Sentiment Analysis of "#COVID-19" Tweets

Nakul Rao[1] and Aidan Carr[1]
[1] *United States Military Academy at West Point*

SARS-COV-2, more colloquially known as COVID-19, was declared a pandemic by the World Health Organization (WHO) on 11 March 2020. Subsequently, the virus has since continued to infect over 200 million people worldwide. The virus has consequently been in the 24-hour news-cycle for the past year-and-a-half. This paper provides insight into the effect that the constant reinforcement of COVID-19 has on the overall perception of the virus through the analysis of tweets related to COVID-19. Our team analyzed 1.27 million tweets and concluded that there was a negative association between the date and time a COVID-19 Tweet was created and the overall sentiment score of tweets containing "#COVID-19." Our analysis further suggests that on days with more tweets pertaining to COVID-19 there is a stronger negative association between date and sentiment score. Prior sentiment analysis research has already been conducted on COVID-19 through deep-learning, however these analyses were conducted in late 2020 and early 2021[1, 2]. There have been future developments since prior research was published such as the deployment of multiple vaccines and the spreading of the Delta Variant. Our team sought to add to the current research by providing an updated verification on the interaction between COVID-19 in the news cycle and the negative sentiment of COVID-19 Tweets. While our analysis and results had limitations which will be discussed later, our team firmly believes that there is sufficient evidence to warrant extensive future research. Note: "Sentiment score" refers to how postive or negative a tweet is and will be elaborated upon in Section II.

I. BACKGROUND AND SIGNIFICANCE

Given the nature of COVID-19 the news surrounding it has been largely negative. Moreover, the rise in cases and imposition of lockdowns worldwide has lead to an increase in anxiety and depression among people of all ages. This association between the mental and physical health of the global population is directly proportional to the current pandemic[1]. Our team believed that the recent developments of COVID-19 in the news has lead to the continuation of this trend.

Social Media has emerged as platforms that allow for the near instantaneous sharing of information regardless of physical distance. Twitter® currently has hundreds of millions of users across the globe and as such makes it a valid platform to test our hypothesis.

The basis for our research lies in the research of Koyel Chakraborty et al. [1] and Furqan Rustam et al. [2]. Both teams analyzed COVID-19 through social media using deep learning in late 2020 and early 2021. Chakraborty et al. [1] analyzed the sentiment of two different types of tweets and Rustam et al. [2] analyzed tweets as negative, positive, or neutral. Our team's research builds off of both of these by providing a foundation for similar research with the current conditions of the pandemic. Thus this paper analyzes how the constant reinforcement of COVID-19 in the news has impacted the overall sentiment of tweets pertaining to COVID-19.

II. INVESTIGATORS

Our research team is made up of **CADET Nakul Rao, US ARMY** and **CADET Aidan Carr, US ARMY**. Nakul Rao is a Yearline (Sophomore) at the United States Military Academy at West Point. He is pursuing two degrees in Computer Science and Mathematics and is involved in research with the Physics and Nuclear Engineering Department concerning effectively simulating the effects of Nuclear Weapons. Aidan Carr is also a Yearling at the United States Military Academy at West Point. He is pursuing a degree in Organizational Psychology and intends to seek out research in motivational psychology with the Behavioral Sciences and Leadership Department.

III. METHODS

Our research sought to analyze aggregated tweets through the Bing and AFINN Lexicons. The Bing Lexicon is a general purpose English Sentiment lexicon that categorizes words as either positive or negative. The AFINN Lexicon is a Lexicon that is rated manually, with words receiving an integer score between -5 (negative) and 5 (positive). The AFINN Lexicon was also designed specifically for microblogs like tweets[1].

From the outset, our team aimed to score each tweet in order to determine how positive or negative each was based on the date it was created, so the AFINN Lexicon was a natural choice. We introduced the BING Lexicon as a secondary sentiment scale in order to provide an additional perspective and verification to our findings.

