Sentiment Analysis of COVID-19 Vaccine Tweets

Jeerthi Kannan, *jeerthi.m.kannan@vanderbilt.edu*
Tabitha S.Y. Lee, *tabitha.see.ya.lee@vanderbilt.edu*

I. Introduction

Sentiment analysis is used to identify the general trend of opinion across different social media platforms, reviews, and other text that contains public thought. COVID-19 has had a widespread effect on healthcare and society as we knew it. In particular, the research efforts and development of a COVID-19 vaccine have led to polarizing sides across many countries worldwide. Thus, we wanted to conduct sentiment analysis on tweets concerning the COVID-19 vaccine, for the purpose of seeing which sentiment is dominant across tweets over time. Results from sentiment analysis could allow us to see how opinions on the COVID-19 vaccine changed from the time before the vaccine was produced, to when it was in limited supply, to its second dose distribution, to now when doses and boosters of the vaccine have become more publicly available. Through seeing the sentiments about the vaccine over time, we could also see if there are correlations to certain events which we can hypothesize to have influenced public opinion. Originally, we wanted to scrape tweets from 2020-2022 using the search endpoint in the Twitter API. However, the Developer Account we were using did not have the necessary permissions to scrape tweets over a few years. As we did not have the necessary items to gain these permissions, we decided to use a publicly available dataset instead. Our work contributes multiple methodologies for performing sentiment analysis, while providing a basis for extracting tweets using only tweet IDs [1].

Our code is available at https://github.com/tabithasylee/CovidVaccinesSentimentAnalysis.

II. Methodology

A. Obtaining the Dataset

We used a dataset from IEEE Dataport, which has tweet IDs related to the coronavirus from 2020 to present day [2]. It is updated every day. The tweets were collected using the keywords relevant to the coronavirus, which are available on the dataset webpage. We used a subset of the dataset spanning 04/18/2020 - 09/01/2022. We chose this timeframe, as the author utilized consistent keywords to collect tweets within this time. This time period is also before the Twitter buyout and takeover by Elon Musk, thus reducing any unnecessary bias. The dataset consists of only tweet IDs, due to the owner following Twitter policy which restricts sharing of Twitter data other than IDs. To get the information from each associated tweet, we had to hydrate the tweet IDs. This means to use the tweet IDs to find tweets through the lookup endpoint in the Twitter API, and retrieve the relevant information [1]. For our purposes, we used an Elevated Developer account. We aimed to extract 12,000,000 tweets out of our subset. 13.8k tweet IDs were randomly chosen from each CSV file that comprised of the subset we used. This was so that we could stay within the limits of our available computing resources. We extracted the date and time the tweet was created, the text within the tweet, number of followers the user has, number of retweets, number of favorites, the user's location, and the tweet's hashtags.

To find only tweets that are relevant to the vaccine, we filtered out the tweets with text and hashtags that had the keywords 'vaccine', 'vaccines', 'vax', 'booster', 'dose', 'pfizer', 'moderna', 'johnson', 'j&j'. This left us with about 1.3 million tweets.

