

Sentiment Analysis on Social Media

E6998 CLOUD COMPUTING

TEAM#14 Project-TA: Peter

Yin (CC4717), Rex (CL4114) Tim (SK4920), Tal (TIZ2102)

Outline

- Introduction
 - Project Purpose
 - Novelty and Value
 - □ Cloud Architecture (Completeness)
 - □ Key Results
- Live Demo
- QA Time

Product URL: https://cc-project-frontend.s3.amazonaws.com/index.html

Novelty and Value

- Analyze stock discussions cross platforms (Reddit, Twitter, Google Search Trend)
- Metrics (metric design based on literature)
 - Emotion analysis
 - Sentiment analysis
 - Popularity analysis
- User subscription for their interested stock keywords and custom thresholds for all metrics

Cloud Architecture Frontend **AWS General AWS Cloud** Gateway GET/POST Trigger every minute data fetch Google Trends Search keyword Amazon CloudWatch **Event Table** Twitter Table User Reddit Table Amazon bscription **Event** Amazon RDS Simple Emal Trigger Populate Data Notification Python Service pakage Text2emotion Web scrape + Data cleaning Trigger every hour Amazon CloudWatch Amazon Comprehend

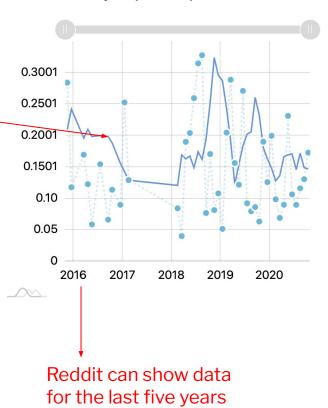
Metrics

- Data Source
 - Reddit: subreddits
 - Twitter: tags
 - Google Search Trend
- Emotion Analysis
 - Data Source: Reddit, Twitter
 - ☐ Type: Happy, Angry, Surprise, Sad, Fear
- Sentiment Analysis
 - Data Source: Reddit, Twitter
 - Type: Positive, Negative, Neutral, Mixed
- Popularity Analysis
 - Data Source: Google Search Trend

Sentiment & Emotion

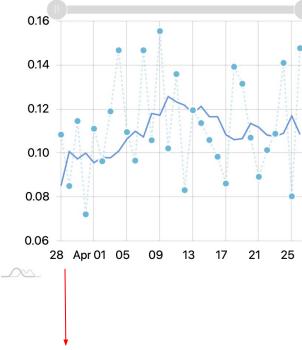
- Moving Average: (blue line) -
- ➤ If 1 ~ 12 months chart: display 7 days moving average.
- If 1 ~ 5 years chart: display 4 months moving average.

Sentiment Analysis (Positive) - Reddit



at most

Sentiment Analysis (Positive) - Twitter

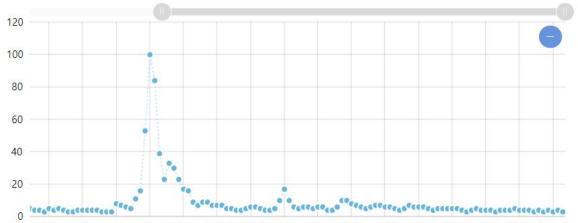


Twitter shows one month's data (due to the high volume of Twitter data)

Correlation (Gamestop)

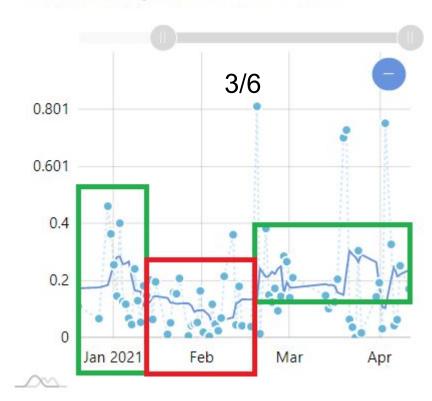


Popularity Analysis - Google Search

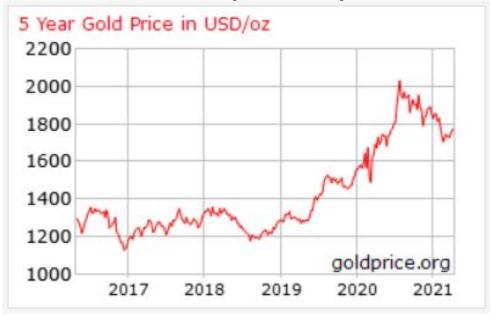


Jan 04 Jan 11 Jan 18 Jan 25 Feb 01 Feb 08 Feb 15 Feb 22 Mar 01 Mar 08 Mar 15 Mar 22 Mar 29 Apr 05 Apr 12 Apr 19