We pulled our tweets directly from the Twitter® API between the dates of 30 October 2021 and 08 November 2021. A majority of our tweets are from between the dates of 07 Nov 2021 and 08 Nov 2021. The reasoning for this will be elaborated upon in Section IV. We wrote 2 scripts in the programming language R to accomplish this: 1) pull and clean the tweets. 2) Parse and analyze

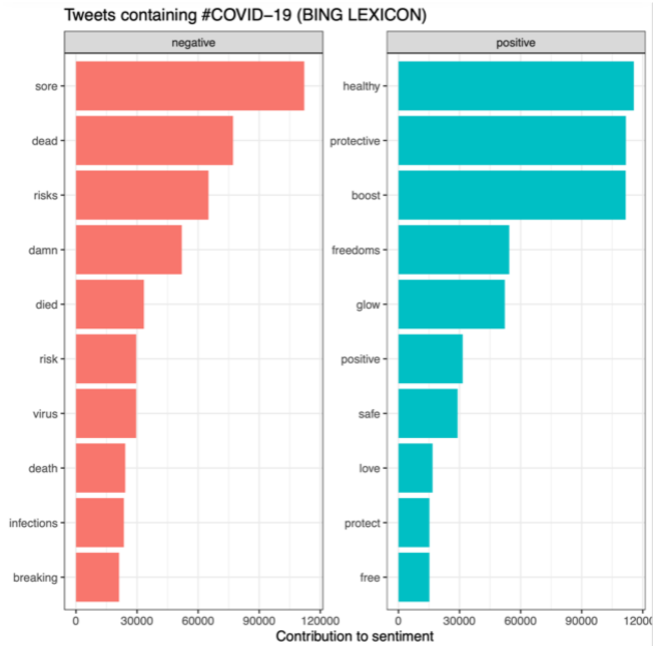

FIG. 1. Most frequent Negative and Positive Words

the cleaned tweets.

A. Pulling and Cleaning Tweets

We utilized the package rtweet to access the Twitter® API and pull 5,000 tweets every 15 minutes. Note: This rate was chosen due to the constraints that were placed on us by the API with respect to the number of tweets that we were allowed to access within a 15-minute time frame.

As we pulled tweets we used Natural Language Text Processing packages such as tidytext and dplyr to clean our data of characters that were unable to be analyzed such as URLs, emoticons, and any special Unicode characters. The script would run automatically every 15 minutes and store the cleaned tweets in a separate Comma Separated Value (CSV) file. Note: Due to the cleaning of the data that our team had to do, we were unable to extract any potential categorical confounding variables.

B. Parsing and Analyzing Tweets

Once tweets were cleaned we pulled them and analyzed them through each of the Lexicons. Each tweet was parsed, and the most frequent negative and positive words were displayed in accordance with the BING Lexicon (FIG. 1). Each tweet was then parsed again and given a sentiment score in accordance with the AFINN Lexicon (FIG. 2).

IV. RESULTS

As predicted, sentiment score was associated with the date at which a COVID-19 tweet was created at.

A. Association Between Date and Tweet Sentiment

In light of these results, we used our sentiment score graph as the basis for a linear regression analysis in order to analyze the statistical significance of our results. Our model can be represented by the following equation:

$$\widehat{score} = \beta_0 + \beta_1 * date$$

where $date$ is the date and time at which a tweet was created at.

With $\beta_0 = 6.305e + 04$, $\beta_1 = -3.853e - 05$, and $p = 7.99e - 15$ our model is statistically significant and demonstrates a negative association between the sentiment score of a tweet and the date at which it was created at. Further specifics on our results can be found in Table I.

B. Validition

After determining that our results were statistically significant, our team tested the validity of our linear regression against Linearity, Independence, Normality, and Equal Variance. The referenced plots for each can be found in Appendix A.

Linearity and Equal Variance were validated through plotting the predicted values from our model against the residual values.

Independence was validated through plotting the residuals in the order that they appear in our dataset and inspecting it to determine that there was no visual pattern.

Normality was validated through inspecting a histogram of our residuals and verifying it was symmetrical and had no outliers.

V. DISCUSSION

From our analysis and linear regression model, our team concluded that there is a slight negative association between the sentiment score of a tweet and the date that it was created. When analyzing our regression model, we noted the stark contrast in number of tweets between 30 October - 07 November and 07 November - 09 October. Although not incuded in this report out of consideration for length, our team recognized that when we isolated the section with more tweets (07 November - 09 November) there was greater negative association between sentiment score of a tweet and date.