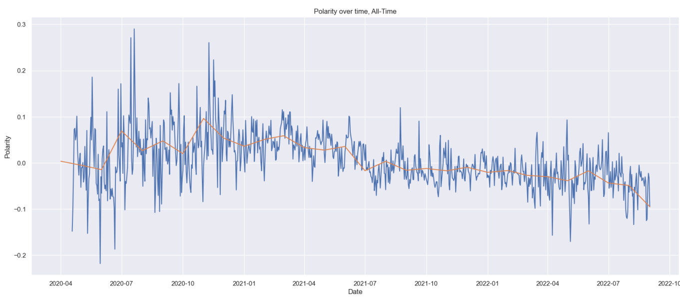

Fig. 1. Polarity of Tweets from April 2020 to September 2022

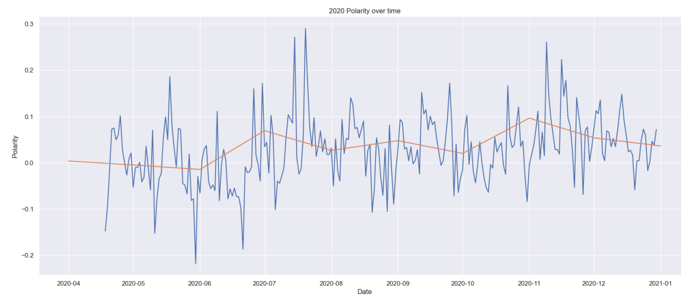

Fig. 2. Polarity of Tweets from April 2020 to December 2021

B. Solutions

1) Sentiment Analyzer: Since our dataset was an unlabeled dataset, our first step for analysis was to identify a natural processing method that would allow us to efficiently analyze the results for such a large dataset. As such, we had two options, category wise, for sentiment analysis techniques: lexicon-based methods, or unsupervised learning methods such as K-means clustering or K-nearest neighbors. Given that our dataset was at the millions-scale, we decided on a lexicon-based method due to the time that training an unsupervised method would likely take.

We decided to choose VADER (Valence Aware Dictionary for sEntiment Reasoning), a rule-based lexicon method for sentiment analysis, which requires no data, does not require more time or computational power to be effective, and most importantly of all, performs well on corpus derived from social media [3]. VADER's lexicon is derived from a number of well known sentiment dictionaries, and further analyzed and validated by human readers. Something important that we noted is that VADER also has sentiment scores for emojis, which are very frequently present in tweets. We can utilize VADER through Python, by installing NTLK's package.

After VADER determines the positive and negative sentiment score for each tweet, the scores are then normalized from 0 to 1 for positive scores and -1 to 0 for negative scores. Those two scores are then combined to give us an overall polarity score, which we can then classify into negative, neutral, and positive if the sentiment scores fit the ranges of [-1.0, 0.05), [-0.05, 0.05], (0.05, 1.0] respectively.

2) Neural Network: For proof of concept, we created a neural network using the Keras package. With the sentiments having now been identified by VADER, we used the polarity labels to create a multiclass classifier. The original dataset with sentiments was split into 80% training and 20% testing. The three classes used by the model were 'negative', 'neutral', and 'positive'. The neural network had an embedding layer with 5000 tokens and 20 dimensions, 5 LSTM hidden layers, and a softmax dense output layer with 3 classes. When training over 20 epochs with a batch size of 2000, the model gave an accuracy of 87%.

III. Results & Evaluation

A. Sentiment Analysis Polarity Results

We compared the polarity against the features of our dataset to examine any potential correlations to the number of followers, number of retweets of the tweet, number of favorites of the tweet, and the number of hashtags that the tweet had. While we also obtained the location of the user that tweeted, since many users do not set their location to a standardized format, or even a string that is a location at all, we discarded that feature.

Furthermore, considering the size of the dataset, all the charts below are binned with regards to the average polarity of the tweet either per day, or per 1k and 10k of retweets and favorites respectively. With regards to concerns that heavily polarized tweets will cancel out their lesser-polarized counterparts, we have taken steps to prevent any outliers from skewing the binned data too much by normalizing the sentiment/polarity scores to a range of (-1, 1) before taking the average.

