



A study on the sentiments and psychology of twitter users during COVID-19 lockdown period

Ishaani Priyadarshini¹ · Pinaki Mohanty² · Raghvendra Kumar³ · Rohit Sharma⁴ · Vikram Puri⁵ · Pradeep Kumar Singh⁶ 

Received: 5 July 2020 / Revised: 17 March 2021 / Accepted: 4 May 2021 /
Published online: 14 June 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The outbreak of the novel Coronavirus in late 2019 brought severe devastation to the world. The pandemic spread across the globe, infecting more than ten million people and disrupting several businesses. Although social distancing and the use of protective masks were suggested all over the world, the cases seem to rise, which led to worldwide lockdown in different phases. The rampant escalation in the number of cases, the global effects, and the lockdown may have a severe effect on the psychology of people. The emergency protocols implemented by the authorities also lead to increased use in the number of multimedia devices. Excessive use of such devices may also contribute to psychological disorders. Hence, hence it is necessary to analyze the state of mind of people during the lockdown. In this paper, we perform a sentiment analysis of Twitter data during the pandemic lockdown, i.e., two weeks and four weeks after the lockdown was imposed. Investigating the sentiments of people in the form of positive, negative, and neutral tweets would assist us in determining how people are dealing with the pandemic and its effects on a psychological level. Our study shows that the lockdown witnessed more number positive tweets globally on multiple datasets. This is indicative of the positivity and optimism based on the sentiments and psychology of Twitter users worldwide. The study will be effective in determining people's mental well-being and will also be useful in devising appropriate lockdown strategies and crisis management in the future.

Keywords Coronavirus · Sentiment analysis · Lockdown · COVID-19 · Cognitive study · Human psychology

✉ Pradeep Kumar Singh
pradeep_84cs@yahoo.com

Extended author information available on the last page of the article

1 Introduction

Ever since its discovery in Wuhan, Hubei Province, China, in late 2019, the novel Coronavirus has been spreading at an alarming rate [23]. The nCov-2019 or COVID-19 which is associated with human respiratory disease [44], is believed to have potentially originated from either bat [46] or the Huanan seafood market in China [15]. Although severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and coronavirus disease 19 (COVID-19) are all associated with bronchitis, pneumonia, or severe respiratory illness [30], due to the increasing number of cases, the World Health Organization declared COVID-19 as a pandemic [6] within three months after its discovery. The outbreak infected more than ten million people across the globe and reported half a million deaths worldwide [7]. The effects of COVID-19 are not only restricted to infected cases and the number of deaths but also have an impact on several businesses across the world. The travel industry perished due to heavy travel restrictions [24]. The stock market and oil prices were hit critically [38]. Other industries affected due to the pandemic are the tourism industry, sports industry, and restaurant industry [32]. A major contributing factor to the uncontrollable spread of infections is the lack of an appropriate vaccine for COVID-19. Although many researchers have suggested a variety of treatments and proposed medicines, the results are not very promising [13, 18, 36, 37]. Since an established cure could not be developed even after six months of the discovery of the nCov-19, social distancing and other lockdown strategies were suggested by various government bodies all over the world [3, 26]. With the lockdown imposed across many parts of the world, restrictions were placed on major facilities and services, such that the general public was not able to avail of any of these amenities [22, 41].