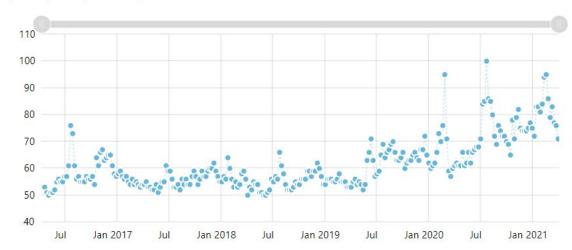
Sentiment Analysis (Positive) - Reddit



Correlation (Gold)



Popularity Analysis - Google Search



Sentiment Analysis (Positive) - Reddit

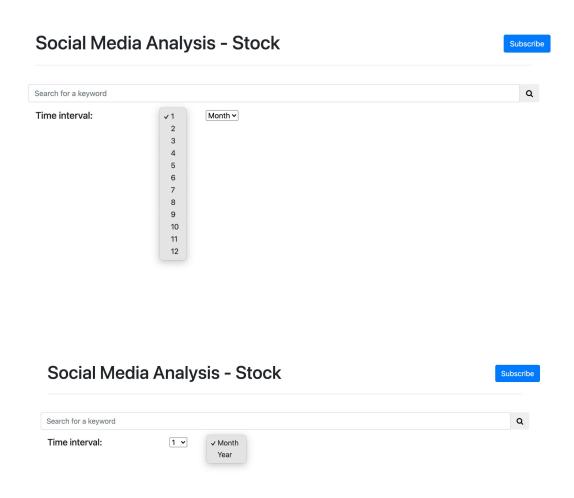


Sentiment Analysis (Negative) - Reddit



□ Front-end interface

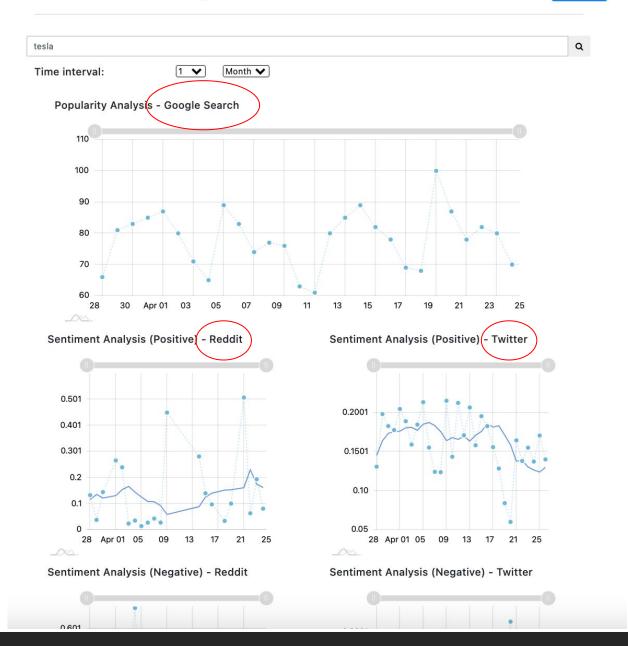
- Available keywords in database:
- 1. tesla
- 2. apple
- 3. google
- 4. amazon
- 5. yahoo
- 6. facebook
- 7. nvidia
- 8. gold
- 9. oil
- 10. GameStop



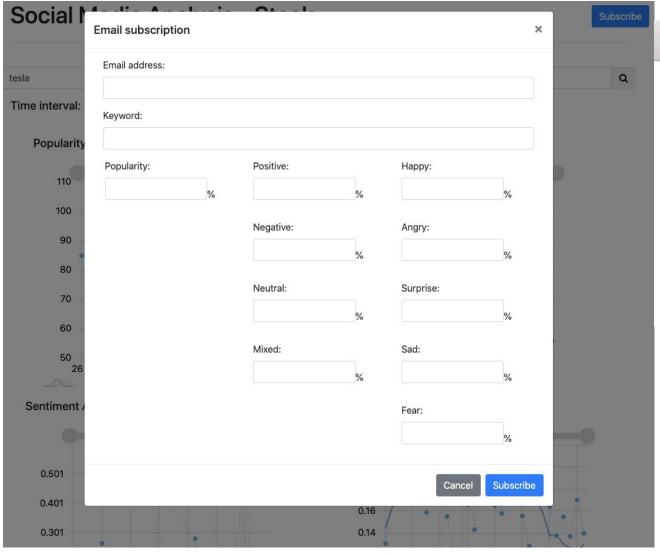
Front-end interfaceResult

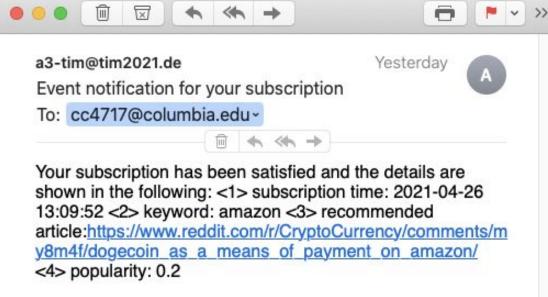
Social Media Analysis - Stock





User Subscription





Future Work/Improvements

- 1. Front End Data Display
- 2. Increase Search Pool
- 3. Noise Filtering
- 4. Linear Regression Analysis
- 5. User Trust Scores

Thank You.

