

Twitter text mining and sentiment analysis on COVID- 19

Aysha Siddiqua¹ and Siddhaling Urolagin²

^{1,2} Department of Computer Science,

²APP Centre for AI Research (APPCAIR),
Birla Institute of Technology and Science, Pilani,
Dubai International Academic City,
Dubai, United Arab Emirates.

Abstract

With the boom of Internet over the past couples of decades, the amount of data being created, generated, stored and recorded has grown exponentially. Twitter is an enormously renowned microblog on which individuals interact. In the last decade, an opinion survey on Twitter data has received considerable attention and it entails the scrutinizing "tweets" (comments) and their contents. In this paper, we examine the outcomes of the different sentiment analysis methods for Twitter data. In this paper, we will look into the sentiments of the Twitter messages (tweets) received during the COVID-19 pandemic. The research focuses on the sentiment analysis of tweets using Python programming language with Tweepy and GetOldTweet3 library package. We will examine the sentiment of Twitter messages (tweets) obtained throughout the COVID-19 pandemic in this study. This paper is intended to conduct a social study of people sharing social media information, particularly Twitter, in order to understand their perceptions about COVID-19. We use an over 50,000 tweet set sent in the year 2020, mainly via hashtags, as a source of data. This study will also aid researchers and policymakers in gaining a better understanding of the emotional impact of a pandemic on society.

Keywords: *sentiment analysis, covid- 19, twitter mining, social media*

Objective

My goal is to try to quantify and interpret early changes in Twitter activity, content, and sentiment about the COVID-19 epidemic, as well as to extract all the hidden information related to each tweet from a specific time period. After that, we must refine the tweets, which is a critical step, so this research paper is attempting to find the most appropriate method for extracting or refining tweets.

Motivation

In the pandemic era, the use of social media and twitter has become increasingly widespread. This study was motivated primarily by the impact of the COVID-19 pandemic on social sentiments, mainly by the tweets and comments of prominent people in society and their reaction to tweets.

1. Introduction

Data mining is an effective tool for finding patterns and correlations among our data. Data mining discovers confidential data from massive databases. People are using social network apps to

voice their opinions on everyday issues, leading to the exponential growth of these networks. Twitter is a hugely successful microblog with over 300 million users monthly where people can express themselves in the form of short messages known as tweets. It is imperative to gather and analyze people's responses to purchasing a product, using public services, and so on. The amount of content related to the pandemic has increased dramatically on social media sites. Sentiment analysis (also known as opinion mining) is a popular dialogue preparation task that aims to explore the sentiments that underpin opinions in texts on a variety of subjects. In recent years, sentiment analysis researchers have focused on analyzing opinions on a wide range of subjects, including movies, consumer goods, and everyday societal issues.

The novel coronavirus, COVID-19, was first discovered in Wuhan City, Chinese Hubei Province, at the end of December 2019 and has spread globally, it was officially named Serious Acute Respiratory Syndrome (SARS-CoV-2)[1][2]. This epidemic frightened the entire world in the space of a few months and aroused concerns regarding the future. Over the course of 2020, COVID-19 was one of the trending areas on Twitter and has been explored further till now[3]. Many federal agencies worldwide use Twitter as a primary means of communication for constantly exchanging policy updates and news to the public during the ongoing COVID-19 pandemic[4]. The data can then be used to extrapolate valuable knowledge that can be helpful to understand the impact of COVID-19 as well as gain an insight into how individuals respond to their respective governments during the global pandemic emotionally[5][6]. Sentiment Analysis is contextual text mining, whereby subjective knowledge is identified and extracted from source documents that can be helpful in many science or business fields such as detection of events, recommending systems, and opinion mining. The goal of this research paper is to gain a stronger grasp of social opinions and perceptions on Covid-19 and also how the thinking of individuals in recent months has transposed.

This study outlines a systematic methodology for analysis of Twitter data in particular to identify sentiment, key words association and trends for crisis scenarios related to the ongoing COVID-19 phenomena. We initiate discussions and quest for insights with comprehensive textual analysis and data visualization, which include Word Cloud exploration.

The organization the paper is as follows. Section 2 presents review of literature on twitter text mining. Section 3 introduces on methodology. In section 4 building the research design on COVID-19 discussed. In Section 5 the experimental setup is presented. The experimental result and discussion in section 6 and conclusion in section 7.

2. Literature Review

Several researchers have worked on sentiment analysis on various social media data, specifically on Twitter, few main contributions that help to discover users' attitudes or sentiments in

different cases when a pandemic occurs worldwide. A number of relevant papers that are used as reference are covered in this section.

The author Turney et al, used bag-of-words method for sentimental analysis approach by which the relationship between words was not at all taken into account and a document is interpreted as a simple set of words. The sentiments of every word were identified and those values are linked to some aggregation functions to determine the sentiments for the whole document.[7]

Twitter released some 500 million tweets a day; Twitter users account for over 22% of the total Internet community in comparison to the number of active Twitter users and the total worldwide Internet users. Two other major reasons were given to choose Twitter in: (1) it is a prominent platform for sharing news, which is developed as an appropriate means for collecting public health information; and (2) tweet-sharing hashtags enable the social network to be a hub-and-spoke.[8]

Bhat et al. (2020) have used two hashtags to carry out sentiment analysis for current coronavirus pandemics: 92,646 tweets were collected with regard to the #COVID-19 and 85,513 tweets were collected using #Coronavirus, the majority of tweets were positive (51.97%). Tweets accounted for 34.05% neutral sentiments and 13.96% for negative sentiments. On the contrary, coronavirus related sentiments were more neutral (41.27%). 40.91% of the total amount of positive tweets and 17.80% accounted for negative. The findings show that the views of users were mostly positive or neutral.[9]