Human beings are social animals [39]; hence, there is a need for them to interact among themselves as well as with the environment. The growing number of cases all over the world, the inability to socialize, and the lack of access to the environment may lead to psychological distress among people [5, 17, 45]. As mentioned before, the lockdown has also resulted in the closing down of several businesses. Thus people whose livelihood solely depended on these businesses may also be helpless economically. Many suicide incidents have also been reported following the lockdown [10]. Worldwide curfews and confinement to their respective homes have made a lot of people dependent on multimedia devices. There is an increase in the number of devices, especially smartphones [14]. Moreover, students and employees depend on Zoom applications for attending virtual classes and meetings [25]. As humans become more and more dependent on multimedia devices, the increased usage of such devices may also have a drastic impact on people's minds, leading to much lesser brain activity, loss of productivity, and device addictions. This makes it imperative to analyze the state of mind of people all over the world. While people around the globe may be dealing with several emotions like anxiety, depression, fear, stress, panic, hope, etc., there is a need to identify how positive or negative people are feeling in a lockdown situation. Therefore, there is a need to identify the sentiments of people in such a situation. While some of the past research works explore sentiment analysis during the COVID-19 lockdown situation, the limitation of many of these research works exists in the form of country-specific sentiment analysis. Moreover, the data is collected over a specific time period. This study will address these limitations by considering a global dataset and performing sentiment analysis on two different days during the COVID-19 lockdown situation. One of the most used and popular global social networking platforms is Twitter. It is a micro blogging system that is used for sending or receiving short posts [42]. We analyze tweets during two of the lockdown days and perform a sentiment analysis on the same.

Sentiment analysis is used to determine the emotional tone behind a series of words. It gives us information about the attitude, opinion, and emotion behind a set of words. The idea is to explore the emotions of people all over the world with respect to lockdown during the COVID-19 outbreak. The key contributions of the study are as follows:

1. A number of research works related to the COVID-19 outbreak and cure have been proposed before. However, little has been done with respect to the after-effects of the lockdown on human psychology.
2. In this paper, we perform sentiment analysis (positive, negative, and neutral sentences), which is an unexplored territory with respect to lockdown situations during the COVID-19 outbreak, for analyzing the state of mind of people.
3. The data collected is not specific to a given country or region, but is taken from a popular global social networking forum, i.e., Twitter. Some of the features of the dataset considered for this study are ‘created_at’, ‘text’, ‘favourites_count’, ‘retweet_count’, ‘country_code’ and ‘lang’.
4. By the beginning of April, most parts of the world had witnessed lockdown [19]. Hence it would be interesting to explore the sentiments of people after a month of lockdown. It may be essentially useful in designing appropriate lockdown strategies and crisis management.
5. Existing research works consider data for either a specific date or a range of dates. In this study, we have considered two dates, i.e., sentiments of people following two weeks and four weeks of the lockdown, in order to perform an in-depth analysis of their sentiments. This assists us in comparing how their sentiments vary over multiple weeks of the lockdown.
6. We compare the results with two other datasets for validating the study.

The rest of the paper is organized as follows. Section 2 presents materials and methods where we discuss the related works and the proposed methodology. Section 3 highlights the experimental analysis incorporating the datasets and the results. Section 4 discusses the overall results obtained, and Section 5 takes into account the conclusion and future work.

2 Materials and methods

This section comprises two parts. In the first part, we present some past related works with respect to the pandemic. In the second part, we discuss the proposed method of our work.

2.1 Related works

In this section, we explore the past research works that have been performed with respect to the COVID-19 outbreak. We also perform a critical analysis of the previous works so as to justify the novelty of our research work.

[11] performed a Twitter sentiment analysis during the COVID-19 outbreak for specific countries like Australia, China, India, United States of America. The study takes into consideration eight basic emotions i.e., fear, joy, anticipation, anger, disgust, sadness, surprise, and trust. Tweets related to Anger are highest in France, and Anticipation is highest in Germany. The USA reports the maximum number of tweets related to disgust, and Switzerland reports

fear. India reports the highest percentage of joy, and Switzerland reports the maximum amount of sadness. The highest number of tweets depicting surprise and trust are from Belgium. The limitation of this study is that it does not take into account the sentiments of people across the globe and is only confined to a few countries. [12] presented a report on analysis and forecast of COVID-19 outbreak for countries China, Italy, and France. The analysis has been performed using a susceptible-infected-recovered-deaths model, indicating that the rate of recovery is irrespective of the country, while death rates and infections vary. The case-fatality ratio is shown to be 4% to 8% in Italy and between 1% to 3% in China. According to the study, 2500 ventilation units are a fair figure for COVID-19 strategic management in Italy. Although the research highlights how the case fatality ratio can be decreased, the limitation of the research is that it is narrow and considers a very small fraction of individuals from only three countries. [4] provided a risk assessment strategy for the novel Coronavirus in the form of a computational tool. The number of cases, connectivity of destination countries with China, and efficacy of control measures of destination countries have been taken into account for performing the risk assessment. The study states that countries with low connectivity to China and higher efficacy measures can reduce the risks better. The limitation of the study is that it considers destination countries from only China's perspective, whereas COVID-19 is a global issue.