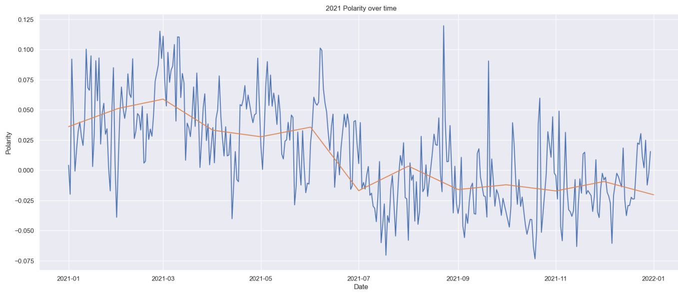

Fig. 3. Polarity of Tweets from January 2021 to December 2021

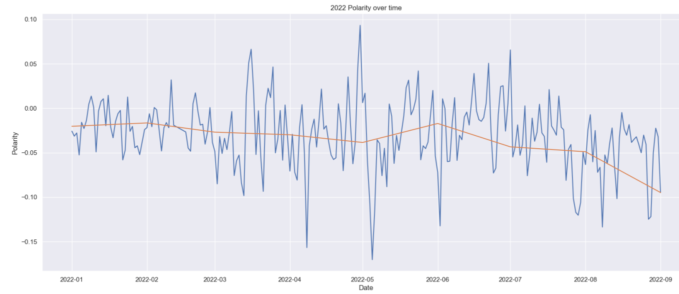

Fig. 4. Polarity of Tweets from January 2022 to September 2022

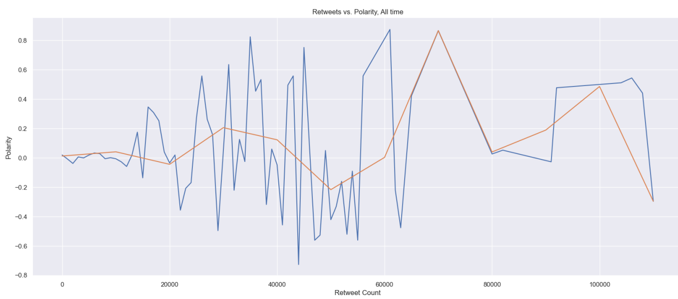

Fig. 5. Polarity of Tweets vs. Retweets from Apr. 2020 to Sep. 2022

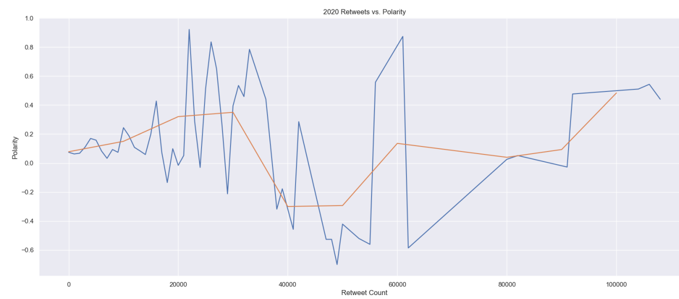

Fig. 6. Polarity of Tweets vs. Retweets from Apr. 2020 to Dec. 2021

We first examine the polarity over time. We visualize these graphs by taking the average polarity of tweets per day, which are plotted in blue, and the average polarity of tweets per week, which are plotted in orange. In Figure 1, we can see that the polarity towards COVID-19 vaccines, overall, are more extreme in either direction, prior to when the vaccines were released, from April 2020 to January 2021. Such spikes reduced from then on, which most tweets staying in the neutral range for the rest of 2021, before a slight downward trend in 2022. In order to get a better picture, we further divide this change over time into each year. In Figure 2, we can see that there is actually a slight upwards trend over time into the end of the year, from what would be considered in the neutral range into something that is more in the positive range. In Figure 3, we can also see that the ranges for the average polarity by day are much smaller than in Figure 2; where in 2020, we previously had a range of around (-0.2, 0.3) for the average polarity, in 2021, the range narrows to around (-0.075, 0.125). Performing t-tests with a threshold of $p < 0.01$ indicate that while there is no statistically significant difference in the polarity from 2020 to 2021, there is indeed a statistically significant decrease in average polarity from 2020 to 2022.

There appears to be a more downwards trend from slightly positively neutral into slightly negatively neutral overall. We could hypothesize these changes in 2020 as to more people warming up to the idea of COVID-19 vaccines as the death tolls and symptoms became more severe, and in 2021, the downwards trend could be attributed to the displeasure in required vaccine mandates as well as the discovery that a booster was needed to retain immunity. We can see in Figure 4 that there is actually a further downward trend of the average polarity that dips lower than 2021, potentially again due to the further announcement of another COVID-19 booster being necessary.

When examining the polarity of tweets versus the amount of retweets they have (Figure 9, we can actually see that more popular tweets tend to be more polarized than less popular tweets. While, naturally, there are more tweets from the (0,10,000) retweets range than the (10,000, 90,000) retweets range, we should keep that in mind as we examine the data, since there are less datapoints, an outlier might have more effect on the average for the higher ranges. Nevertheless, looking at the average polarity for tweets from (0, 10,000) even compared to the (10,000, 20,000) range still indicates an increase in polarity. While it might be too early to jump to the conclusion that more popular tweets tend to be more polarized,

