

WallStreetBets Sentiment

**Retail Investor Sentiment Analysis
and Stock Movement Prediction**

Kai Cansler





Outline

01

Business
Problem

02

Data
Understanding

03

Visualizations

04

Modeling

05

Conclusion

06

Recommend

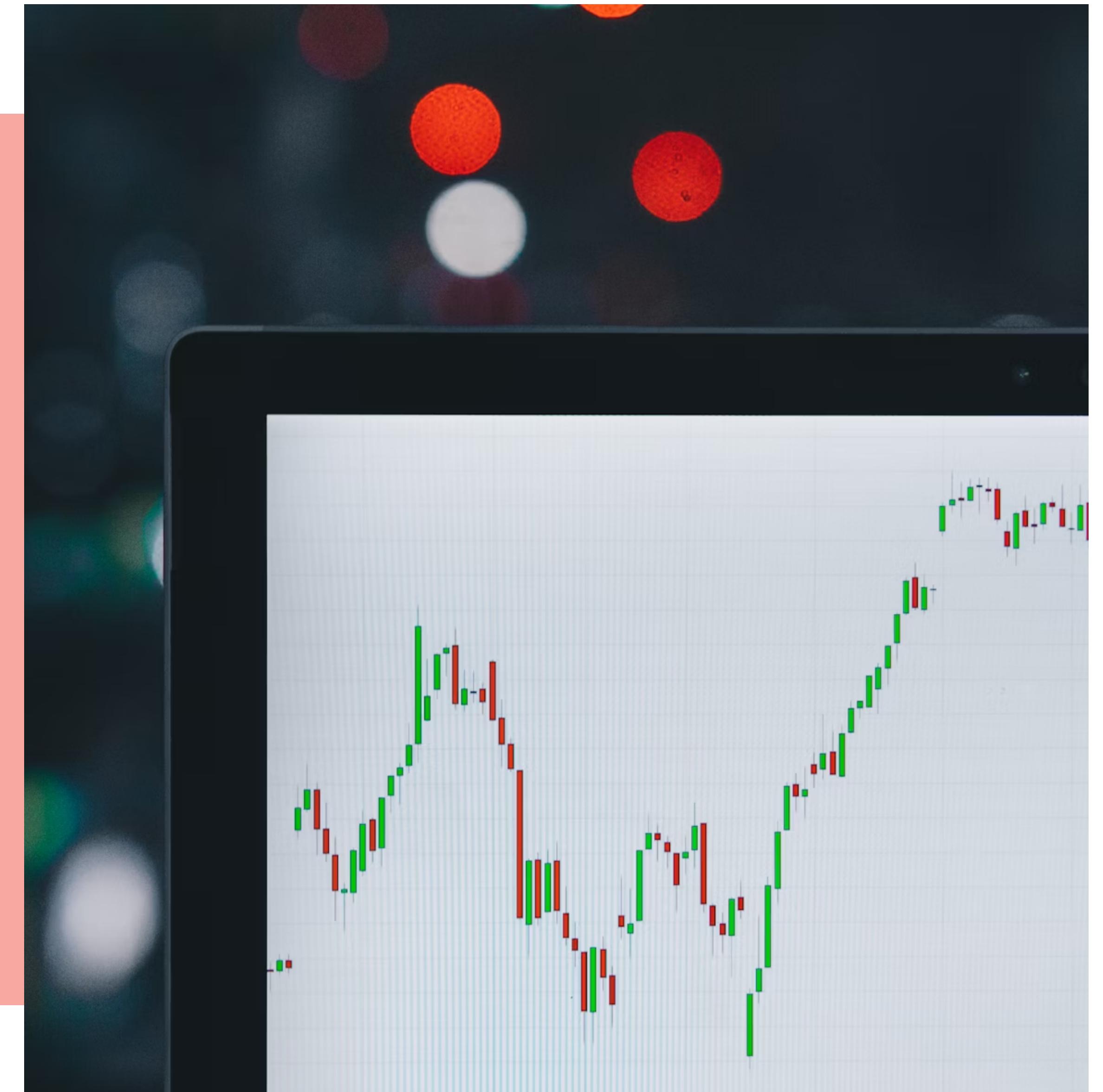
Business Problem

- Retail Traders
- Sentiment Analysis
- Market Movements
- Risk



Data Understanding

- reddit_wsb.csv
 - 1. 50,069 Posts
 - 2. 2020 ➔ 2021
- AlphaVantage
 - 1. Daily Price Data
 - 2. RSI, CCI, and ADX
- Tickers
 - 1. Stock Symbols



