University of Illinois at Urbana-Champaign

CS410 – Text Information System

Fall 2021

Student: dangn2

**Tech Review**

*Can monitoring r/wallstreetbets subreddit predict a meme stock surge?*

1. **Abstract**

End of year 2020 and beginning of 2021, r/*wallstreetbets* subreddit members were orchestrated by a group of retail stock holders caused phenomenal stock rallies for two tickers GME, AMC. There was no real performance in those two companies which led to their stock surges. The phenomenon simply occurred due to great orchestration via meme in social media. There never been one or many true factors that can absolutely predict stock price trend, this tech review is purely an attempt to do so by looking into historical data of subreddit comment and stock price at the time to find the correlation. Based on that correlation, a prediction may be achieved.

1. **What available tool can make the prediction?**
2. Public stock data from Yahoo Finance API.
   1. What data the API provides for a ticker

* Date
* Open (price)
* Close
* High
* Low
* Volume

There are a few more parties which provides free/subscription-based services where we can pull stock historical data and/or live data. For example,

[Yfinance API - Realtime & Historical Data API - polygon.io](https://polygon.io/stocks)

* 1. Sample retrieval data

Text

Description automatically generated

1. Reddit API
   1. Based on an input subreddit, it can provide

* Author
* Score
* Upvote\_ratio
* num\_comments
* comments (list)
  + author
  + body
  + score
  1. Sample retrieval data

Text

Description automatically generated

1. **Implementation**
2. Obtain historical data
3. Pull historical ‘meme’ stock data.
4. Pull historical reddit comment about ‘meme’ stock
5. Study correlation between stock price and high score, related comment posted in Sub-Reddit ‘*wallstreetbet’* based on date chronologically
6. **What left to be done to improve the prediction?**

Since learn correlation between stock price and r/*wallstreetbet* subreddit by using its historical data may form a good prediction. Spotting a surge in meme stock ahead of time can be very tricky. It may require more data sources and activities to improve quality of the prediction. Some ideas as below.

* Pull data from Google Trends
* Pull data from News
* Watch r/*wallstreetbet* subreddit in real-time

1. **Sources**

* Yahoo Finance API

[*https://reasonabledeviations.com/2017/07/30/yahoo-historical-prices/*](https://reasonabledeviations.com/2017/07/30/yahoo-historical-prices/)

* Reddit API

[*https://praw.readthedocs.io/en/stable/*](https://praw.readthedocs.io/en/stable/)

[*https://towardsdatascience.com/how-to-use-the-reddit-api-in-python-5e05ddfd1e5c*](https://towardsdatascience.com/how-to-use-the-reddit-api-in-python-5e05ddfd1e5c)

* Screenshots of retrieved data is from/*credited to* team yoda

1. **References**

* Yahoo Finance API - A Complete Guide - AlgoTrading101 Blog [*https://algotrading101.com/learn/yahoo-finance-api-guide/*](https://algotrading101.com/learn/yahoo-finance-api-guide/)
* Google Trends  
  Unofficial API for Google Trends [*https://pypi.org/project/pytrends/*](https://pypi.org/project/pytrends/)
* Newscatcher  
  [*https://github.com/kotartemiy/newscatcher*](https://github.com/kotartemiy/newscatcher)