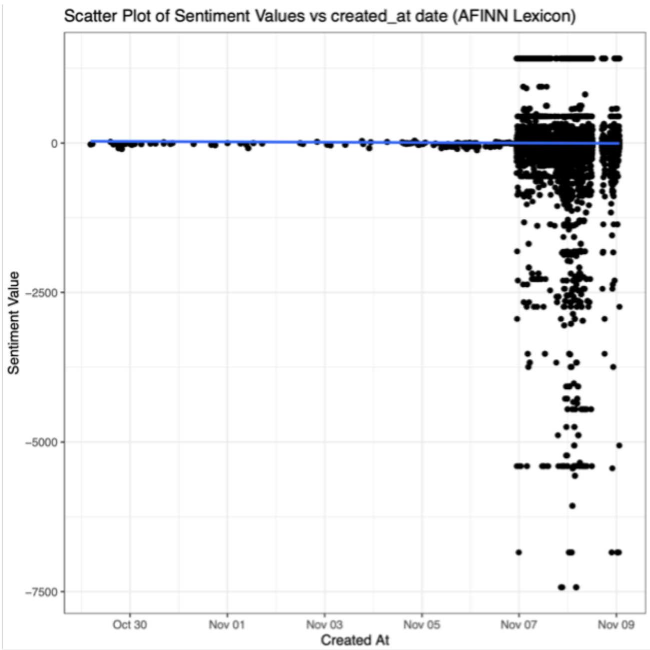

FIG. 2. Sentiment Score for Each Tweet Graphed with Linear Regression

Variable name	Estimate	Std. Error	t value	$Pr(>\|t\|)$
β_0	$6.305e+04$	$8.116e+03$	7.768	$8.00e-15^{***}$
β_1	$-3.853e-05$	$4.960e-06$	-7.768	$7.99e-15^{***}$
Multiple R-Squared: $9.258e-05$				
Adjusted R-Squared: $9.105e-05$				
F-Statistic: 60.34 on 1 and 651676 DF				
p-value: $7.99e-15$				

TABLE I. Lin Reg Results $p < 0.05^{*}, p < 0.01^{**}, p < 0.001^{***}$

A. Limitations

The difficulty for our team lied mainly in the retrieval, cleaning, and parsing of our data. As stated before, the tweets that were pulled contained many characters that we were unable to analyze. Moreover, given the status of our developer account, we were only able to retrieve tweets from certain dates, hence the concentration of tweets from 07 Nov to 08 Nov 2021. We were also unable to detect if there were any spam or bot accounts that were making up a larger proportion of the negative tweets that we parsed. Note: Spam or Bot accounts are accounts that are not people but are computers whose job is to tweet predefined statements and flood other users' Twitter® Feed.

B. Future Work

Future work in this field of research could include analyzing spam and bot accounts. If these accounts were able to be identified and a network were to be constructed, then researchers would be able filter tweets out and gain a more accurate reading. Additionally, researchers could analyze the broad psychological affects of the constant reinforcement of negative sentiments throughout the news and social media spheres. Finally, researchers can utlize machine learning to identify how to combat misinformation within social media spheres.

C. Conclusion

We recognize that our results do not confirm our hypothesis fully, however we believe that this work serves as the basis for extensive future reseach, in the same vein of Chakraborty et al. [1] and Rustam et al. [2].