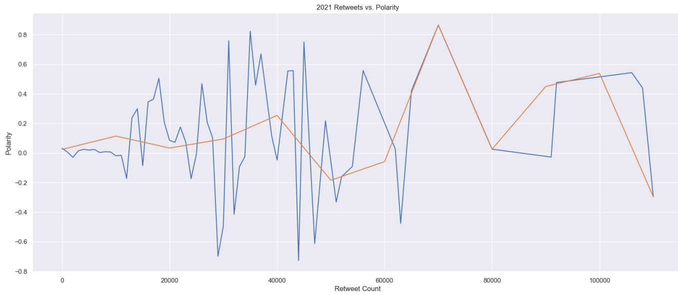

Fig. 7. Polarity of Tweets vs. Retweets from Jan. 2021 to Dec. 2021

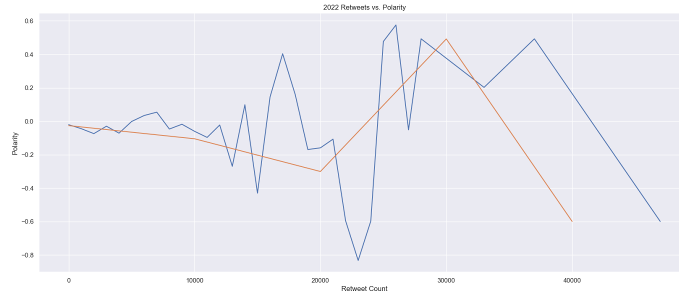

Fig. 8. Polarity of Tweets vs. Retweets from Jan. 2022 to Sep. 2022

Fig. 9. Polarity of Tweets vs. Favorites from Apr. 2020 to Sep. 2022

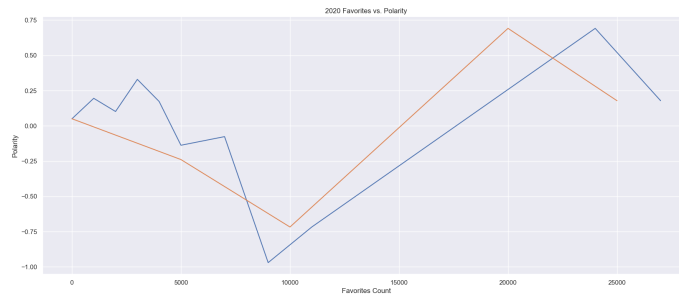

Fig. 10. Polarity of Tweets vs. Favorites from Apr. 2020 to Dec. 2021

considering that the Pearson correlation coefficient for retweets and overall polarity is 0.02, it is still of interest to note that polarized tweets *can* become extremely popular. While we do divide the data into subgraphs based on the year (see Figure 6, Figure 7, Figure 8), this relationship does not seem to change overall. However, the more popular tweets tend to have more negative sentiments in 2021 and 2022, following our analysis earlier that there is a drop in sentiment towards COVID-19 vaccines in 2021 and 2022. When we perform a similar analysis of tweets by favorites, however, even looking at the average of tweets with less favorites, they do still get rather polarized starting at 5,000 favorites in 2020, 7,000 favorites in 2021, and as early as 2,000 favorites in 2022. This does lend more evidence to the our potential hypothesis that more popular tweets tend to favor a particular sentiment and draw engagement. However, the overall correlation coefficient for polarity and favorites remain extremely low at 0.0004, which disproves that as well.

Next, considering that hashtags are frequently used to give tweets more visibility as well as target them towards a specific audience, we decided to examine if any correlation could be found. Considering that the polarity stayed in the range of (-0.05, 0.15), this seems unlikely, although tweets with more hashtags did have more positive sentiment than less hashtags, this range is still rather small. Furthermore, the Pearson correlation coefficient for these two variables is 0.03, indicating a lack of strong relationship.