The author in this paper, the feelings about COVID-19 were investigated, and different people's feelings about the pandemic were examined. Therefore, with the aid of machine learning techniques, the Twitter API obtained useful corona virus tweets and then analyzed them based on positive, negative, and neutral emotions. In addition, authors used the NLTK library to pre-process fetched tweets and the Textblob dataset to evaluate tweets, since the exciting results suggest that they have positive, negative, neutral feelings in different views.[10]

In this paper, twitter has proved to be helpful for a variety of activities, including emergency communication network, public emotion monitoring, anomalies detection and pre-warning, etc. Twitter was used as a source of data for the monitoring of public and health disaster responses (e.g. hurricanes, floods, earthquakes, disinformation spreads, public health and so on) and disease outbreaks.[11]

Lisa et al. and Ramez et al. presented their works on the spread and quantification of misinformation on twitter relevant to COVID-19. The authors conclude that there is an unprecedented rate of medical misinformation and unbelievable material posted and distributed over the pandemic on the twitter. In order to avoid excessive distress and medical harmful practices in the battle against COVID-19, the disinformation on social media should be measured and actions taken.[11]

There is a latest and extensive literature on sentiment analyzes (SA) and their applications in different fields is available. Social networking data are also investigated and widely used by organizations to analyze consumers' opinions and feelings of goods, services and company strategies.[12]

The identification of events is also a research field related to twitter data mining. The researchers in this paper suggested a general structure in social media documents for recognition of events. They used methods of similarity metric learning to yield high-quality results. They also indicated that similarity metric methods provide better results than traditional approaches that take text-based approaches into consideration.[13]

In[13], a method for identifying events in real time was suggested. They proposed an Algorithm to classify target events in real time and included tweets related feature such as keywords, number of words and their contexts in order to detect target events were considered. Identifying earthquakes is the main subject of their research.[13]

R. Medford, S. Saleh et al. tweets with COVID-19 based hashtags were extracted and assessed frequency of keywords related to infection prevention, vaccination and racial prejudice. They conducted a sentiment analysis in order to identify emotional valence and predominant emotions with the use of topic modeling to identify and address issues over time.[6]

The author, in this article examined US tweets which contained "Chinese-virus" or "Wuhan-virus" for the purpose of performing a user characterization, referring to the COVID-19 pandemic. The findings were compared to users who did not use this controversial language. The results show that variations in age, geolocation or followed politicians came into account.[1]

The author has reviewed state and country level tweets from the USA. First, they found that people tweet more about COVID 19 during the work hours as the pandemic spread, and that they were able to identify variations in temporal patterns of tweets. In addition, they performed an over-run sentiment analysis that included an event specific subtask reporting negative sentiment when the 1000th death was confirmed and positive when the lockdown measures were eased in the states.[14]

In this paper, the author analyzes Twitter messages obtained in Europe in accordance with sentiment in the first months of the COVID-19 pandemic. This is achieved using multilingual sentence embedding with a neural network for sentiment analysis. This will provide a clear picture of the situation's impact on the mood of the people.[3]

3. Methodology

The dataset is gathered using online API tweets from twitter data. Data containing tweets IDs and sentiment scores related to pandemic COVID-19 and is processed through a set of five phases. The phases include Tweet collection & pre-processing, Tweet's cleaning, Feature selection, Modelling and Evaluation[15].

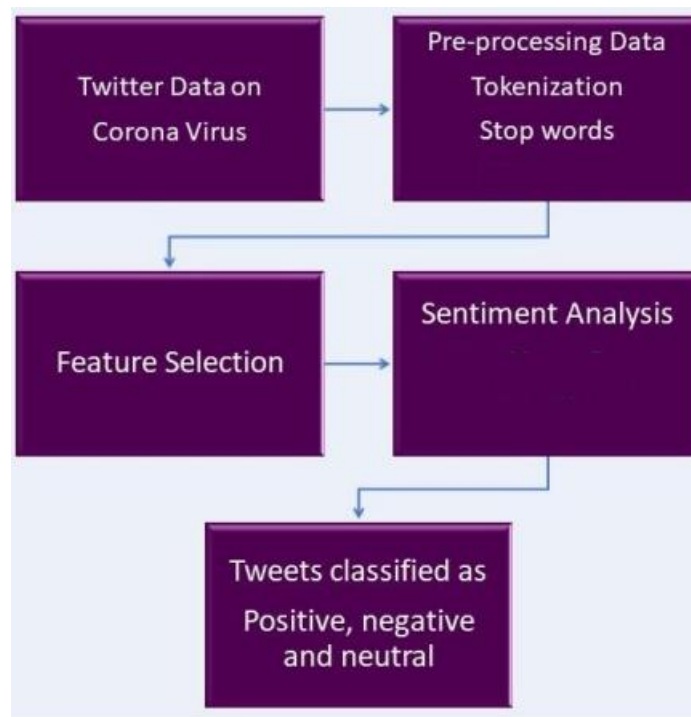


Figure 1

Sentiment Analysis Model

Tweets Collection and Pre-processing- The initial step involves extracting data from csv files. This is achieved through the Pandas library. The CSV file used is 'covid19_tweets.csv', file that contain numerous tweets and information related to tweets, such as the username, the number of retweets or likes, etc. Pre-processing is a method of converting unstructured data into structured data. Hashtags, hyperlinks, whitespaces, usernames, URLs, stop words etc[15]. are removed from tweets to reduce the text volume to a manageable quantity and reduce the time required for the model to execute.[10], [11]

Tweet's cleaning- Tweets cannot be assessed for analysis in their original form. Tokenization is a splitting of string into smaller terms called tokens. Tokenization involves identification of nouns, verbs, adverbs etc. Stop words are words which don't benefit the sentiment of text and have little value in a program for sentiment analysis. Therefore, they are just noise in the text which may be omitted.