[32] performed an analysis of the outbreak of the COVID-19 and presented a comparison of the cases confirmed, recovered, and deaths with respect to China using data visualization. The study also stated some industries that were severely affected by the pandemic, like tourism, travel, and the sports industry. The study fails to discuss many other global industries like real estate, retail, supply chain, etc. that may have also been affected due to the pandemic. [1] presented an infoveillance study highlighting the top concerns of tweeters during the COVID-19 pandemic. The analysis identified specific themes, i.e., origin of the virus, its sources, its impact on people, countries, and the economy, ways of mitigating the risk of infection. Most of the tweets were classified as positive. The negative tweets were based on deaths due to COVID-19 and increased racism. The highest mean of tweets was for economic loss and the lowest mean of tweets was for travel bans and warnings. The analysis has been performed with respect to the COVID-19 pandemic, and not with respect to the psychology of the public following the lockdown.

[16] proposed a strategy for real-time estimation of risk of death from COVID-19 infection. The study relies on the exponential growth rate of incidence to estimate the basic reproduction number and confirmed case fatality risk. The results estimated a growing number of cases by the end of January suggesting the potential possibility of the infection culminating into a pandemic. The study takes into account limited empirical data confined to analyze only the case fatality risk. The veracity of the assumed date of onset is questionable and detection window time is uncertain. [31] suggested a binary classification and regression analysis methodology for investigating the COVID-19 outbreak for the sustainable development process. The classification model produces accurate results with respect to confirmed cases. The regression analysis has been used for comparing the fluctuations of parameters like wind, humidity, and average temperature. The results manifest that maximum daily temperature and relative humidity have the highest impact on confirmed cases. The case study is restricted to China, hence the number of cases considered is a fraction of the total number of cases. Hence, the results may be specific. [29] conducted a sentiment analysis of students Synchronous Online Delivery of Instruction in the Philippines as a result of community quarantine during the COVID-19 pandemic. Most of the respondents forecasted that they may face issues and

many respondents are worried about network connectivity. A negative sentiment analysis (66%) indicates that most of the students may not be able to adapt to the new trends of education. The limitation of this study is that the dataset considered is immensely small focusing on a specific institution. [34] recommended a quicker method of identifying COVID-19 using an artificial intelligence framework. The framework is based on a mobile phone survey using which data may be collected during the quarantine. The data comprises information related to an individual's location details, demographic information, travel information, signs, symptoms, etc. The only limitation of the study is that it is a device and survey-based application dependent. [20] presented a review of modern technologies that may assist in tackling the COVID-19 pandemic. The methods mentioned in the study are Diagnosis using radiology images, disease tracking, Prediction outcome of patient's health condition, Computational Biology and medicine perspective, Drug discovery, protein structure predictions, and awareness through the internet. The methods prescribed are certainly interesting, however, according to the past research works, the accuracy of the suggested methods is not 100%. [33] presented an analysis of 'Fear Injury Kidney' theory, according to which long term or excessive fear could lead to damage in the neuro-endocrine-immune system. This in turn leads to diseases, and several groups are observed to have the influence of stress and fear on the body which cannot be ignored. [27] analyzed the psychological impact of COVID-19 on the elderly population in China. More than fifteen hundred elderly people representing various ages were considered for the study. More than 37 % of seniors experienced depression and anxiety due to the pandemic. The study also shows that women experience more anxiety and depression as compared to men. The suggestions presented for such a situation included calling psychological hotlines, concentrating on seniors' mental health, and considering counseling and psychological crisis intervention. [9] presented a study on the impact of COVID-19 on the psychology of more than fifteen thousand university students. Epidemic related stressors are observed to be related to the level of anxiety. The stressors revolve around the impact of the pandemic on the economy, daily lives, academic delays, and social support. [40] explored the effect of COVID-19 lockdown on the psychology of people based on law enforcement. The enforcement of law led a large population to follow rules and restrictions during emergency situations, and quarantine people affected by the disease. The laws stated that citizens not abiding by laws would be penalized. All these impositions have been observed to cause distress among the general public (Table 1).