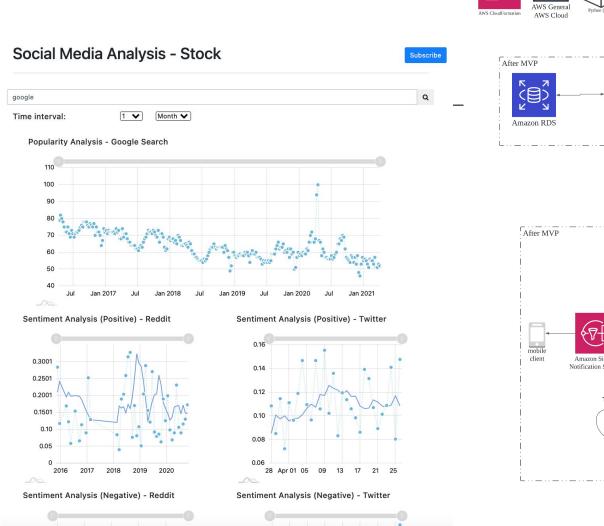
Live Demo and QA

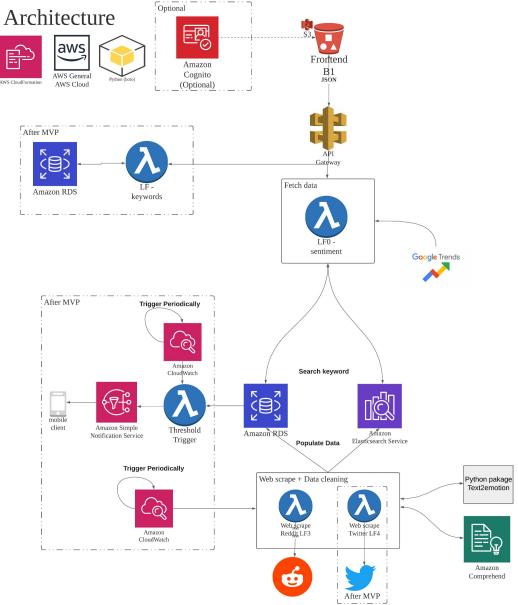
MVP slides

Team14

Metrics

- 1. Popularity
- 2. Sentiment
 - 1. Positive
 - 2. Negative
 - 3. Neural
 - 4. Mixed
- 3. Emotion
 - 1. Happy
 - 2. Angry
 - 3. Surprise
 - 4. Sad
 - 5. Fear





Demo/Service URL: https://cc-project-frontend.s3.amazonaws.com/index.html

Social Media Opinion Analysis - Stock

Tim (sk4920), Tal (tiz2102), Rex (cl4114), Yin (cc4717)

About Our Project:

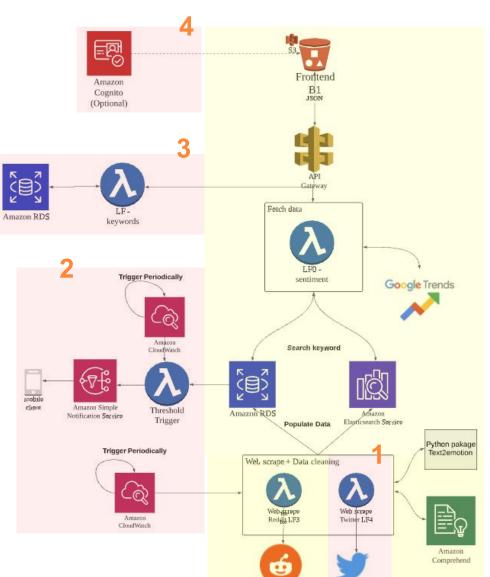
- 1. Analyze posts on social media
- 2. Metric: emotion, sentiment, and popularity analysis
- 3. Google Search, Reddit, Twitter

What We've Done So Far:

- 1. Finished frontend
- 2. Scraped from Google Search, Reddit)
- 3. Built metrics

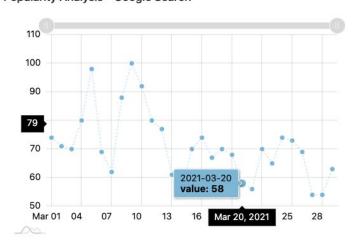
What We're Planning To Do:

- 1. Scrape tweets from Twitter
- 2. Create event trigger for scraping
- 3. Stock search history (Ranking)
- 4. Cognito user login

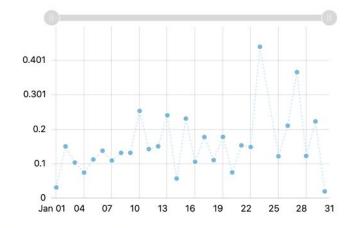


Popularity Analysis - Google Search

TESLA



Sentiment Analysis (Positive) - Reddit



Emotion Analysis (Happy) - Reddit