Building tweet dictionary & determining word density- The word dictionary is created after text processing. All terms in all datasets are classified into positive, negative and neutral. The number of occurrences of all unique words in all of this dataset is then calculated as the density of each word. This dictionary can be used to evaluate the test set.

Feature Selection- is based on the density of word. The Term Frequency Inverse Document Frequency (TF-IDF) method determines the word density. TF-IDF distinguishes the words in the tweet that occur frequently and those which do not occur frequently in the other training sets. Word density aids in knowing about the tweet polarity.

The tweet polarity is measured using TF-IDF by means of the term weight. The positive and negative characteristics of the word are determined according to the number of times the word appears in a given tweet dataset. To find the rank of each term, TF-IDF is implemented in every term in the dataset. A high ranking in TF-IDF indicates that the word is relevant in the tweet and is important to the polarity of the tweet.

The overall methodology works as below:

Given a Data Set D , a term t , and an individual tweet (dt), $dt \in D$, we calculate[15]:

$$\text{Adjusted TF-IDF} = f_{t, dt} * \log (|D| / f_{t+s, D}) \text{ [15][16]}$$

Where $f_{t, ds}$ equalant to the no. of times term t appears in d_{s1}

$|D|$ is the size of the Data set, and $f_{t+s, D}$ equals the number of tweets in which the term t and its corresponding synonym word appears in D .

The corresponding synonyms in the word dictionary are collected for each term in the term data. Synonyms are regarded as terms corresponding to the original term and are thus used in the calculation of $f_{t, D}$. Set of terms that are extracted using Adjusted TF-IDF are used to judge the tweet polarity[15].

4. Research design

In this paper, an observational study design was used and an purposive sampling method to select all the specified hashtag tweets (e.g., #2019nCoV) relating to COVID-19 in Twitter . For impactful topics and terminology specific to COVID-19, we used natural language processing methods. Data preparation and data analysis is part of our Twitter data mining strategy. Three steps involved in data preparation: (1) sampling; (2) data collection; and (3) raw data preprocessing. We continued the data analysis, including (1) sentiment analysis, following the pre-processing of the raw dataset. Every message-level tweet posted on Twitter was the unit of analysis.

5. Experimental setup

The primary method of data collection was through Twitter. Using a python package called GetOldTweets3, we were able to scrape old tweets from Twitter without needing any API keys. Tweets were collected around significant dates related to COVID-19 shut downs and re-openings. It focused on countries such as *New York, New Jersey, Texas, Florida, Ohio, Georgia, Michigan and Illinois*. And later, Tweepy python library was used for extracting data from twitter API for further research. The dataset was collected from January to August, 2020. Tweepy allows

appropriate data retrieval by searching via keywords, hashtags, timelines, trends, or geo-location. All the tasks have been done in Anaconda using python programming language. While collecting the data non-English keywords, stop words, @, #, RT, emojis has been removed and the sentence was tokenized. The tweets are filtered with the hashtags “#COVID-19 and #coronavirus”. The gathered data is saved in CSV format, and fed for the Sentiment Analysis using VADER. Sentiment analysis is a part of NLP that identifies the general opinion of text. VADER which stands for (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is used to analyze sentiments expressed on social media. Vader categorizes its results into 4 scores:

- Positive—shows how positive the text is (0.05 to 1)
- Negative—shows how negative the text is (-1 to -0.05)
- Neutral—shows how neutral the text is (-0.05 to 0.05)
- Composite—combines the three scores into one general/overall score
- Ranked from -1 to 1

6. Results and Discussions

In this paper, python is been used to implement sentiment analysis using VADER. Some packages have been utilized including GetOldTweet3 and Tweepy. Installation of required libraries can be done:

- `pip install tweepy`
- `pip install GetOldTweet3`
- `pip install wordcloud`

Following necessary libraries have been imported:

```

#! pip install branca==0.4.1 #0.3.1
#! pip install wordcloud

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import seaborn as sns
import os
import string
import re
from datetime import datetime
import matplotlib.pyplot as plt
import seaborn as sns
import branca.colormap as cm
from mpl_toolkits.basemap import Basemap
import requests
import folium
from folium import plugins
from folium.plugins import HeatMap
import branca.colormap
from nltk.tokenize import TweetTokenizer
from nltk.corpus import stopwords
from nltk import pos_tag, ne_chunk
from nltk.sentiment.vader import SentimentIntensityAnalyzer
from wordcloud import WordCloud
from tqdm import tqdm, notebook
%matplotlib inline

```

Figure 2

As the Covid-19 pandemic continues to be the topic of conversation across all dinner tables and of course TV screens in the United States, general sentiment regarding COVID-19 announcements have become a point of interest.

Since most of the response to this pandemic has been mostly regional and state based, we wanted to focus on collecting tweets from several high-profile states that have been impacted by the virus, compare how the states reacted to the news of their respective shut down and reopening announcements, and determine if states reacted differently to the announcements.

These states include but are not limited to: - New York, New Jersey, Texas, Florida, Ohio, Georgia, Michigan and Illinois.