In this section, we presented the past relevant works performed with respect to the COVID-19 outbreak. Based on the previous work, we can say that COVID-19 research is multi-faceted. Most of the research work done aims at analyzing the outbreak and finding a cure for the same. Very little work with respect to human psychology has been done in the past. One way to explore the domain is by performing sentiment analysis. Another shortcoming of the previous research works is that many research works fail to conduct the studies globally, and many times the study is limited to specific regions and a specific number of cases. In our research, we will be handling this shortcoming by targeting data from all over the globe on a social networking platform.

2.2 Methodology

As is evident from the previous sections, a significant number of research works related to COVID-19 outbreak analysis and clinical trials have been proposed before. However, the after-effects of the lockdown on human psychology due to the pandemic still needs to be

Table 1 Summary of exiting related works

Author and proposed work	Datasets	Advantages	Limitations
[11], Twitter Sentiment Analysis	Tweets from twelve countries, between 11th March 2020 to 31st March 2020	France, Switzerland, Netherland, and US show signs of anger and distrust	The study is confined to specific countries like Australia, China, India, USA
[12], Forecast COVID-19 outbreak in China	Epidemic data from Johns Hopkins University (22 Feb 2020 to 15 March 2020)	Provides information on case fatality of China, Italy, and France	Considers a small fraction of people within three countries
[4], Risk assessment strategy for COVID-19	International Air Transport Association (IATA) and infectious disease vulnerability indexes (IDVIs)	Presents a study on the efficacy of control measures of destination countries	Considers destination countries only with respect to China
[32], Outbreak and global impacts of COVID-19	Time series data from Johns Hopkins University (22 January 2020 to 4 April 2020)	Industries affected: Tourism, Restaurants and Leisure, Entertainment, Travel, Sports	Limited industries identified for the study
[1] Analysis of Tweeters' sentiments	Extracted text from Twitter API	Most tweets are positive	The tweets are not specific to lockdown or psychological stress
[16], real-time estimation of risk of death from COVID-19	Data collected from government websites or from media quoting government announcements	The study asserts the potential possibility of the infection culminating in a pandemic	The assumed date of onset is questionable and detection window time is uncertain
[31], COVID-19 outbreak analysis	Data collected from 42 provinces in China, Japan, South Korea, and Italy	Accurate results with respect to confirmed cases	The study is confined to China only
[29], Sentiment Analysis	Social media group created for gathering data	66.55% had negative sentiments, 4.05% has positive sentiments	The study is confined to the Philippines
[34], identifying COVID-19 using Artificial Intelligence	Data collected from a mobile phone survey	Cost-Effective, efficient diagnoses and decision making	The study is device-based, survey-based, and application dependent
[20], diagnosing COVID-19	Extracted information by performing a review	Methods: radiology images, disease tracking, Prediction outcome of patient's, Computational Biology, etc	The methods may not be very accurate
[33] 'Fear Injury Kidney' theory for analyzing stress during COVID-19	Data from telephone interviews and questionnaire surveys	Fear psychology is common during the COVID-19 epidemic	The study is restricted to a small number of patients, only in China
[27], psychological impact of COVID-19	More than 1500 elderly people considered for the study	At least 37% of seniors anxiety and depression, women experience more anxiety as compared to men	The study is restricted to a small number of people, only in China
[9], psychological impact of COVID-19	More than 1500 university students	Stressors identified are economy, daily lives, academic delays, and social support	The number of students considered for the study is limited
[40], psychological impact of	No dataset (only enforcements of law)	People are affected by impositions and law enforcement	Scope of the study is limited as psychological stress

Table 1 (continued)

