Sentiment Analysis with Machine Learning for Stock Predictions

•••

by

Ramgopal Gollakota, Taylor Little, Timothy Prost & Celeste Terrell

Executive Summary

Concept

- Generate stock ideas from up/down votes through social media
- Motivation for development? Looking for unique ways to use the wisdom of the crowds to make predictions
- The Meme stock craze

Concept

- We used traditional algorithmic trading algorithms, specifically Bollinger Bands, to analyze stock prices and identify potential trading opportunities based on stock price volatility. We used Random Forest Algorithm to produce classification Report
- Then we incorporated sentiment analysis using upvote scores from WallStreetBets
- We ran the Bollinger Bands analysis to identify the entry and exit points
- We also performed Random Forest predictive model and its classification report
- We compared both the approaches to see what difference we can find by incorporating sentiment analysis data

Data Techniques

- *Data sources:* Reddit's subreddit Wall Street Bets (WSB) and Alpaca Trading
- *Reasoning for data selection:* WSB is a popular forum for stock ideas.
- Used Alpaca API to get daily close and other data on AAPL stock
- *Collection, exploration and cleaning process:*
  - Web scraping on WSB to get upvote ratios and score data
  - Merge the Alpaca daily trading data with the sentiment scores from WSB
  - Overcame issues of merging two different time and date indexed dataframes
  - Had to address missing data in the sentiment analysis

Approach

- *Libraries and APIs:* Pandas, Numpy, Finta, Alpaca Trade API, Requests, PMAW, Panel
- *Breakdown of tasks and roles:* Ram - Trading and machine learning, Tim - Visualizations, Taylor - Web scraping, Celeste - Git support, Presentation
- *Challenges:* Scraping WSB subreddit for ticker symbols, dates, times, and up/down votes
- *Successes:* Figuring out how to scrape a website for data, working with indices, and getting time stamping across platforms to work!

Summary

- Stock prediction performance decreased slightly by adding sentiment analysis. The precision, recall scores were **4-5% lower**.
- The number of entry points were **3x** more with the sentiment analysis than without.
- Overall our conclusions:
  - It is possible to incorporate sentiment analysis from sites such as WSB in a trading algorithm
  - However, the resulting model may be biased toward more “buy” signals than trading “exits”

Demo

Next Steps

- *Additional topics to research:* Does a stock move because it is mentioned in social media or does it move because it is a good or bad stock?
- *Plan for future development:* Could this be a good method for selecting stocks to sell or buy puts on?

Links

- https://github.com/guitchik/Sentiment_Analysis_with_Machine_Learning

Thank You!

Sources

https://towardsdatascience.com/the-unexpected-lesson-within-a-jelly-bean-jar-1b6de9c40cca

https://www.reddit.com/r/wallstreetbets/