Positive Sentiment

Words used in Positive Posts

- Daily and Weekly
 - Retail Investors Discuss
 - Rocket, Buy, and Moon

WordCloud of Positive Post Bodies



Negative Sentiment

Words used in Negative Posts

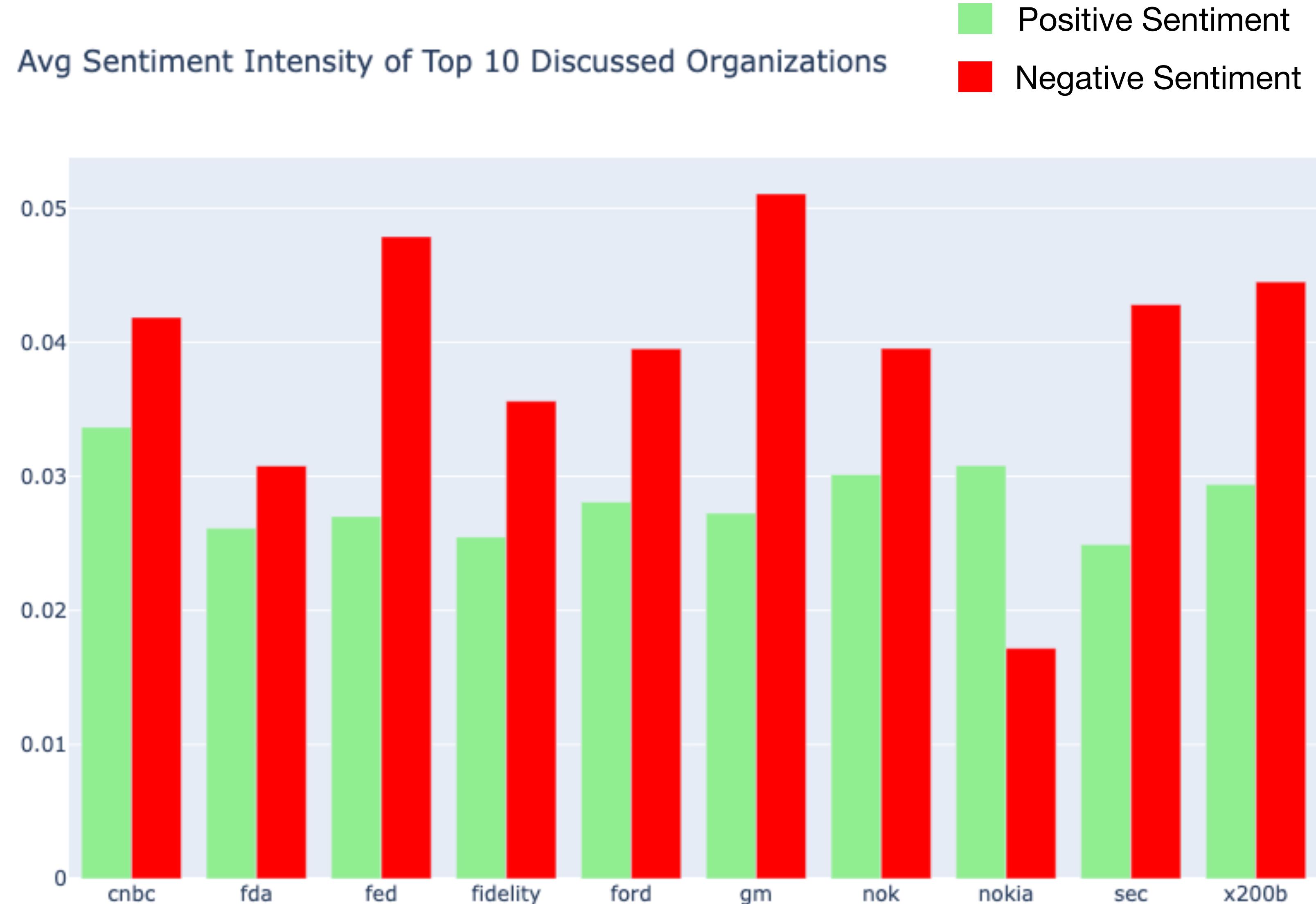
- Loss and Sell
 - Negative view on Robinhood
 - Short, Hold, and Moon

WordCloud of Negative Post Titles



Top 10 Discussed Sentiment on Organizations

- Nokia with higher positive.
- 9 out of 10 have higher negative.
- Fed, GM, & SEC