Evaluating the shutdown rate

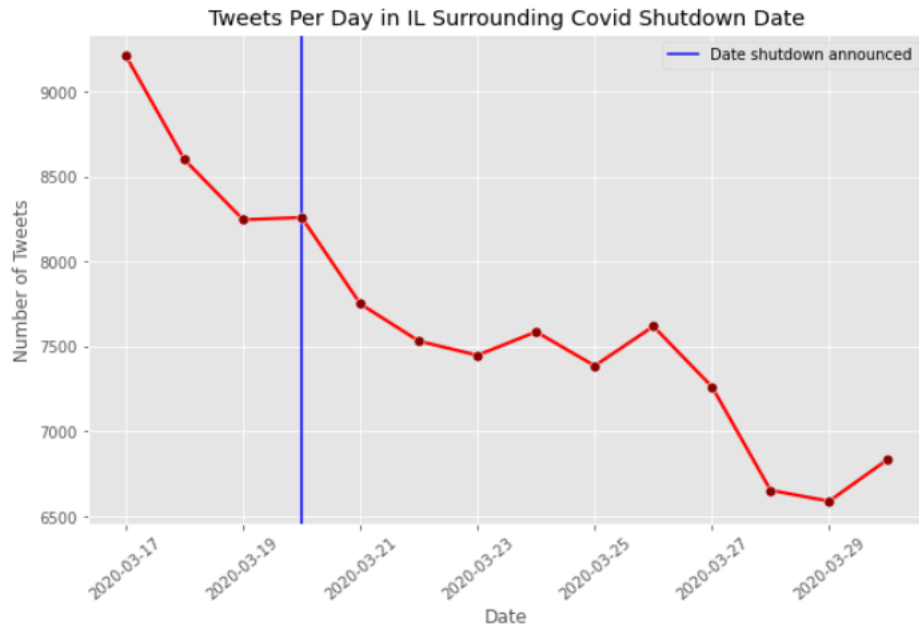


Figure 3

Tweet activity steadily decreased in the weeks following the announcement of the shut-down in Illinois on **March 20, 2020**.

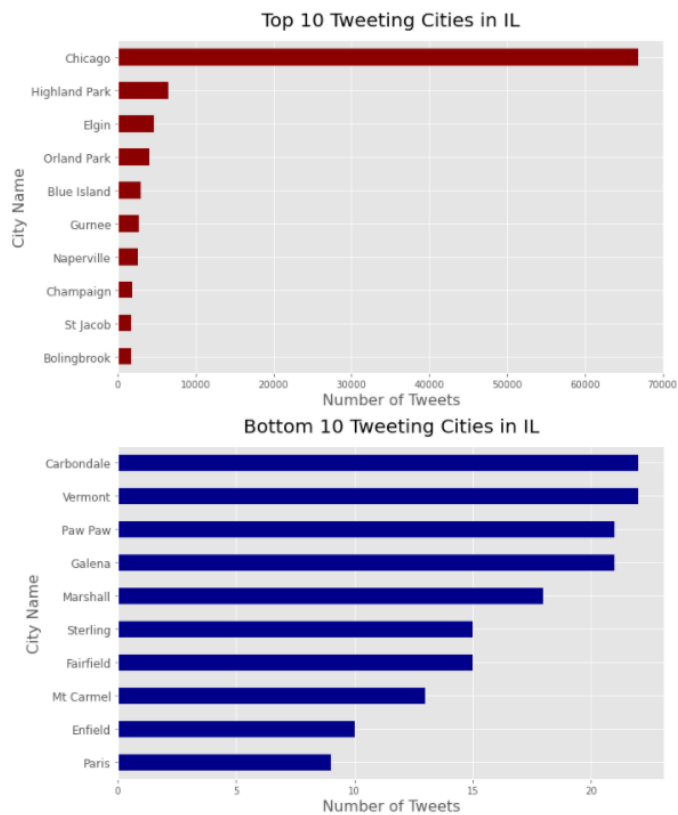


Figure 4

The top three tweeting regions are Chicago, Highland Park, and Elgin. These may not be precise city areas, but regions, as collection method involved the radius surrounding a city.