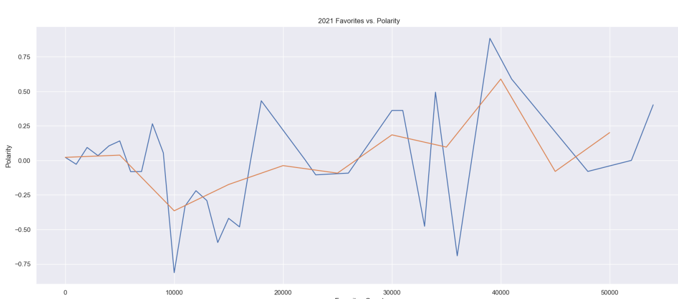

Fig. 11. Polarity of Tweets vs. Favorites from Jan. 2021 to Dec. 2021

Fig. 12. Polarity of Tweets vs. Favorites from Jan. 2022 to Sep. 2022

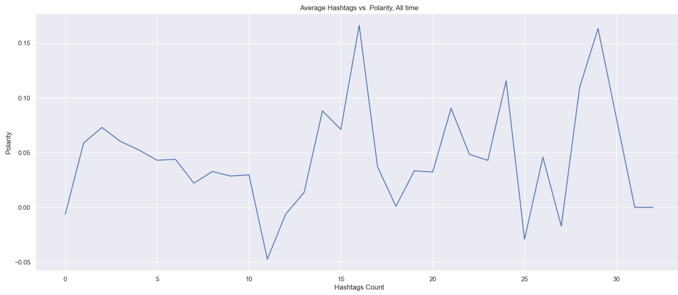

Fig. 13. Polarity vs. Number of Hashtags

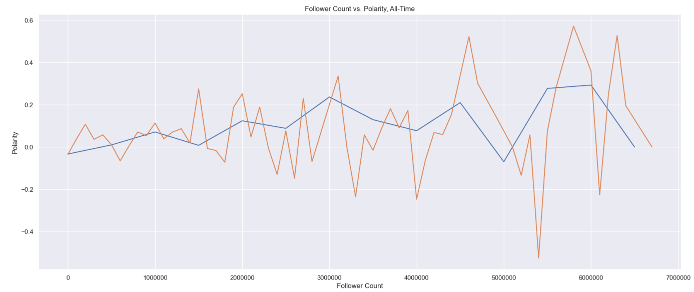

Fig. 14. Polarity of Tweets vs. Followers Count of the User

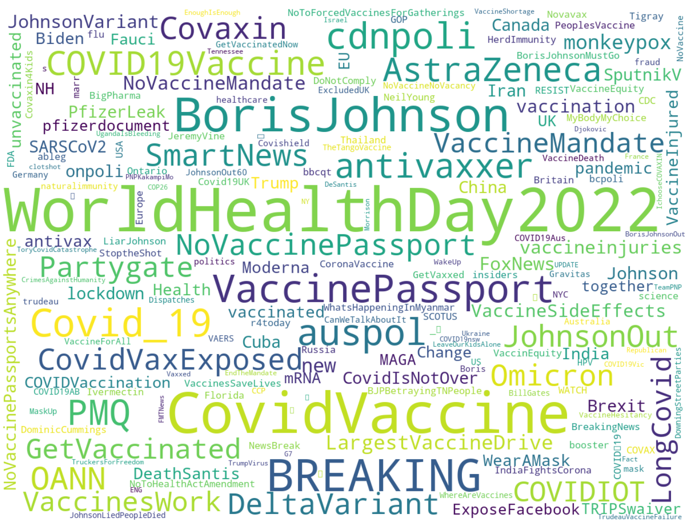

Fig. 15. Top Hashtags from Negative Tweets

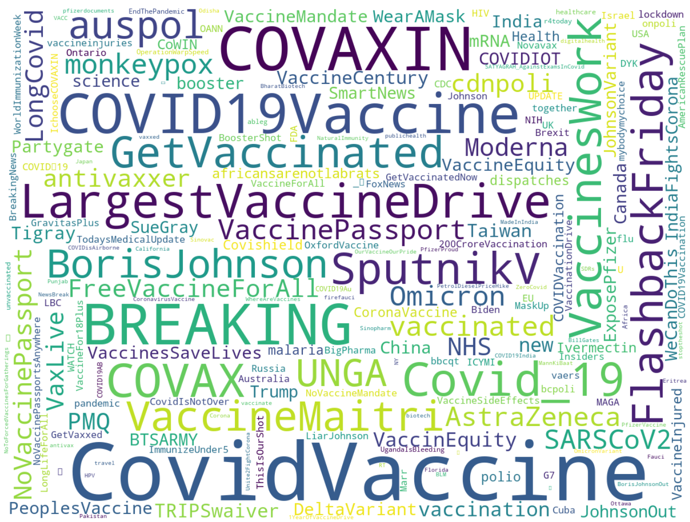

Fig. 16. Top Hashtags from Positive Tweets

Finally, we were curious to see if more popular users tended to post more polarized opinions. Again, it is still important to note that since there are more users with less followers, the average is likely to remain more neutral as a result. This is further proven by the correlation coefficient being 0.006. Despite this, we can still see a similar relationship on the extreme higher end of follower counts with the follower count of users to the amount of retweets/favorites - users who post polarized tweets indeed are able to obtain followers.

