



# A Sentiment Analysis of the Stock Market

Chuck Nadel



Point72



# The Business Problem

- Problem: Point72 wants to improve their systematic trading algorithms
- Task: Use Natural Language Processing to help predict whether or not the S&P 500 will increase or decrease
- Goal: Create a model to gauge retail investor sentiment



Point72



# Agenda

1. Data Overview
2. Data Exploration
3. Sentiment Analysis
4. Model Results
5. Next Steps

# Data Overview



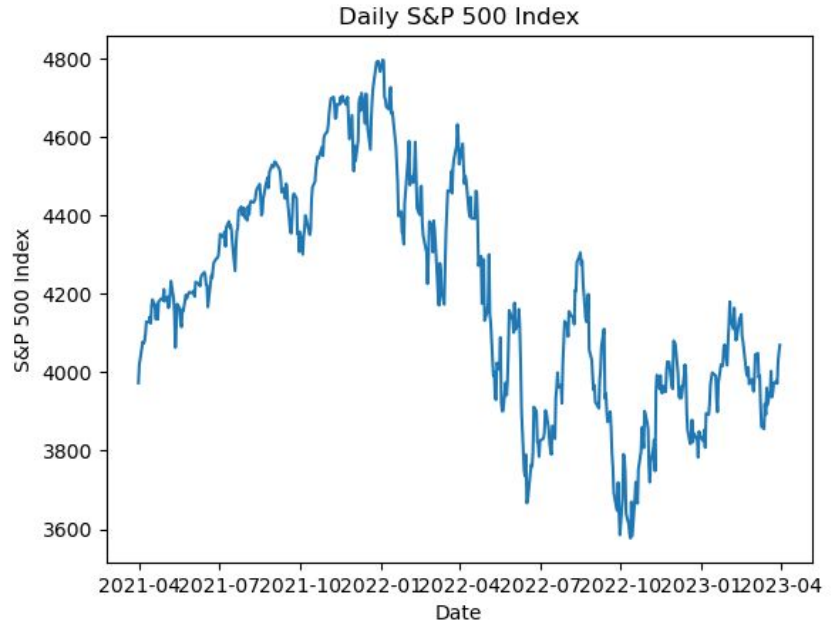
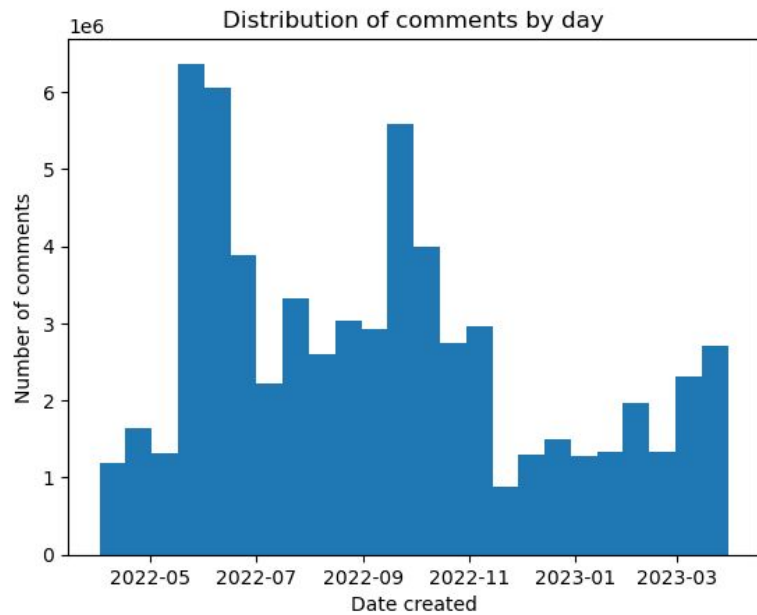
## R/Investing Data

- Scraped Reddit for the past 250 r/Investing Daily Discussion Posts
- Yielded over 100k comments for our analysis