Appendix A: Validity Figures

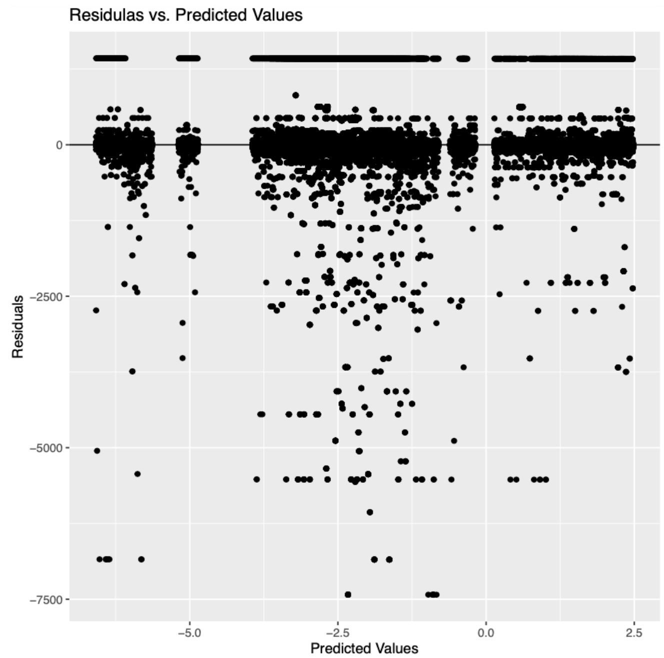

FIG. 3. Validating Linearity and Equal Variance

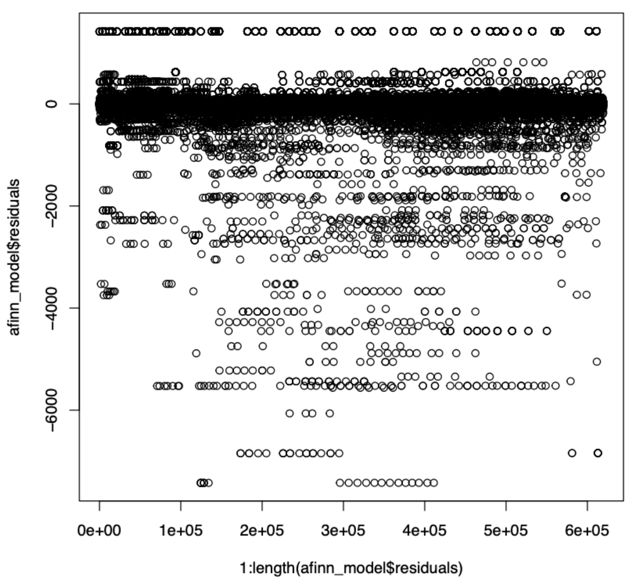

FIG. 4. Validating Independence

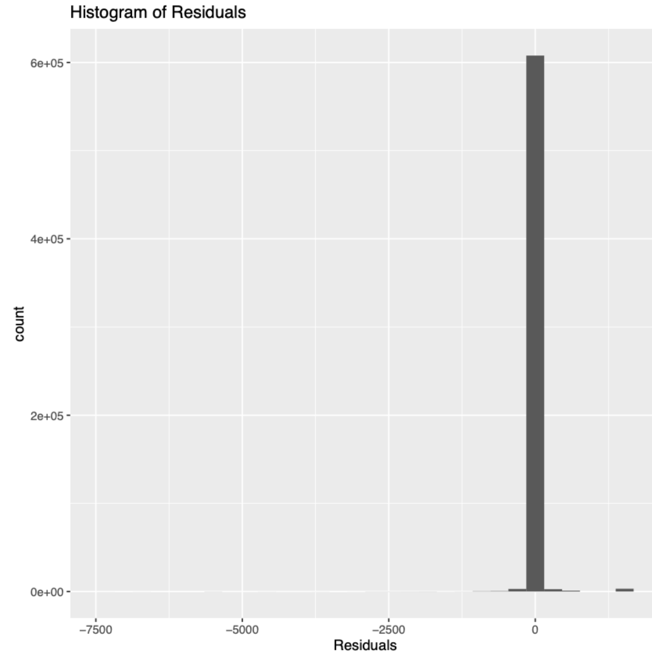

FIG. 5. validating Normality

REFERENCES

[1] K. Chakraborty, S. Bhatia, S. Bhattacharyya, J. Platos, R. Bag, and A. E. Hassanien. Sentiment analysis of covid-19 tweets by deep learning classifiers—a study to show how popularity is affecting accuracy in social media. *Applied Soft Computing*, 97:106754, 2020. ISSN 1568-4946. doi:https://doi.org/10.1016/j.asoc.2020.106754. URL `https://www.sciencedirect.com/science/article/pii/S156849462030692X`.

[2] F. Rustam, M. Kahlid, W. Aslam, V. Rupapara, A. Mehmood, and G. Sang Choi. A performance comparison of supervised machine learning models for covid-19 tweets sentiment analysis. *PloS One*, 16, 2021.