twitter

Company Trends

Andy Li, Bisakh Das, Emmanuel Teferra, Georges Elizee, Karen Herrera, Season Chowdhury

Go to the page →

twitter

Purpose

- Our goal is to create a web application that users can use to find current, real-time stock and sentiment information about a company

Discover →

Features

Users will input a stock ticker of their choice, then the app will first find a context, which will be represented by some amount of most recent tweets, using the Stream API to retrieve real-time tweets including the ticker.

The app will then perform a sentiment analysis on the ingested tweets and display a pie chart showing the positive versus negative sentiment. A line graph can also be shown to display the stock data

Features cont.

The app will also show the user the selected stock's previous day information alongside the sentiment analysis graph via AlphaVantage's finance API

Stock Information returned
- Stock Price
- Change percent
- Latest trading day
- Stock Volume

The app will also display a line graph of the selected stock's recent activity to the user, giving them full access to the stock's entire behavior for the past 24hrs

Wordnet Lemmatizer and Sentiwordnet Lex.

Lemmatizer

- NLTK's lemmatizer allows the application to sort tweets based on inflected/varied forms of the same word or phrase
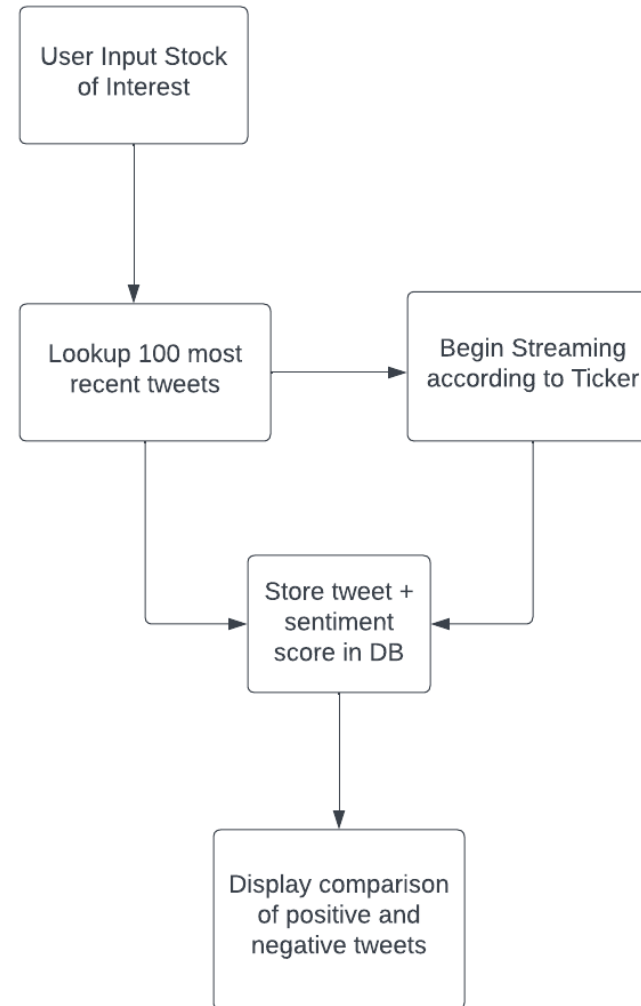
Sentiwordnet

- Another NLTK lexicon, Sentiwordnet allows us to perform more robust opinion mining on the gathered tweets while also screening out tweets that would give users a false positive
  - "$AAPL is so hot my phone's gonna crash"
  - "Hoping this $TSLA stock collapses so I can buy in! "

```python
from requests.structures import CaseInsensitiveDict
from .models import Tweet, Stock
import nltk
from nltk.stem import WordNetLemmatizer
nltk.download('wordnet')
from nltk.corpus import sentiwordnet as swn
nltk.download('sentiwordnet')


def get_sentiment_analysis_score(tweet_text):
    tt = nltk.TweetTokenizer()
    tokens = tt.tokenize(tweet_text)
    lemmatizer = WordNetLemmatizer()
    sentiments = []
    for token in range(len(tokens)):
        lemma = lemmatizer.lemmatize(tokens[token])
        synsets = list(swn.senti_synsets(lemma))
        sentiments.append([])
        if synsets:
            for synset in range(len(synsets)):
                sentiment_score = synsets[synset].pos_score() - synsets[synset].neg_score()
                sentiments[token].append(sentiment_score)
        else:
            # if the token is not in SentiWordNet, ignore it
            pass
```

Program Flow

User Input Stock of Interest

Lookup 100 most recent tweets

Begin Streaming according to Ticker

Store tweet + sentiment score in DB

Display comparison of positive and negative tweets

Live Demo

Season0727 Add component refresh   59fd735 13 hours ago   🕐 **27** commits

📁 data_mining	removed files ignored by the gitignore from the repo	4 days ago
📁 myapp	Add component refresh	13 hours ago
📁 static	Add component refresh	13 hours ago
📁 templates/myapp	Add component refresh	13 hours ago
📄 .DS_Store	Commit 4	3 weeks ago
📄 .gitignore	add gitignore	4 days ago
📄 README.md	added requirements.txt, updated readme	3 days ago
📄 manage.py	2nd commit	3 weeks ago
📄 requirements.txt	updated requirements.txt to reflect prettytable addition	14 hours ago

README.md

Company-Trend-Analyzer

Instructions

1. pull code from repository
2. pip install -r requirements.txt
3. python manage.py migrate
4. python manage.py runserver

CIS 400 Social Media and Data Mining Project

📖 Readme
⭐ 0 stars
👁 2 watching
🍴 0 forks

Report repository

Releases

No releases published

Packages

No packages published

Contributors 4

👤 ge-ogre

👤 Season0727

👤 andyLiSyr Andy Li

👤 bisakhdas