Author and proposed work	Datasets	Advantages	Limitations
COVID-19 on people			can be caused by many other factors

explored. To explore the psychological effects of the pandemic on people's minds, we consider the natural language processing technique of sentiment analysis. Sentiment analysis leads to the exploration of subjective opinions and feelings with respect to a particular subject by collecting data from various sources. Sentiment analysis can be of various types like fine-grained, emotion detection, aspect-based, and intent analysis. For our study, we determine the polarity of the sentences and perform aspect-based sentiment analysis. It is a classification technique that categorizes sentences into positive, negative, and neutral categories based on their polarity. Polarity or orientation is the emotion expressed in a sentence. For analyzing the sentiment of texts, there is a need to analyze individual words or phrases' sentiment. There will be several words in a sentence that are positive, and several words are negative. Presenting emotions and sentiments in a sentence is a challenge. Words like 'good' seem positive, but bear a neutral meaning when used in phrases like 'Good Morning.' Hence, positive and negative sentences are sprinkled throughout phrases. When a text is long, positive and negative sentiments tend to average towards neutral. Hence it is important to separate genuine sentiments from basic phrases. Polar words and phrases are used to denote strong and clearly defined sentiments. On the other hand, non-polar words and phrases are used in everyday language. Thus polarity approximates the sentiments within a given text. In order to determine polarity, we rely on the python library, Textblob, which is a tool built on top of the Natural Language Toolkit. Based on the polarity, it is possible to determine whether a text is positive, negative, or neutral. For TextBlog, if the polarity is greater than 0, the text is considered positive, while a text is considered negative if the polarity is lesser than 0. If the polarity of a text is equal to zero, the sentence is considered neutral. The following operations have been conducted for performing the sentiment analysis.

Step 1: Data Collection - The data is collected from appropriate sources

Step 2: Text Preprocessing- This step refers to cleaning and preparing text data. It may include several operations like noise removal (removing unwanted data), tokenization (breaking the text into smaller components for accessing each word), stopword removal (removing common words in the language that do not provide relevant information), etc.

Step 3: Calculating polarity- Using Textblob, we calculate the polarity of the words

Step 4: Calculating the total sentiment score of the text

Step 5: Sentiment Results

Figure 1 depicts the steps followed to perform the analysis.

The data has been collected from a popular global social networking forum, i.e. Twitter, and is not specific to any country or region. We consider the dates April 16, 2020, and April 30, 2020, i.e., the tweets after two weeks and four weeks of the lockdown. The steps for pre-processing of the dataset are as follows:

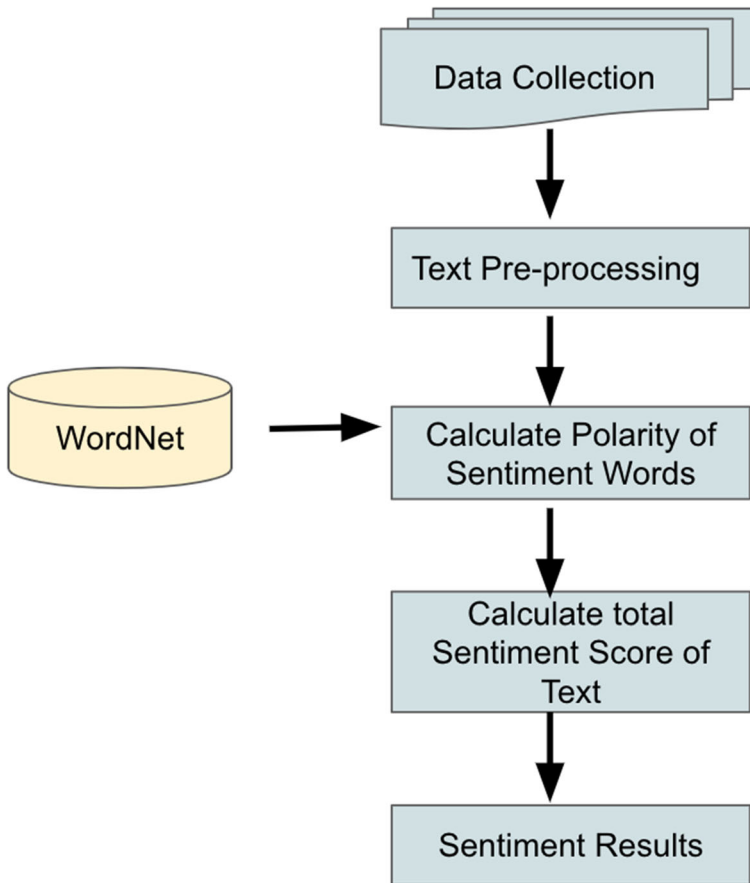


Fig. 1 Steps for performing sentiment analysis

2.2.1 Noise removal

Once the data is collected, we observe that there are a total of 21 columns namely ‘status_id’, ‘user_id’, ‘created_at’, ‘screen_name’, ‘text’, ‘source’, ‘reply_to_status_id’, ‘reply_to_user_id’, ‘reply_to_screen_name’, ‘is_quote’, ‘is_retweet’, ‘favourites_count’, ‘retweet_count’, ‘country_code’, ‘place_full_name’, ‘place_type’, ‘followers_count’, ‘friends_count’, ‘account_lang’, ‘account_created_at’, ‘verified’, ‘lang’. For our study, we do not require all these columns, hence we preprocess the data or perform noise removal. Once we remove the unwanted data, we are left with five columns namely ‘created_at’, ‘text’, ‘favourites_count’, ‘retweet_count’, ‘country_code’ and ‘lang’. Since the most commonly used language on Twitter is English, we set the language preference to English. Figure 2a and b depict the twitter data for April 16, 2020 and April 30, 2020 respectively.

2.2.2 Favorited tweets

We extract the most favorited tweets from the ‘favourites_count’ column.

	created_at	text	favourites_count	retweet_count	country_code
0	2020-04-16T00:00:00Z	G20 agrees to debt relief for poorest countrie...	1075	0	NaN
1	2020-04-16T00:00:00Z	B.C. health officials are urging people to kin...	501	0	NaN
2	2020-04-16T00:00:00Z	The Lovely Lisa Ann Joins The Show LIVE From N...	13891	0	NaN
3	2020-04-16T00:00:00Z	@allisonnamias @ArriolaLR As @ethanbdrm referen...	327	0	NaN
4	2020-04-16T00:00:00Z	Study suggests higher rates of co-infection be...	1357	130	NaN

a

	created_at	text	favourites_count	retweet_count	country_code
0	2020-04-30T00:00:00Z	Attention Seattle shoppers! Grocery stores are...	935	1	NaN
1	2020-04-30T00:00:00Z	Microsoft sees digital reboot from pandemic, p...	5448	1	NaN
2	2020-04-30T00:00:00Z	Despite #COVID19 impacts, the Victoria Street ...	507	0	NaN
3	2020-04-30T00:00:00Z	Right Now on NightSide - China Expert BU Prof....	21	1	NaN
4	2020-04-30T00:00:00Z	If there's an ideal person to provide insight ...	19559	2	NaN

b

Fig. 2 a. Tweets for April 16, 2020. b. Tweets for April 30, 2020

Three most favorited tweets on April 16, 2020.

0] Ever since the #COVID19 epidemic began i think I've only taken 2 naps ☐ the whole time that's how much my anxiety has been bothering me it's so not fun at all!! ☐☐
1] Horrible what @FashionNova is doing to people during the #COVID19 epidemic especially trying to get people to spend their stimulus checks on them that's not right!! ☐☐ https://t.co/Y51WA03UYs
2] This is a wonderful story of surviving #COVID19 and of hope https://t.co/2fhFVPV8BV

Three most favorited tweets on April 30, 2020.