Modeling

Classification Models

- Gradient Boosting

Success Metric

- Accuracy Score

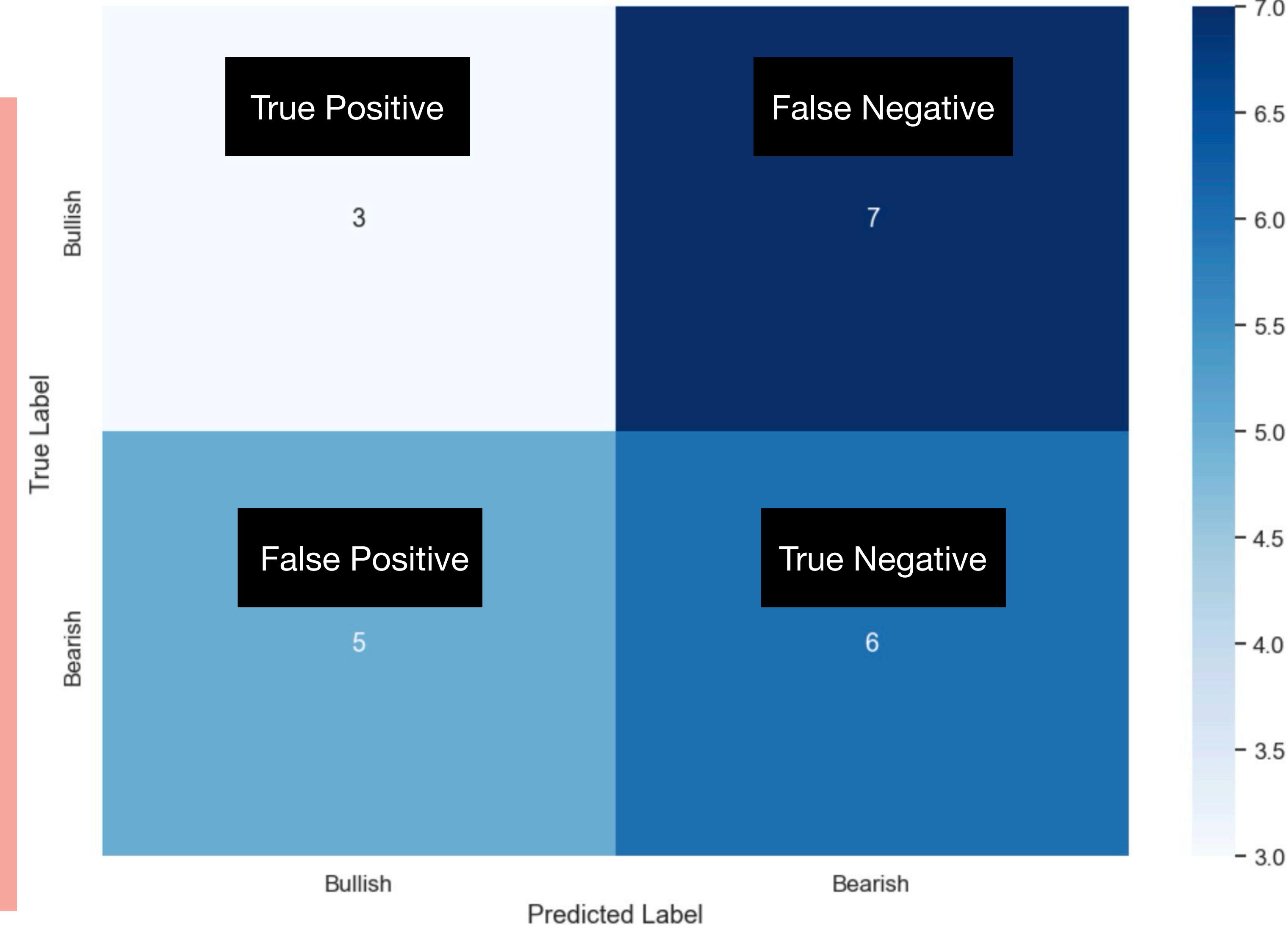


Classifier

Gradient Boosting

