Sentiment Analysis Tool

Jeet Shah

Coleman Lyski

Tim McCabe

Advised by Dr. Ali Minai

Our Platform

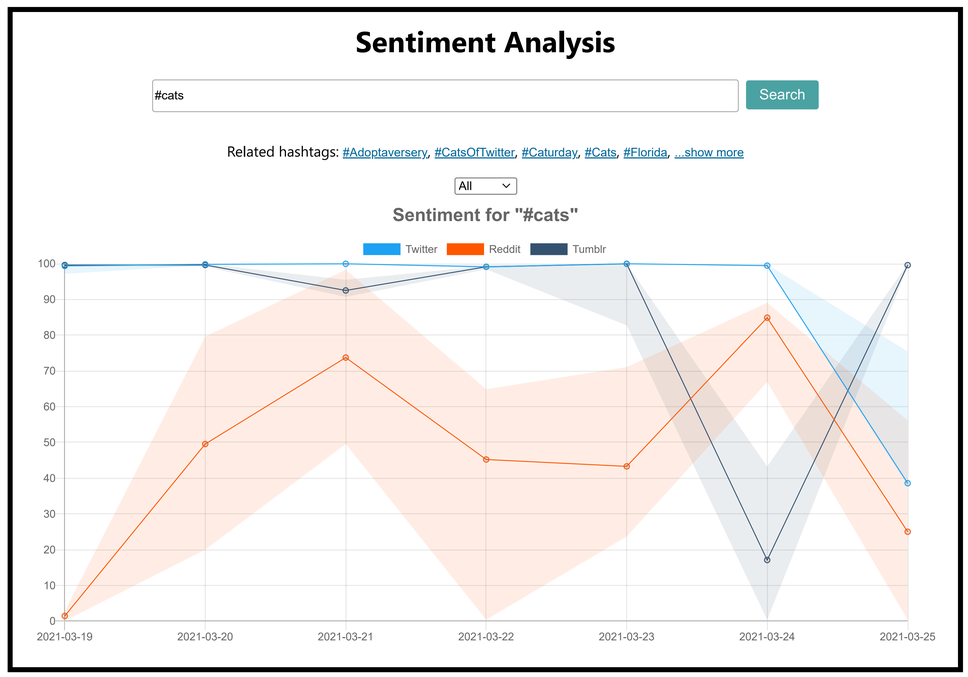
This project is a sentiment analysis and aggregation web application. Users can search for any set of keywords (e.g. a brand name or topic) and they will see a graph of public opinion regarding their search. The result data presented to the user is constructed from social media posts that are processed through a natural language processing engine. Users will see public sentiment changes over time and can break down results by social media platform.

Background

Currently, there are many existing solutions to perform Sentiment Analysis to differentiate between positive and negative text. For this project, the goal was to utilize Natural Language Processing to create our own engine that will determine the overall sentiment trend of keywords across a variety of social media. This tool can then be utilized to provide further insight into decision making for marketing executives and others in that respective industry.

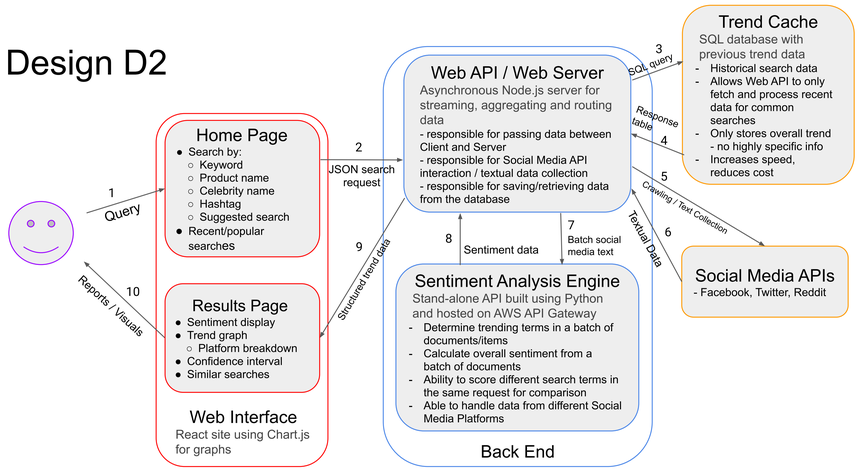
Sentiment Analysis

The methodology of the Sentiment Analysis engine used in this project is based on a very large pre-classified dataset and a BiLSTM neural network. Many different options were tested in this context, such as basic word scoring, Naive Bayes, Logistic Regression, and multiple types of neural networks. By incorporating the Universal Sentence Encoder from Google to create embeddings, the BiLSTM was able to be trained using the Sent140 dataset from Stanford, a dataset containing over 1.6 million text objects. The final classification rate for the test data was over 80% for the 100,000 records tested.



A screenshot of the tool in action

Design Diagram



Technologies

React, node, python, Flask (web development, one drop at a time), TensorFlow

Technical Design

- The React site displays the data based on the user's selections.
- The Node server manages communication throughout the application.
- The Python engine API prepares and schedules the sentiment analysis and aggregates the results.
- The Python sentiment engine assesses the positivity or negativity of social media posts.

Achievements

- Realtime, concurrent data aggregation and processing from three different social media platforms
- A robust Sentiment Analysis API that allows the analysis part of our application to run standalone without the need for a frontend
- Implemented Googles Universal Sentence Encoder along with a Recurrent Neural Network to build a classifier with nearly 80% sentiment accuracy

Future Goals

- Create a customizable date range for the sentiment graph. This will allow users to view trends over a longer period of time.
- Implement targeted sentiment analysis. This will enable our engine to assess the sentiment of text with respect to the search terms.