0] The @WhiteHouse Gift Shop is selling "World vs. Coronavirus" coin for \$100...and the description says there's only 1,000 but somehow I feel like this is another @realDonaldTrump scam to cash in on #COVID19 Surprised it's not a "Trump vs. Coronavirus" coin tbh https://t.co/FTEKXaEKBB
1] "I don't think anybody else can feel any worse than I do about all the death." I think the family members of #coronavirus victims feel worst than you do @realDonaldTrump Stop being a narcissistic asshole making everything about yourself.
2] People under lockdown are showing their gratitude to front-line healthcare workers worldwide by applauding them. #coronavirus #COVID19 #QuedateEnCasa #StayAtHome #RestezChezVous #sanitarios @famartinez2001 https://t.co/FvYzuvv1V1

2.2.3 Retweet count

We extract the most retweeted tweets from the 'retweet_count' column.

The three most retweeted tweets on April 16, 2020.

0] I'll give \$1,000 total to 10 people (\$100 each) who retweet this within the next 48 hours. Must be following so I can DM. Hope this helps a few more people dealing with #Covid_19 or any hardships. ☐
1] House Democrats are now considering #coronavirus legislation that PAYS states to let criminals out of prison. But those same Democrats also argue that gun stores and your #2ndAmendment rights are "non-essential" to protect yourself. You can't make this up! https://t.co/GaW18YK5T7
2] #COVID19 will not stop until we put a stop to it. The time to work together is now. #TaiwanCanHelp Full text: https://t.co/8N6n3vJvdu

The three most retweeted tweets on April 30, 2020.

0] a conversation with dr raghuram rajan former rbi governor on dealing with the covid19 crisis
 1] at 9 am today you can watch my conversation with dr raghuram rajan former rbi governor on dealing with the covid19 crisis amp a new vision for india on my twitter facebook amp youtube platforms this is the first of a series of conversations with global amp indian thought leaders
 2] I've said it before, and I'll say it again — Michigan is an extraordinary place to live because of the people who call it home. There are millions of Michiganders doing their part to slow the spread of #COVID19 every single day. We are going to get through this together.

2.2.4 Number of tweets per hour

Within the span of 24 h, we can analyze the number of tweets generated.

The number of tweets per hour for April 16, 2020, and April 30, 2020, are depicted as follows (Fig. 3).

From the number of tweets generated it is evident that there are fewer tweets in the first few hours, and the number of tweets increases between the 10 to 20 h. For April 16, 2020, the maximum number of tweets per hour is close to 20,000, whereas, for April 30, 2020, the maximum number of tweets per hour is slightly more than 14,000.

2.2.5 Overall WordCloud

Word clouds are used to depict the important words in the collection of texts. One of the popular ways of communicating important information at glance is by using data visualization techniques that promote charts, infographics, graphs, etc., especially when raw data is text-based. In order to perform an in-depth analysis of text-based raw data, it may be difficult to discern which points are the most important. In such situations, word cloud generators can simplify the process. Word clouds or text clouds or tag clouds are images of words used in specific texts or subjects, such that the size of each word indicates its frequency of importance. The more a specific word appears in a source of textual data the bigger and bolder it appears in the word cloud. Word clouds are clusters of words that vary in sizes. They are appropriate for pulling the most pertinent parts of textual data. Word clouds may also be used to compare and contrast two different pieces of text to find any similarities and differences between the two. In this study, we use the word cloud data visualization for determining which words were most frequently used during the COVID-19 lockdown period across social media. Since words have

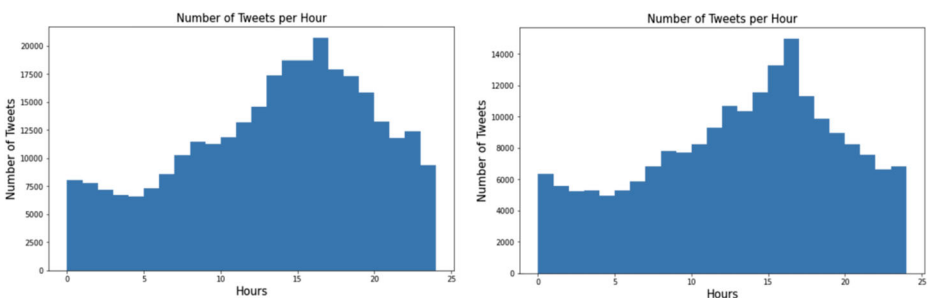


Fig. 3 Number of tweets per hour for April 16, 2020, and April 30, 2020

their own polarity, it will be beneficial in understanding the sentiments behind the tweets, and the overall population. The most important words for April 16, 2020, and April 30, 2020, are as follows (Fig. 4).