- 7 Incorrect Predictions
- Accuracy Score = 42.06%
- Best Model

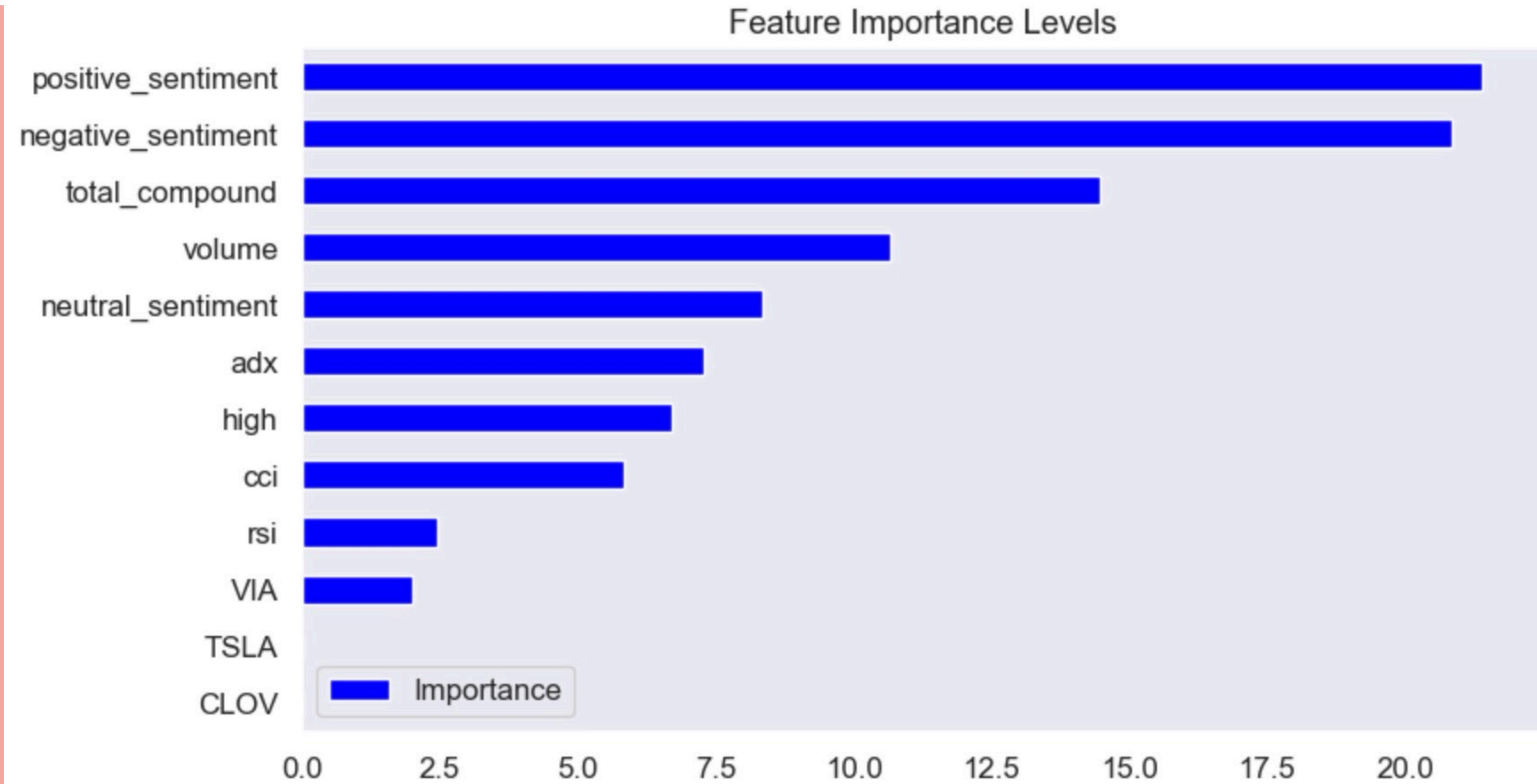
Confidence Matrix



Feature Importance

What has the greatest effect?

