Twitter *Sentiment Analysis of a Trending Topic Using An Unsupervised Lexicon-based Approach*

Adebusola Adewole
Informatics. Univeristy of Louisiana at Lafayete: *Univeristy of Louisiana at Lafayette,*
Lafayette, USA
adebusola.adewole1@louisiana.edu

***Abstract*— The outbreak of the COVID-19 virus has no doubt had an enormous impact on people's lives. It has led to social distancing and enforcing of the mask mandate. Twitter, a microblogging site where people share their views and opinions on social and political issues offer a way to share sentiment. In this paper I analyzed 5000 tweets from 2022 relating to the lifting of the mask mandate to understand people's public opinion on wearing masks. The results from this study show that more people have positive sentiment regarding the lifting of the mask mandate.**

Keywords—sentiment analysis, Textblob, social media, hashtags

I. Introduction

Social media and the Internet has been instrumental in changing the way people express their views and share information[2]. Twitter is one of the most popular microblogging websites where users are allowed to post their opinions on different social and political issues [9]. Founded in 2006, it has risen to have 290.5 million active users making it one of the most visited websites. Users are allowed to comment, share and post messages referred to as 'tweets' of no more than 280 characters. Twitter has an average of 500 million tweets per day and users can post from their mobile app, desktop or the web. These tweets represent a unique and large data set that can be analyzed to generate insights and understand people's sentiment about a particular topic [1][7].

Sentiment analysis is one of the most popular uses of Natural Language Processing (NLP) [3]. It has a wide range of application from businesses, movies, launching a new product to better understand reception based on people's opinions [7]. The purpose of sentiment analysis is to understand opinions on issues and to identify and classify those opinions into positive, negative or neutral based on text data. Sentiments are classified and quantified based on a polarity and subjectivity score. Sentiment analysis can also provide a way for us to predict future trends and how users will respond to social and political issues, and help businesses improve their strategy. [5].

Sentiment analysis approaches can either be considered as lexicon based or supervised machine learning based. Lexicon based approaches are unsupervised [6]. Popular lexicon used for sentiment analysis includes AFINN lexicon, Bing Liu's lexicon, MPQA subjectivity lexicon, sentiWordNET, VADER lexicon, and TextBlob lexicon.

Machine learning algorithms are usually applied to these text data mined from social media sites. These algorithms help classify the opinions into negative, positive or neutral. Popular machine learning approaches that have been used include Naïve bayes, Support vector machine, Maximum entropy method, Artificial Neural network, Decision trees [5][6]. Supervised machine learning approaches require the use of feature extraction and a training model using feature set and some pre-labelled data set.

The purpose of this paper was to analyze views and sentiments of twitter users when they lifted the Mask mandate using Textblob. The emergence of the virus put the entire world in panic mode requiring us to protect ourselves. I evaluated the polarity and subjectivity of those tweets.

II. Method

A. Python

Python is a high-level programming language with a range of applications. It also provides different libraries which can be used for NLP, data analysis etc. I used the following libraries during the study Tweepy, WorldCloud, PIL, Pandas, Numpy, Re, Os, Matplotlib.

B. Data Collection

5000 Tweets were collected from twitter using Tweepy, a client for Twitter Application Programming Interface (API). To get the tweets from the Twitter API I had to register an App through my twitter account. This enabled me fetch my consumer keys, consumer secret keys, access token keys, and access Token secret. These keys were used as authentication for Twitter API in the code to collect keys anytime it was run. I gathered tweets in English language that had the key word #mask mandate from 2022. I made sure to filter out retweets to remove duplicates

C. Preprocessing in Python

I stored the tweets collected in a data frame and proceeded to preprocess the tweets to get them ready for sentiment analysis. I created functions to clean the data that removed links, hashtags, punctuations marks, and numerical values.

D. Tweetblob

I used Tweetblob, an outstanding open-source library for performing sentiment analysis. It gives a polarity and subjectivity score. These scores have a normalized scale; a

Identify applicable funding agency here. If none, delete this text box.

polarity score is a float within the range [-1.0, 1.0] while a subjectivity score is a float within the range [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective.

III. Results and Discussion

Each tweet was classified as negative, positive or neutral based on the polarity score of the words used. A score of 0 indicated that it was a neutral tweet, a score less than 0 is a negative tweet and a score above 0 is a positive tweet. In this study a positive sentiment means that they are in support of the mask mandate, a negative tweet indicates that they have negative sentiments towards wearing a mask while a neutral tweet indicates that they are indifferent.

The hashtag #maskmandate became a trending topic in March 2022 when a federal judge lifted the mask mandate, after requiring that people wear mask in public places due to the COVID-19 pandemic. According to my study, 40.34% of people where in favor of ending the mask mandate. It has been 2 years since the pandemic started and a lot of people are already vaccinated. They find it extremely uncomfortable to wear the mask (Fig.1). On the other hand, 25% of people were against ending the mask mandate probably due to the fact they still want to protect themselves from the deadly virus. However, 34.66 % of people were neutral about the mask mandate.

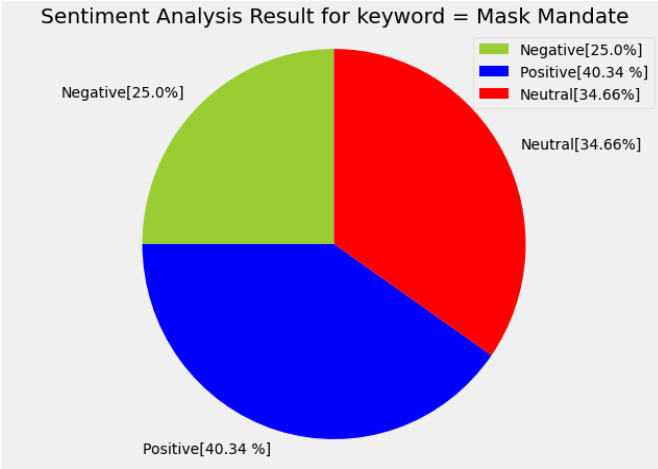

Fig. 1. Sentiment Analysis for the Keyword #MakMandate

The sentiment analysis scatter plot revealed that majority of the tweets were concentrated around -0.25 to 0.50.

