TweetSpeak: Delving into Public Sentiment of a Topic through Twitter Threads

Stuti Agrawal
stutia3@illinois,edu
University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

Syed Ahmad Raza
syedr3@illinois,edu
University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

Zhou Ong
czo2@illinois.edu
University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

Rutva Pandya
rutvadp2@illinois,edu
University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

ACM Reference Format:
Stuti Agrawal, Syed Ahmad Raza, Zhou Ong, and Rutva Pandya. 2024. TweetSpeak: Delving into Public Sentiment of a Topic through Twitter Threads. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

This project aims to streamline the process of distilling sentiment analysis from the vast sea of social media data, intending to provide users with a comprehensive understanding of public opinion surrounding their queries. In this project, we create an end-to-end pipeline that takes a user-provided query, return an overall sentiment related to the query and provide the top 5 positive and negative tweets relevant to the query. We use Kaggle Twitter Sentiment Analysis Dataset [1] to create a database of tweets that we use in our project. The tweets are treated as documents in the retrieval task. Throughout the project we use class concepts like MapReduce, TF-IDF with BM25 scoring, document length normalisation to preprocess the tweets and produce relevant tweets related to a query. In this project, we also compare two state-of-the-art models that are trained to perform sentiment analysis. We use IBM Watson NLU model[2] and a HuggingFace Transformers model[3] fine-tune on the sentiment analysis task. We use different evaluation metrics to compare the two models and present the results in the paper.

2 DESCRIPTION

Within the vast field of sentiment analysis, the abundance of data present in social media platforms such as Twitter offers a wealth of insights that are just waiting to be discovered.

2.1 Dataset

We have used a publicly available dataset on Kaggle[1]. This dataset has 13.2k unique tweets which have a variety of topics ranging from video games like Call of Duty to furniture stores like Hope Depot and big tech companies like Amazon and Verizon. This offers a diverse database to experiment with a range of different queries in different domains. This is a very popular dataset, and many projects have been built upon this dataset. A major advantage of the project is that it has ground truth sentiment values for each tweet which makes it easier to evaluate the results.

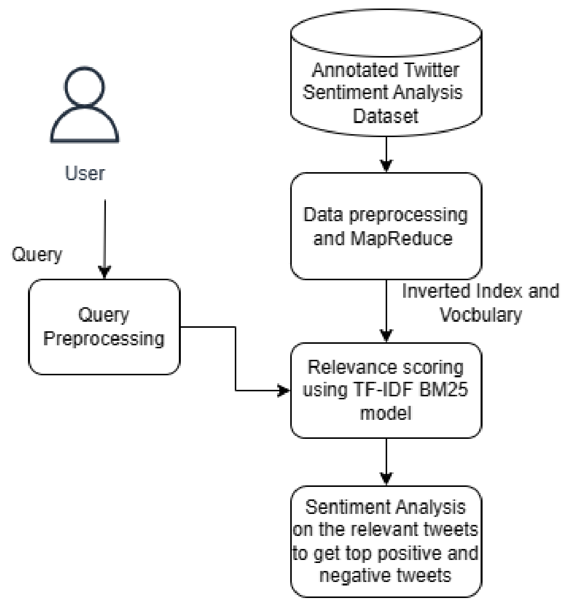

Figure 1: Project architecture

2.2 Data Preprocessing and Retrieval

The pipeline begins with preprocessing the data to make future operations more efficient. We employ similar strategies to what we have used previously in class and in the homework. We begin by removing stop words, punctuations, numbers, and special characters, stemming the tweets. We use Hadoop for implementing the MapReduce algorithm, similar to the assignments we worked on previously. We created an inverted index using the mapper and reducer script, setting the stage for effective document retrieval.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

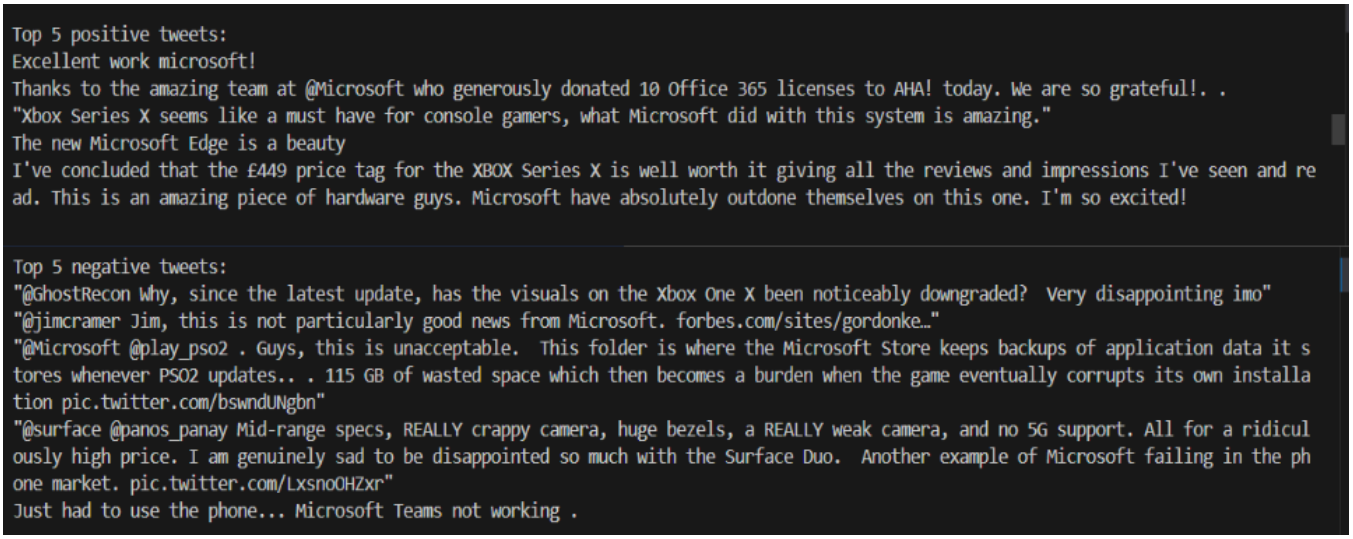

Figure 2: This figure shows the top 5 positive and negative tweets related to the query: Microsoft Technology

