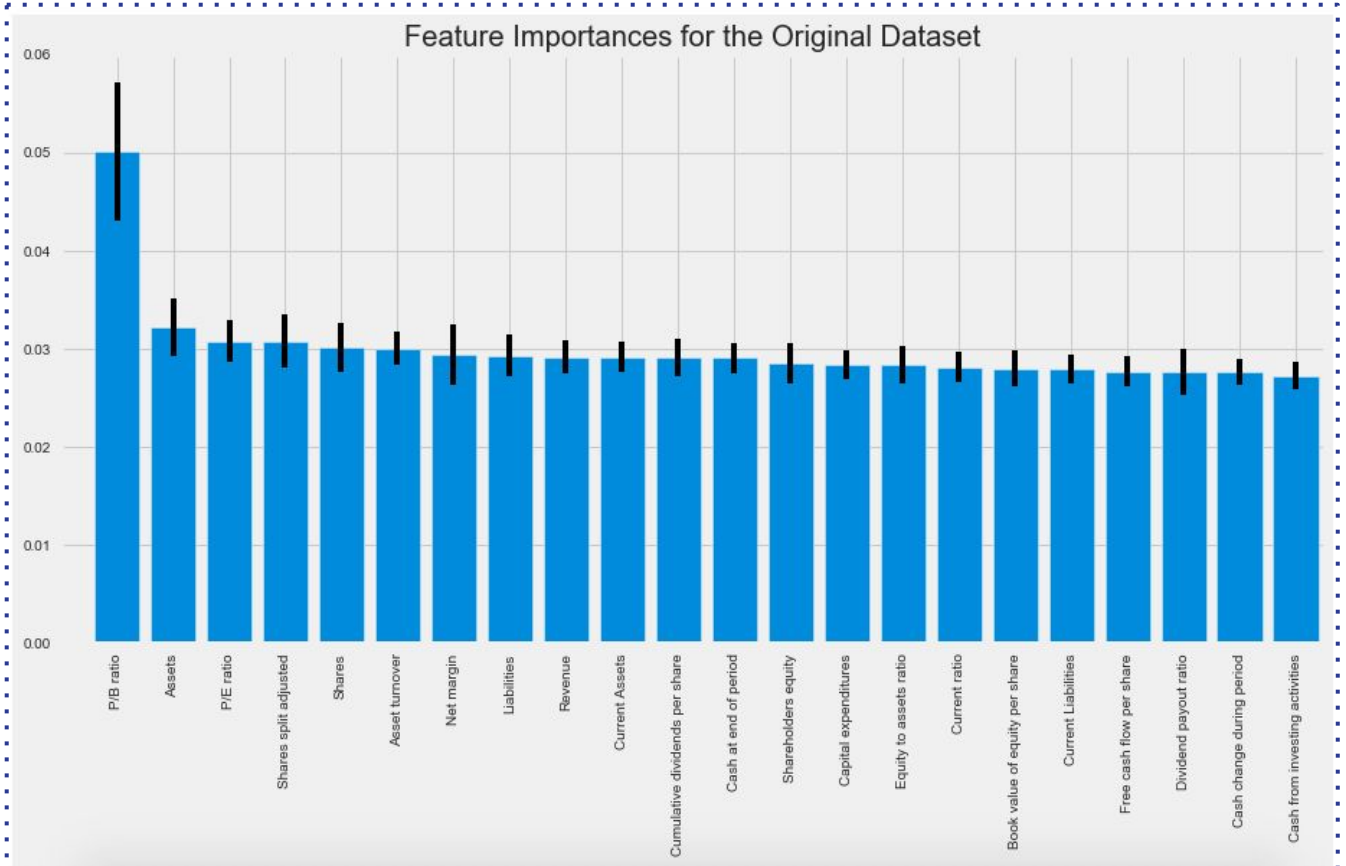
- Testing: 45%
- Training: 47%



Feature Importance - Heatmap



Feature Importance



Top 10

Most Important Features

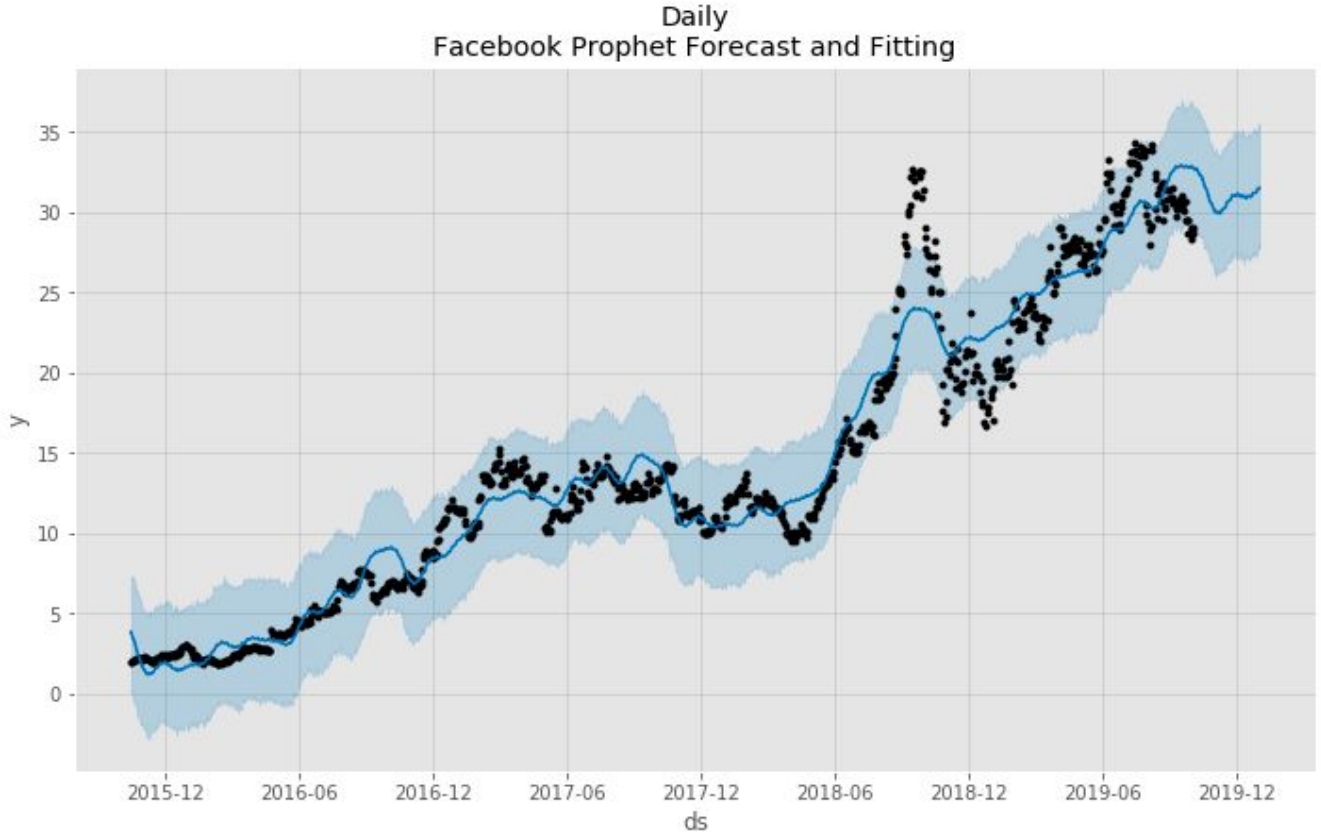
1. P/B Ratio
2. Assets
3. P/E Ratio
4. Shares split adjusted
5. Shares
6. Asset Turnover
7. Net Margin
8. Liabilities
9. Revenue
10. Current Assets

Time Series

Time Series for Technicals

- Observing only the Historical Daily Closing prices of a selected stock.
- Modeling will fit and train to that specific stock and forecast the next desired periods.

Time Series Modeling - Facebook Prophet



Facebook Prophet

- At the core of FBProphet is an **additive regression model.**
- Able to produce quality forecasts with much less effort.
- Quick and efficient compared to other time series models such as SARIMAX.

Deep Learning

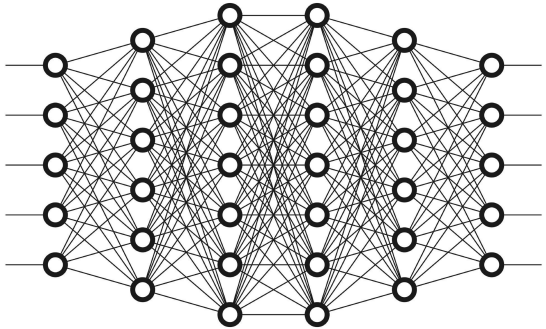
Neural Networks

With Time Series Data

Recurrent Neural Network

- Used a Long Short Term Memory (LSTM) network.
- Trained on historical daily closing prices.
- Only one stock was learned (AMD)