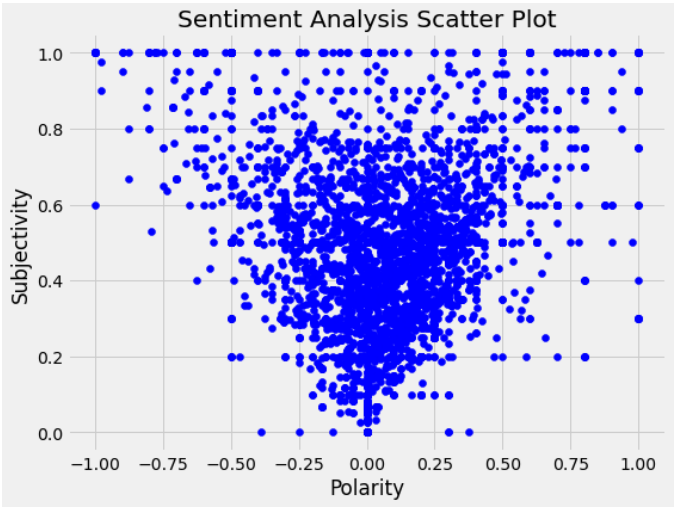

Fig. 2. Scatter plot showing Subjectivity against Polarity

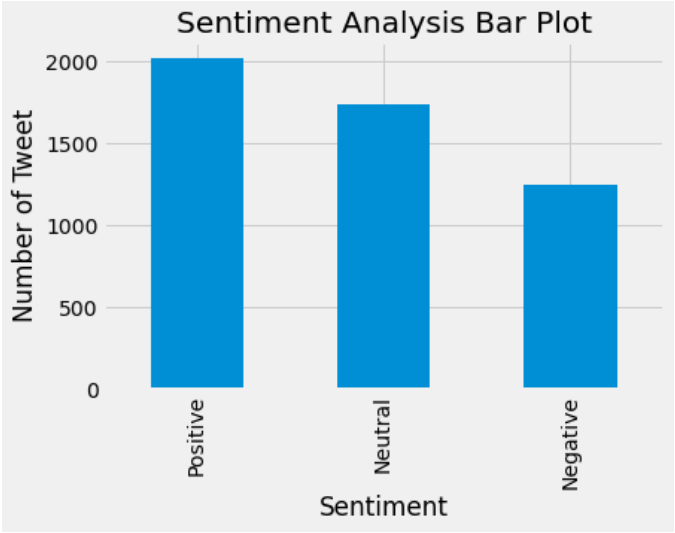

Fig. 3. Bar Plot showing number of number of tweets asscoiated with a sentiment

Fig. 4. World cloud showing frequent word in the Positive tweets

World cloud shows the frequent words used in the tweets that had positive sentiment about ending the mask mandate. Words like ‘right’, ‘free’, ‘mask madness’, ‘joy’ were frequently used denoting their relief and happiness about ending the mask mandate.

Based on the results from this study, Textblob is a powerful tool that can be used to learn the sentiment on people’s opinion. Further research can be done by comparing value of accuracy of supervised and unsupervised approaches to sentiment analysis.

Literature cited

[1] B. Le, H. Nguyen, “Twitter Sentiment Analysis Using Machine Learning Techniques”. In: Le Thi, H., Nguyen, N., Do, T. (eds) Advanced Computational Methods for Knowledge Engineering. Advances in Intelligent Systems and Computing, 2015, vol 358. Springer, Cham.

[2] M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT*)*, 2013, pp. 1-5,

[3] I. Hemalatha, G.S. Varma, and A. Govardhan,. “Sentiment analysis tool using machine learning algorithms”. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS),2013, *2*(2), 105-109.

[4] R. S. Jagdale, V.S Shirsat,, and S.N. Deshmuk, “Sentiment analysis on product reviews using machine learning techniques” In Cognitive Informatics and Soft Computing, 2019, Springer, Singapore pp. 639-647

[5] B. Gupta, M. Negi, K. Vishwakarma, G. Rawat, P. Badhani, and B.”Study of Twitter sentiment analysis using machine learning algorithms on Python” International Journal of Computer Applications, 165(9), 29-34.

[6] D.P. Turney,”Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews.” Proceedings of the 40th annual meeting in association for computational linguistics. Association for Computational Linguistics, 2002.

[7] Z. Rasheed, A.N. Ibupoto, S.S. Bano, S. Ahmed. “Sentiment Analysis of Current Trending Topics on Twitter User Base”. Journal of Mechanics of Continua and Mathematical Sciences, 2021,vol16 pp 22-31

[8] J.P. Gujjar and H.P. Kumar. "Sentiment Analysis: Textblob For Decision Making." International Journal of Scientific Research & Engineering Trends 7,2021, vol 7 pp1097-1099.

[9] R. A. Laksono, K. R. Sungkono, R. Sarno and C. S. Wahyuni, "Sentiment Analysis of Restaurant Customer Reviews on TripAdvisor using Naïve Bayes," *2019 12th* International Conference on Information & Communication Technology and System (ICTS), 2019, pp. 49-54

[10] S. Madhu “An approach to analyze suicidal tendency in blogs and tweets using Sentiment Analysis” Int. J. Sci. Res. Comput. Sci. Eng. 2018, 6(4):34-6.