Tweets per day

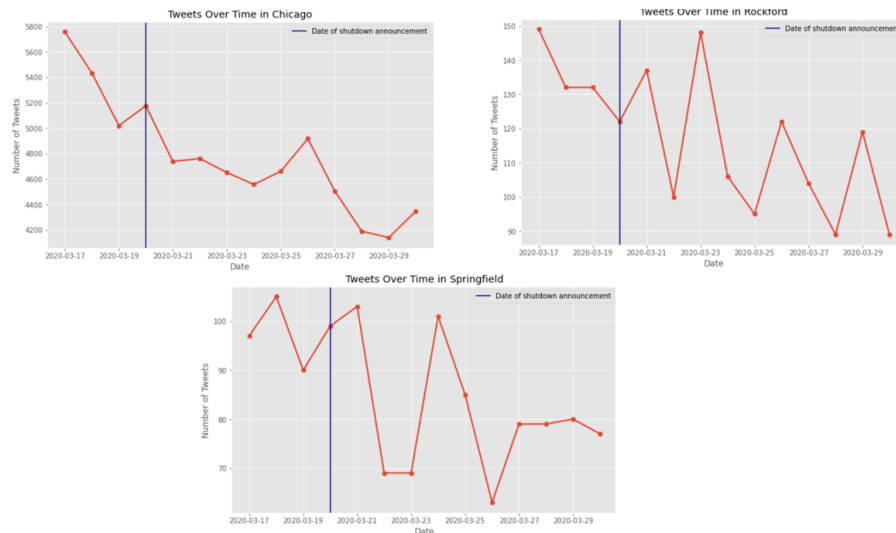


Figure 5

Getting positivity rate at shutdown announcement date

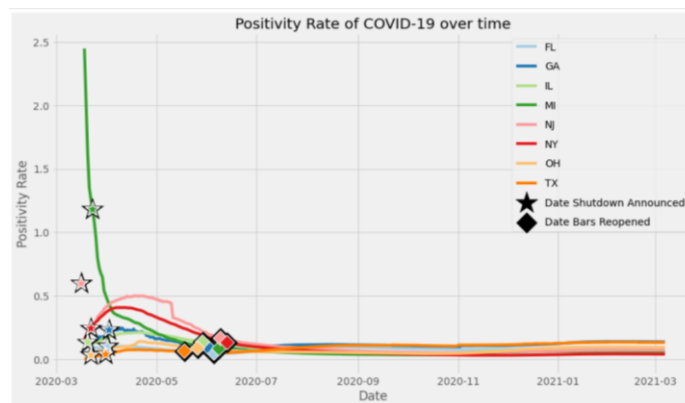


Figure 6

Note the stars on this plot not connected to lines belong to NJ and OH. These states had positivity rates that did not represent the natural trend at the onset of their Covid shutdown.

Insights:

- States allowed bars to reopen within a month window around the beginning of June
- States whose **positivity rates went down** after reopening bars include: OH, NY, IL, MI, and NJ

Plotting the heatmap to see the geographical distribution based on number of tweets

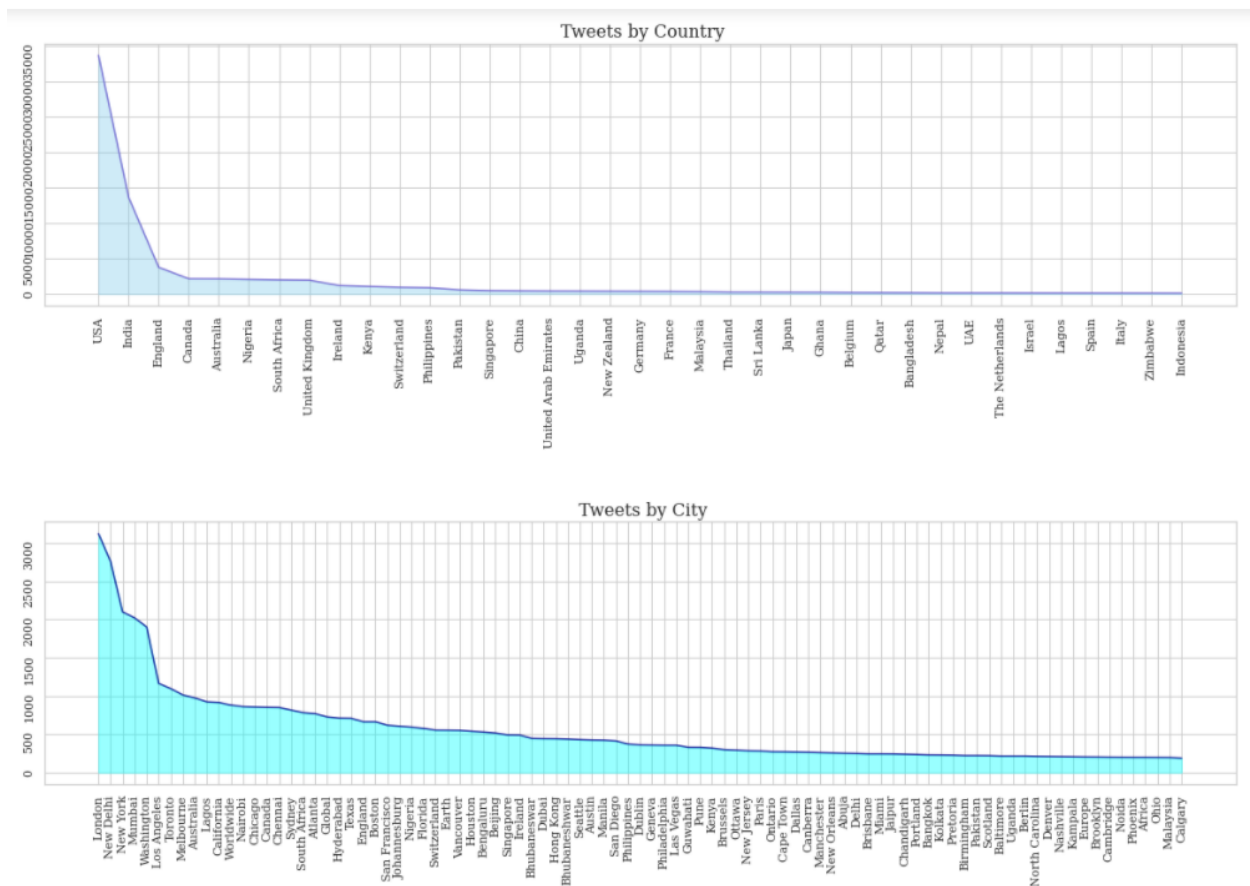


Figure 9

The figure depicts that that US, India, and England are the top 3 countries with the highest tweets with USA having significantly higher tweets (~20%). The curve is almost flattened beyond the top 3 countries with ~ 2k tweets (or less) tweets per country.

The top 5 cities with highest tweets are London, New Delhi, New York, Mumbai, Washington with each having 2k-3k tweets.