Issues with Deep Learning

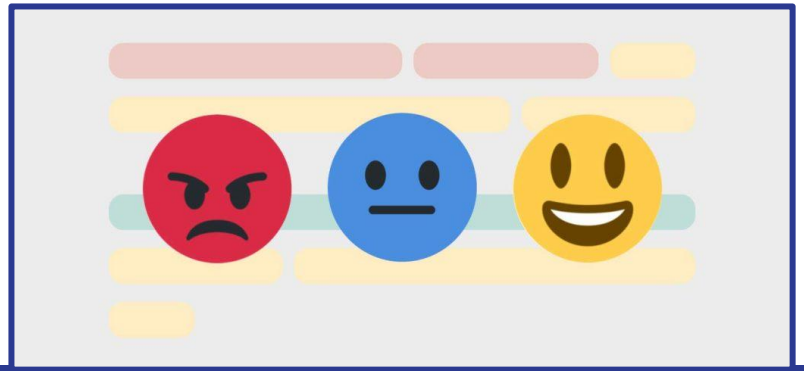


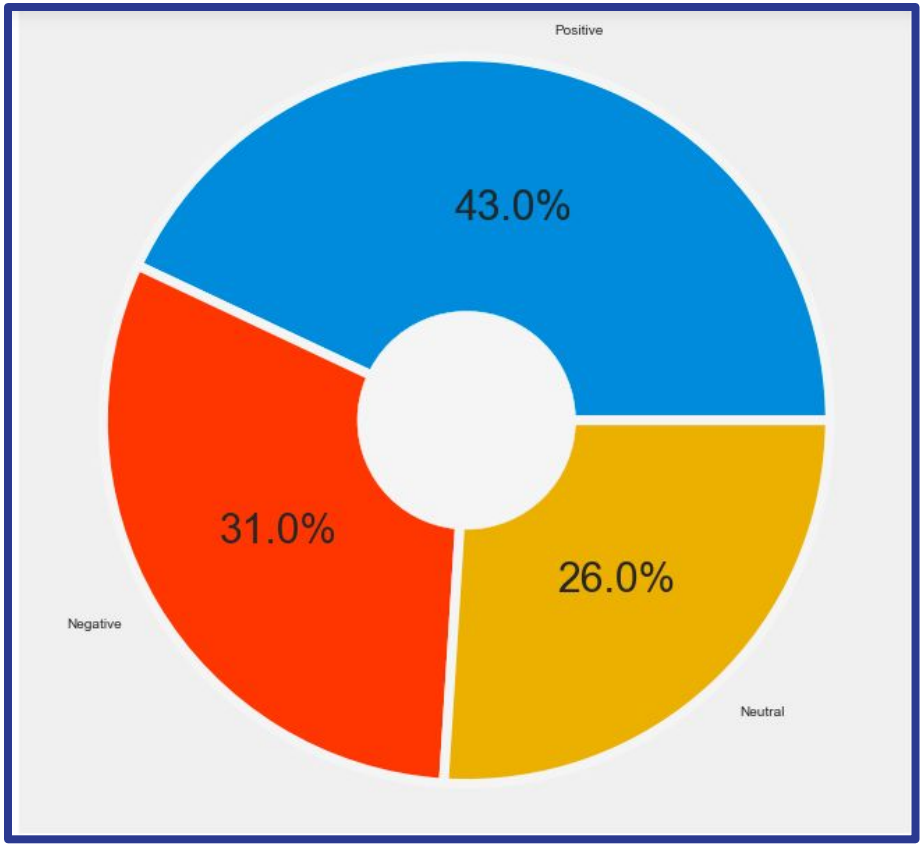
- CPU consumption
- Time to train
  - Much longer than any regular time series model.
- Requires more tuning and experimentation
- Not possible to run a NN on every stock; must be dedicated to only one stock at a time.

Sentiment Analysis

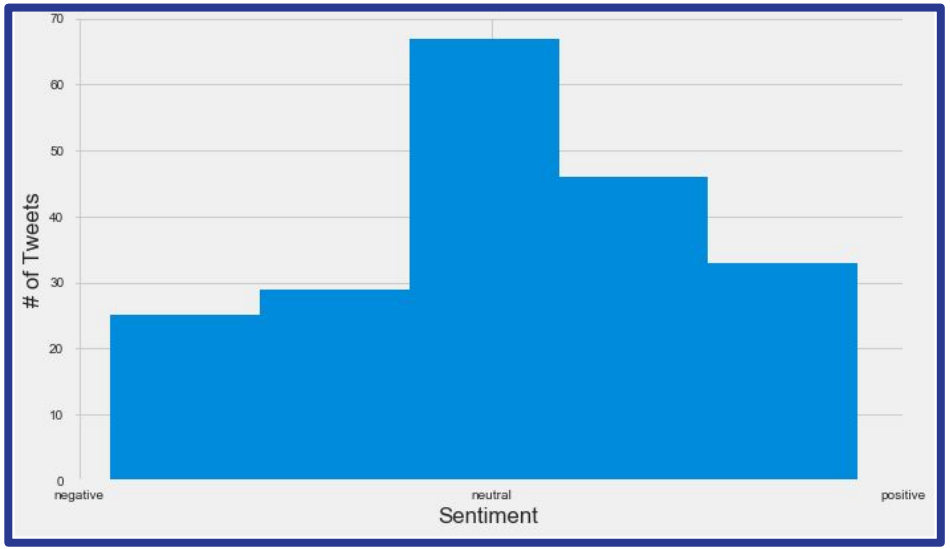
Sentiment Analysis

- Used NLTK and VADER's SentimentIntensityAnalyzer for the compound scores.
- Created simple class labels for Negative, Neutral, and Positive based on the scores.





Percentage of Tweets that leaned towards a specific sentiment



Distribution of Sentiment Scores

Examples of the Sentiment Analysis EDA

Potential Improvements

- Up to date quarterly reports by scraping other sites.
- Add other technical indicators (SMA, MACD, etc.) as exogenous variables.
- More time for neural network training.

Closing

- Stocks can be analyzed in various different ways.
- Use machine learning/deep learning to assist in those way.
- Able to choose stocks without the harm of emotional trading.