- Sentiment is strong
- ADX, CCI, and RSI

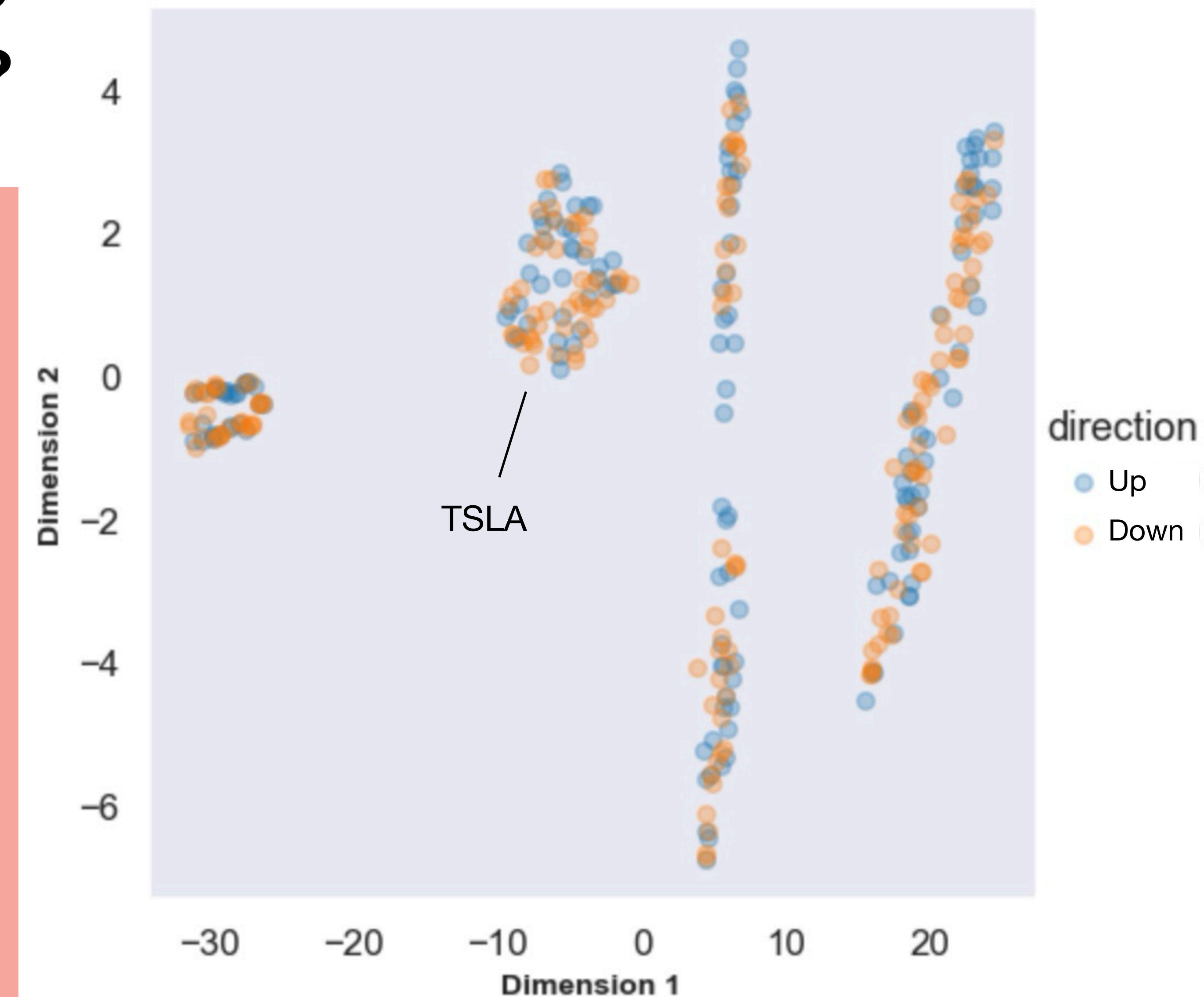


Feature Importance

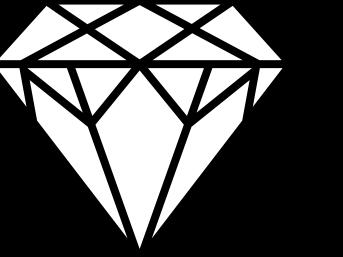
What has the greatest effect?

- Each Cluster is a Stock.
- Differences in Stock Dimension show Price movement.
- t-SNE finds points close together.

t-SNE Results: Stock Sentiment



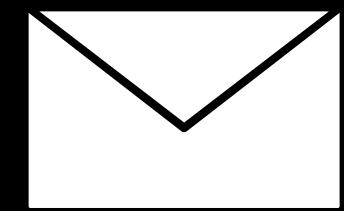
Next Steps



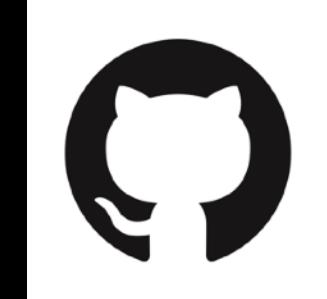
1. Collect More Data - Other Subreddits and Twitter
2. Include additional Stock Indicators
3. Use model to identify next movement trends

Thank you.

Kai Cansler



Email:
cansler.kai@gmail.com



GitHub:
<https://github.com/kai-cansler>