Tweets sentiment distribution

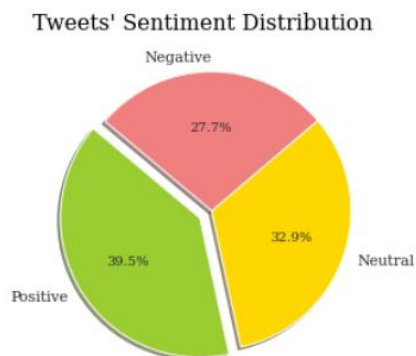


Figure 10

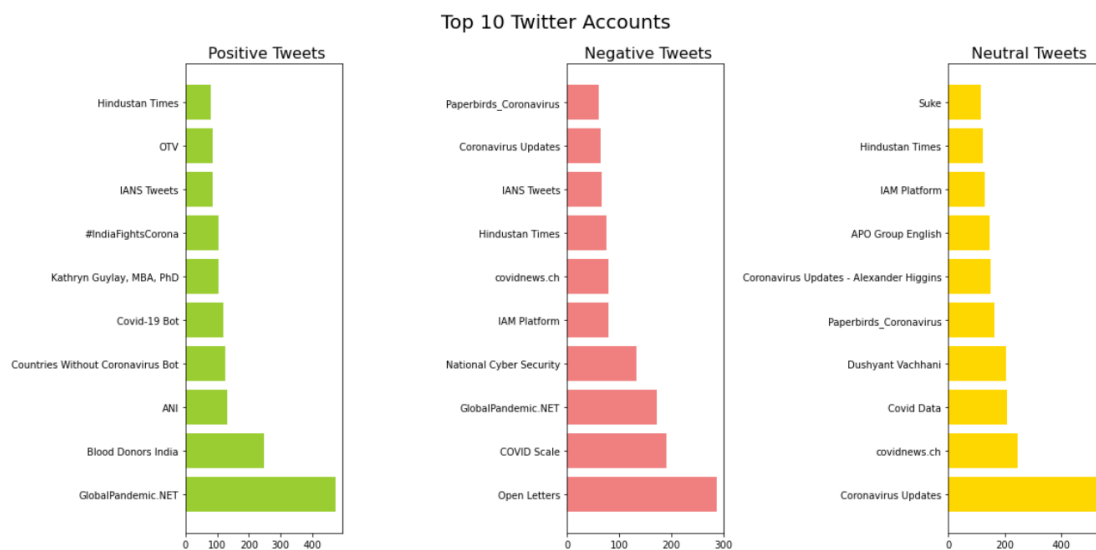
A total of ~70k tweets have a positive sentiment making the largest pie in the chart with ~40% tweets.

The top 10 twitters account being followed

user_name	user_followers
CNN	49442559
National Geographic	24359165
CGTN	13892841
NDTV	13568396
The Times Of India	13057020
United Nations	12806543
China Xinhua News	12681192
CNN International	11081219
WIRED	10380131
ABP News	10363089

Figure 11

CNN, and National Geographic are the top 2 twitter accounts to be followed the most with CNN having over 50m followers, and National Geographic having ~25m followers. Next in line are CGTN, NDTV, and Times of India with each having slightly over 13m followers.



'GlobalPandemic.NET' has the highest number of positive tweets while 'Open Letters' leads the list for having maximum negative tweets and 'Coronavirus Updates' for neutral tweets.

Figure 12

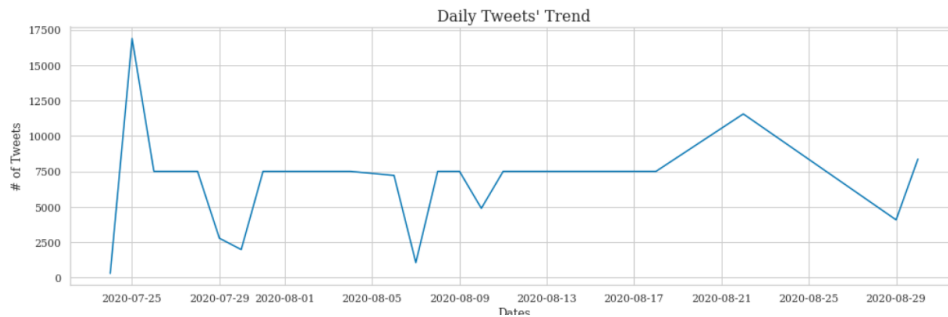


Figure 13

The figure shows a spike in number of tweets in the last week of July. This could possibly be because of the highest number of corona cases during that time. July 24, had the highest daily corona cases till date with ~290k globally and ~80k in the US. This could probably explain the high volume of tweets in that week and especially on July 25th.

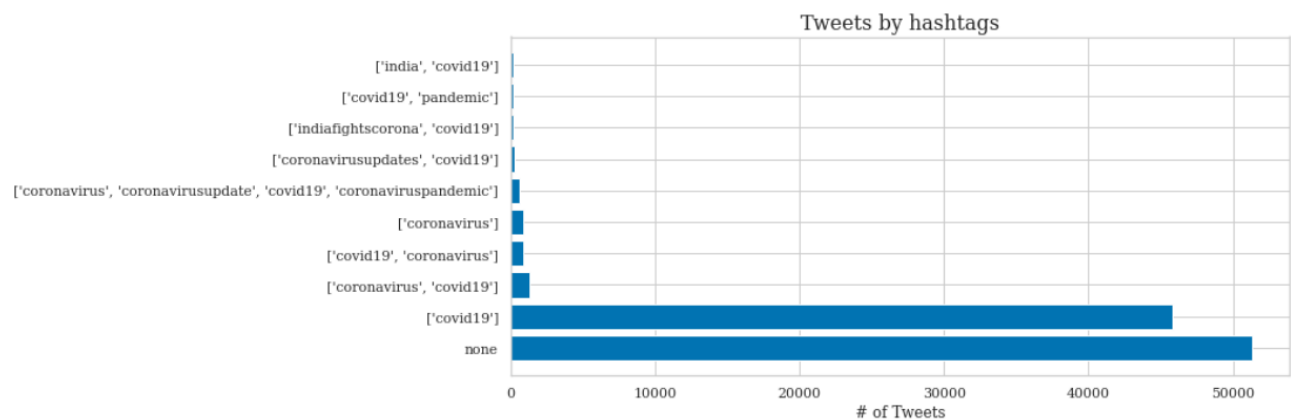


Figure 14: Tweets by Hashtags

As expected, ~70% of the tweets mention either covid19 or coronavirus as the hashtag.