## Yahoo! Finance Data

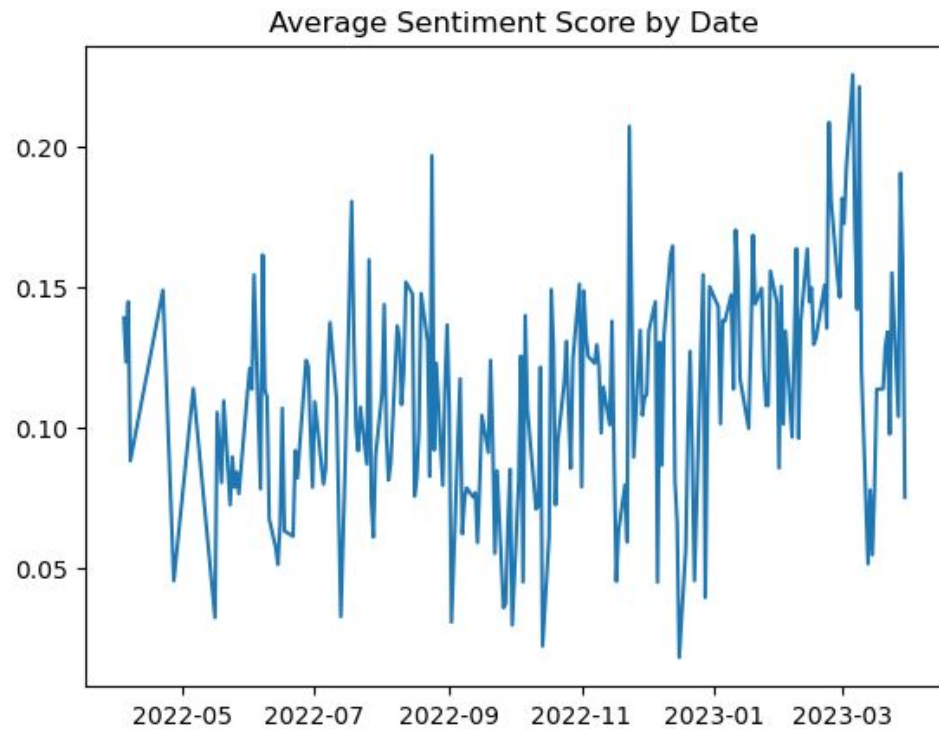
- Highly utilized financial data API with Historical Data
- Collected S&P 500 data for the dates of the discussion posts

# Data Exploration



# Natural Language Processing

Vader Sentiment Library  
Tailored for use on social media





# Model Overview

Model 1: K Nearest Neighbors

Model 2: Naive Bayes

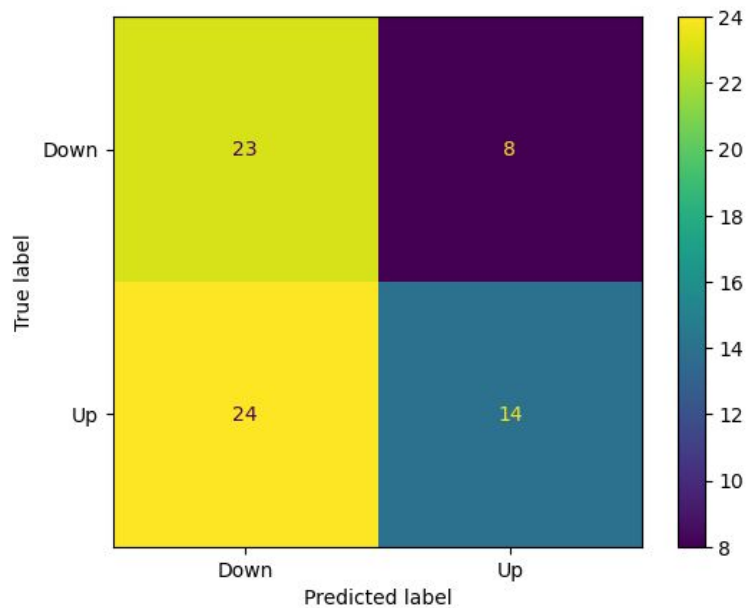
Predictive Variables:

- Number of Comments from that Day
- Average Compound Score of each comment

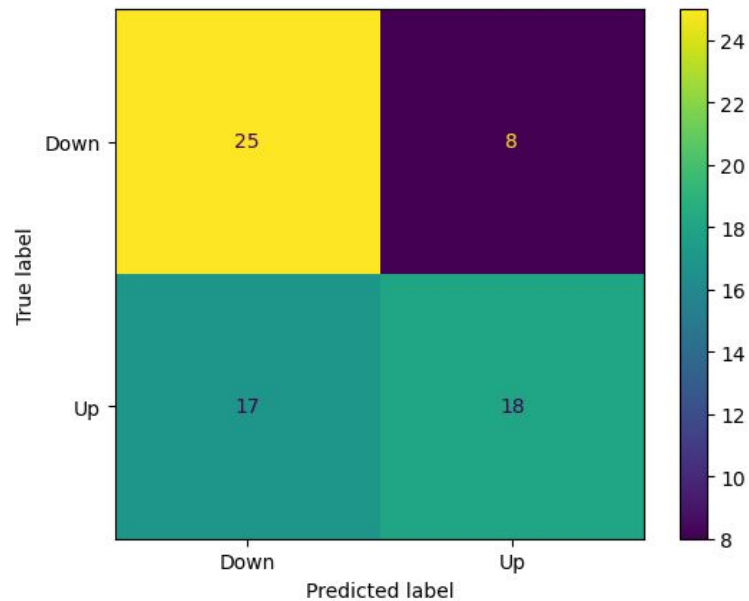
# Model Results



Naive Bayes Model



K Nearest Neighbors Model







## Next Steps...

- Stack Classification Models to improve accuracy
- Increase Sample Size- 250 days not great
- Apply to specific tickers and other websites
- Implement on top of more formal trading algorithms



# Thank You!

Email: [cnadel14@gmail.com](mailto:cnadel14@gmail.com)

Github Repository: <https://github.com/ChuckNadel/StockMarketSentimentAnalysis/>

LinkedIn: <https://www.linkedin.com/in/chuck-nadel-2ab034173/>