This inverted index records all the unique words and the documents they appear in and ensures faster access to all the words in the documents. This fundamental stage improves the overall efficiency of our sentiment analysis pipeline by simplifying the process of finding relevant tweets for user queries. Once a user enters a query, we preprocess the query to prepare it to use TF-IDF relevance mode with BM25 and document length normalization to get relevant tweets. This script takes 0.0672 seconds on average to retrieve relevant tweets once a user enters a query. The Fig.2 shows that the user is prompted to enter a query.

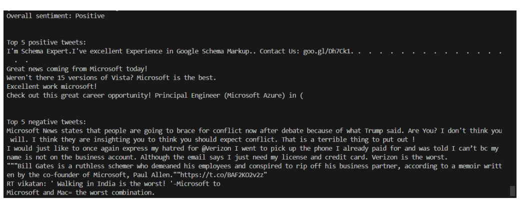

Figure 3: Entering a query to get a list of relevant tweets

2.3 Sentiment Analysis

After retrieving relevant tweets, we use two different sentiment analysis models to get the overall sentiment and the relevant tweets. We use IBM Watson NLU model fine-tuned on the sentiment analysis task to get the sentiment of each tweet and then we compare the weighted sum of the sentiment values to get the overall sentiment using confidence scores returned by the model as weights. The IBM Watson model labels the tweet as Positive, Negative or Neutral which is the same as the dataset. We use the free tier of the model that has limited compute units per month. Another model we use is the HuggingFace Transformers library which has a model that is fine-tuned on the sentiment analysis task and pretrained for natural language processing tasks. This model labels tweets only as Positive or Negative. It is an open-source model and is free to use.

Essentially, our comprehensive method yields the extraction and display of the top five most favorable and unfavorable tweets along with a general evaluation of the sentiment polarity of the query. Figure 3 shows an example of the output.

The image above provides an overall sentiment about the query entered. Along with this, the top 5 positive and negative tweets are printed to give the user an insight into the context of the tweets that are influencing the overall sentiment score.

3 EVALUATION

We use a variety of metrics to compare the two models. We use Sklearn to get the F1-Score, Precision, Recall for the two models. We also compare the response times for the two models to get the sentiment labels. We make use of the ground truth labels present in the dataset to calculate the F1-Score, accuracy and precision. Table 1 and Table 2 show the results. We curate 3 queries based on popular topics in the dataset. We then get the evaluation metrics for each query and average them over the 3 queries to get the evaluation numbers.

4 DISCUSSION

In this section, we will interpret our results. As the numbers in the table suggest IBM Watson has a much higher accuracy as compared to the Transformers model. One reason for this can be that the IBM Watson model has more sentiment labels as compared to the Transformers model. Based on accuracy, we would choose the IBM Watson model for our project pipeline. In terms of the average response time of the model, Transformers is a lot faster about 4 times faster with an average response time of 0.097 seconds per tweet making it more desirable when the latency is concerned. Additionally, it does not have any API rate limits and is free to use for an unlimited number of tweets and queries. The IBM Watson model has a restriction of 20000 NLU compute units in the free tier making it a more expensive option. However, for our project,

Model	F1 Score	Precision	Recall
IBM Watson NLU	**0.587**	0.575	0.599
Transformers	0.472	0.405	0.564

Table 1: Comparison of sentiment analysis models

Model	Average Response Time	Max Response Time	Min Response Time	Median Response Time
IBM Watson NLU	0.375	3.135	0.284	0.343
Transformers	0.0971	5.885	0.0252	0.068

Table 2: Comparison of response times for sentiment analysis models

we prioritize accuracy over latency and cost. Assuming unlimited resources, we would choose IBM Watson for our project.

5 SUMMARY

In order to gain insight into sentiment trends on Twitter, we experimented with different strategies to perform sentiment analysis on tweets to get an overall sentiment on a particular topic. We used different technologies, and models to accomplish our goal. Our results revealed the pros and cons of both models. We concluded by deciding to use the IBM Watson NLU model instead of the HuggingFace transformers model as it had better accuracy. In the future, we also plan to extend this project to use real-time data using some sort of Twitter API to experiment in a more real-world setting. Below is a link to the GitHub repository documenting our progress. The instructions for running it can be found in the README file. TweetSpeak GitHub Repository.

REFERENCES

[1] [n. d.]. Sentiment Analysis - Twitter Dataset — kaggle.com. https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis. [Accessed 16-02-2024].

[2] IBM Corporation. 2022. IBM Watson Natural Language Understanding. https://www.ibm.com/products/natural-language-understanding.

[3] Thomas Wolf, Julien Chaumond, Alexandre Defossez, Victor Sanh, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Hendrik Brew, Jens Hawkins, Patrick von Platen, Sharan Ma, Yacine Jernite, Julien Plu, Alex Casale, Mohammad A Patwary, Zsolt Kiss, and Inc. Hugging Face. 2019. Transformers: State-of-the-art Natural Language Processing. https://huggingface.co/docs/transformers/en/index.