Sentiments of top 10 favorited tweets

	user_name	sentiment	clean_tweet
	Chelsea Anderson.▼	Negative	1
	Patty	Negative	3
		Positive	2
	paolo ignazio marong	Neutral	4

In [286]:	top10tweets
-----------	-------------

Out[286]:	
-----------	--

	user_name	clean_tweet	user_verified	sentiment
14504	Chelsea Anderson.▼	So Trump wants everyone else s child to get COVID19 but his son Baron gets an exception now that s messed up	False	Negative
172825	Patty	So many conflicting reports so much we don t know why is it that the USA can t contain COVID19 yet other count	False	Negative
172861	Patty	Interesting since some say asymptomatic carriers are not the main spreaders of COVID19 Which is it other report	False	Positive
173409	Patty	How do you pioneer a growing fatality rate and act like that s an accomplishment COVID19	False	Negative
148145	Patty	Herd immunity at 20 is still too high and would risk too many lives A vaccine or therapeutics to lessen the sympt	False	Negative
148540	Patty	Sounds like the need for a booster shot once a vaccine is made for COVID19 might be necessary based on this info	False	Positive
177848	paolo ignazio marong	Nuovo rimbalzo del covid19 in Turchia Coronavirus live news UK records 1 108 new confirmed cases as Myanmar rep	False	Neutral
168379	paolo ignazio marong	Sotto i 300 casi covid19 SudCorea S Korea counts 299 more COVID 19 cases on first day of stricter distancing	False	Neutral
156746	paolo ignazio marong	Il rapporto tra casi accertati e mortalit elevato covid19 Victoria Australia Coronavirus updates LIVE Victor	False	Neutral
141621	paolo ignazio marong	Focolai legati a chiese CoreadelSud Covid19 S Korea adds 197 more cases of COVID 19 total at 15 515	False	Neutral

Figure 15

Amongst the top favorited tweets, three (non-verified) accounts pop-up Patty, Paolo, and Chelsea. The negative tweets here showcase concerns about the rising COVID-19 cases and US' inability to contain it. On the other hand, there are some positive tweets talking about how the situation might improve once a vaccine is made. The other neutral tone tweets are mostly an update on the number of covid19 cases.

Looking into tweets talking about 'Trump'

```
Out[299]: 179008      WalpurgaWeller realDonaldTrump The two atomic bombs detonated over Hiroshima and Nagasaki killed between 129 0
00
179023      why is all the violence happening in Democrat run cities with Dem Mayors Govs Trump2020 BidenHarr
is
179026      realDonaldTrump Meanwhile the maskless people who attended your pathetic superspreader convention are now conta
mi
179039      the resistor realDonaldTrump it s two lives If over 185 000 totally preventable COVID19 dead Americans
do
179040      realDonaldTrump Disgraceful 176 000 plus Americans dead from COVID19 and this is what the president is worri
ed
179056      ProjectLincoln it looks like realDonaldTrump literally signed his name to the Iowa COVID19 death cou
nt
179065      realDonaldTrump LawlessTrump is getting beaten like a dog in the polls Sadly the only failed reaction which h
as
179091      dougmillsnyt realDonaldTrump FLOTUS The Trump Family contributed to the already 180 000 lives lost amp fami
li
179092      politvidchannel The Trump Administration s Incompetence led to the highest number of COVID cases in the world
U
179097      politvidchannel Just not the lives of COVID19 sufferers coronavirus kills Trump
Virus
Name: clean_tweet, dtype: object

In [300]: trump_tweets.sentiment.value_counts(normalize=True)
Out[300]: Negative    0.472854
Positive    0.266447
Neutral    0.260699
Name: sentiment, dtype: float64
```

Figure 16

The figure shows that Out of all the tweets that mention 'Trump', almost 50% have a negative tone. When inspected most of these tweets had people talking about the increasing corona cases. Some of the tweets talk about country's inefficiency in promoting awareness about coronavirus, having faulty ventilators, lack of much needed preventive measure etc.

Tweets by source

In order to get information from which device are used by the users for writing their tweets.

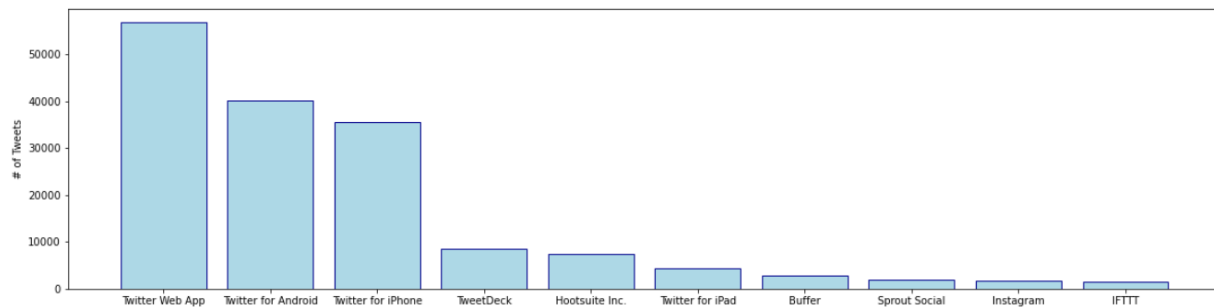


Figure 17: Tweets by Source

32% of people have tweeted using the Web App, closely followed by Android users with 22%, and iPhone users with 20%.