B. Hashtag WordClouds

With the hashtags from the tweets, wordclouds were constructed to show most used hashtags associated with negative and positive tweets. Common hashtags were found by making lists of the first 10 commonly occurring hashtags for both positive and negative tweets, and then comparing them to see which ones both lists share. The most common hashtags shared among both lists were removed to produce wordclouds with more unique results and further identify differences between the two. Figure 15 shows more hashtags related to political figures, while Figure 16 shows more hashtags related to organizations and movements.

C. Neural Network Model Evaluation

As shown in Table I, precision, recall, and F1-scores across all three labels ranged from 84-90%. The network was trained with just the full text and polarity labels. Yet the performance can be improved through adding more data, changing the architecture of the network and adding additional features like number of retweets and number of favorites. For future directions, we can increase the amount of tweets

Labels	Precision	Recall	F1-Score
Negative	0.8492	0.8847	0.8666
Neutral	0.9085	0.8492	0.8778
Positive	0.8638	0.8715	0.8676

TABLE I
NEURAL NETWORK MODEL EVALUATION RESULTS

we hydrate, add additional features when training the neural network, and experiment with other neural network architectures like RNN and CNN.

IV. CONCLUSION

Through utilizing VADER, we were able to identify sentiments across tweets from early 2020 to late 2022. Through hydrating, filtering, and formatting tweets, we created a dataset of 1.3 million tweets which we used for sentiment analysis. From our results, we saw that polarity significantly decreased over time. There were stabilizing and downward changes correlated with acceptance of vaccines and vaccine mandates, respectively. More popular tweets tend to be more polarized than less popular tweets. The amount of hashtags on tweets did not seem to have any significant effect on the sentiment of a tweet. From the users we have in our dataset, those who post more polarized tweets tend to have more followers. Some hashtags were shared between positive and negative tweets, but the wordclouds show distinct differences. The neural network we trained only classified polarity of tweets based on their text, but it can be improved when given more data, changing its architecture, and adding additional features like number of retweets and favorites. Due to our available computing resources, we only analyzed 1.3 million tweets, but we could analyze more given more resources. There may be differences or similarities when analyzing data at the billion-scale compared to a million. In the future, the performance of our neural network could be compared with an RNN or CNN to see which architecture may be better suited for sentiment analysis of tweets.

V. ANSWERS TO QUESTIONS DURING PRESENTATION

The wordclouds we used during our presentation did not show significant differences between them due to sharing common hashtags. To address this issue, we removed common hashtags and re-made the wordclouds as stated in section III-B. Compared to our neural network, an RNN, and CNN, we think CNN may perform the best, due to its performance with other known sentiment analysis applications. CNNs are fast and can quickly learn to classify sentences, compared to RNNs. This comparison is outside the scope of our current work, and can be addressed with future efforts.

VI. TEAM MEMBER CONTRIBUTIONS

Both Jeerthi and Tabitha worked on obtaining and hydrating the COVID-19 tweet dataset together. Jeerthi worked on filtering the tweets with vaccine keywords, formatting the resulting dataset, conducting sentiment analysis of the hydrated dataset, creating the neural network model, and training the model using our dataset with VADER sentiment labels. Tabitha worked on the data visualization, statistical analysis of results, and the demo showing live hydration and sentiment classification of a tweet given a tweet ID.

REFERENCES

[1] R. Lamsal, "Hydrating tweet ids," https://theneuralblog.com/hydrating-tweet-ids/, 2021, accessed: 2022-11-28.
[2] ——, "Coronavirus (covid-19) tweets dataset," 2020. [Online]. Available: https://dx.doi.org/10.21227/781w-ef42
[3] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," 01 2015.