Closing thoughts

US (New York, Washington), India (New Delhi, Mumbai), and England (London) have the highest tweets. These are also the most populated and metropolitan cities clearly indicating that people in these cities are more active on Twitter.

People using twitter on mobile are 1.5x of the ones using web when combined the Android and iPhone users collectively making ~45% of tweets.

Looking at the tweets' trend, people have tweeted more following a spike in covid19 cases especially when cases on June 24 were the highest till date, we see a huge spike in number of tweets on June 25.

As expected, most of the news channels like CNN, National Geographic, CGTN, NDTV, and Times of India are the top 5 twitter accounts to be followed with CNN having over 50m followers, and National Geographic having ~25m followers.

While doing text analysis, we captured the words like covid19, coronavirus, pandemic, vaccine, death, mask, etc. which were mostly talked about.

Overall, there were more positive tweets (~40%) than negative or neutral, indicating people still have great hopes for the world to become normal soon again. When analyzing positive tweets further, we found 'GlobalPandemic.NET' leads the list for having the highest number of positive tweets while 'Open Letters' leads the negative tweet list and 'Coronavirus Updates' for neutral tweets.

Amongst the top 10 favorited tweets, tweeted by three (non-verified) accounts pop-up Patty, Paolo, and Chelsea:

- The negative tweets showcased concerns about the rising COVID-19 cases and US' inability to contain it.
- The positive tweets talked about how the situation might improve once a vaccine is made.

- The neutral tweets were mostly an update on the number of covid19 cases.

7. Conclusion

in this paper, the study was carried on twitter data from January to August 2020 on the epidemic of corona viruses, the virus spread through various nations and the outbreak became a pandemic. This study helps to gain better understanding on how the people have perceived about coronavirus and how its impact has effect on the public. The sentiments were downloaded during the period and the public's reaction was analyzed towards the outbreak. Sentiment analysis using VADER is applied for data analysis. During the study, nearly all the countries had a tweeting of COVID-19 with a positive sentiment, since all those people became accustomed to COVID-19 and also the rate of recovery has been improved over the course. Similarly, it has been concluded during the analysis of word clouds that people have various tweets with different thoughts, such as pandemic, COVID, case, coronavirus, etc. This study offered a clear overview of people's feelings and sentiment at Covid-19 and allowed us to fully understand that people around the world are almost on the same level of thought.

8. References

- [1] H. Lyu, L. Chen, and J. Luo, "Sense and Sensibility : Characterizing Social Media Users Regarding the Use of Controversial Terms for," 2020.
- [2] R. J. Medford, S. N. Saleh, A. Sumarsono, T. M. Perl, and C. U. Lehmann, "An ' Infodemic ' : Leveraging High -Volume Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak," 2020.
- [3] H. Matthias, A. Kruspe, and I. Kuhn, "Cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic," 2020.
- [4] R. Lamsal, "Design and analysis of a large-scale COVID-19 tweets dataset," no. October, 2020.
- [5] A. D. Dubey, "Twitter Sentiment Analysis during COVID19 Outbreak," no. March, pp. 1–9, 2020.
- [6] B. P. Pokharel, "Twitter Sentiment Analysis During Covid-19 Outbreak in Nepal Twitter Sentiment analysis during COVID-19 Outbreak in Nepal Bishwo Prakash Pokharel M . Phil in ICT 3 rd semester Nepal Open University , Nepal Keywords," no. January, 2020.
- [7] V. A. Kharde, "Sentiment Analysis of Twitter Data : A Survey of Techniques," vol. 139, no. 11, pp. 5–15, 2016.
- [8] A. Addawood and A. Alohal, "Tracking And Understanding Public Reaction During COVID-19 : Saudi Arabia As A Use Case," vol. 2020, 2020.
- [9] and B. A. Muzafar Bhat, Monisa Qadri, Noor-ul-Asrar Beg, Majid Kundroo, Naffi Ahanger,

"Sentiment analysis of social media response on the Covid19 outbreak," no. 2020 May 8, 2020.

- [10] M. A. Kausar, A. Soosaimanickam, and M. Nasar, "Public Sentiment Analysis on Twitter Data during COVID-19 Outbreak," vol. 12, no. 2, pp. 415–422, 2021.
- [11] Y. Kabir, "CoronaVis : A Real-time COVID-19 Tweets Data Analyzer and Data Repository."
- [12] H. Drias, T. Houari, and Y. Drias, "1 Introduction," 2020.
- [13] N. Azam, M. Abulaish, and N. A. Haldar, "Twitter Data Mining for Events Classification and Analysis," no. November, 2015.
- [14] Y. Feng, "OBSERVATIONAL STUDY BASED ON A LARGE GEO - TAGGED COVID-19 TWITTER DATASET," 2020.
- [15] K. B. Priya Iyer and S. Kumares, "Twitter sentiment analysis on coronavirus outbreak using machine learning algorithms," *Eur. J. Mol. Clin. Med.*, vol. 7, no. 3, pp. 2663–2676, 2020.
- [16] Y. Jain and V. Tirth, "Sentiment Analysis of Tweets and Texts Using Python on Stocks and COVID-19," vol. 16, no. 2, pp. 87–104, 2020.