From the word clouds generated on April 16, 2020, and April 30, 2020, we observe that some words that have the highest frequency are g20, health, relief, COVID19, lockdown, pandemic, etc.

2.2.6 Data cleaning by removing stopwords

Stopwords refer to data that must be filtered out before natural language processing. In order to determine the polarity of the texts, there is a need to remove the stopwords. After removing stopwords, the data for April 16, 2020, and April 30, 2020, is as follows.

Data for April 16, 2020.

- 0 G20 agrees debt relief poorest countries amid ...
- 1 B.C. health officials urging people kind jump ...
- 2 The Lovely Lisa Ann Joins The Show LIVE From N...
- 3 @allisonnamias @ArriolaLR As @ethanbdm referen...
- 4 Study suggests higher rates co-infection #SARS...

Data for April 30, 2020.

- 0 attention seattle shoppers grocery stores work...
- 1 microsoft sees digital reboot pandemic profits...
- 2 despite covid19 impacts victoria street west p...
- 3 right now nightside china expert bu prof 8 you...
- 4 s ideal person provide insight effects covid19...

3 Experimental analysis

This section is divided into two parts. In the first section, we describe the dataset and in the second section, we perform the experimental analysis.



Fig. 4 Word Clouds depicting the important words from tweets for April 16, 2020, and April 30, 2020

3.1 Datasets

The datasets considered for the study have been acquired from Kaggle, and is referred to as ‘Coronavirus (covid19) Tweets - late April Tweets using hashtags associated with Coronavirus’. The data collected is not specific to a given country or region, rather it is taken from Twitter. The dataset incorporates the Tweets of users supported by the following hashtags: #covid19, #coronavirus, #covid_19, #ihavecorona, #coronavirusoutbreak, #coronavirusPandemic, #StayHomeStaySafe, #epitwitter, #TestTraceIsolate. In order to conduct the analysis, we consider global tweets from Twitter on April 16, 2020, and April 30, 2020. We specifically choose the dates April 16, 2020, and April 30, 2020. The lockdown restrictions were imposed during the end of March or the beginning of April, hence we analyze the tweets after two weeks and four weeks of the lockdown. As we know, lockdowns were imposed almost all over the world by the end of March or early April. It would be interesting to analyze the sentiments of people on April 16, 2020, which marks the end of two weeks after the lockdown was imposed, and April 30, 2020, which marks the end of four weeks since the lockdown was imposed. April 30, 2020, also happens to be the quadrimester of the year 2020, which means the end of the first four months and nearly a month after the worldwide lockdown was imposed. Since the Coronavirus was discovered in late December 2019, it would be interesting to analyze people’s state of mind after four months of the outbreak. We use statistical learning methods to analyze the tweets.

In order to validate the study, we consider two more datasets from Kaggle, namely Coronavirus tweets NLP - Text Classification, and COVID19 Tweets respectively. We refer to them as Dataset 2 and Dataset 3. While Dataset 2 incorporates Corona virus tagged data, Dataset 3 considers tweets with the hashtag #covid19. Dataset 2 has a total of 87,446 unique tweets, while Dataset 3 has 44,955 tweets. Some of the features for Dataset 2 are UserName, ScreenName, Location, TweetAt, OriginalTweet, etc. Dataset 3 is characterized by user_name, user_location, user_description, user_created, user_followers, user_friends, etc.

3.2 Experimental results

In this section, we present the analysis results for dates April 16, 2020, and April 30, 2020, respectively.

3.2.1 Calculating the polarity of sentences

In order to classify sentences as positive, negative, and neutral, there is a need to calculate the polarity of the texts. The polarity has been determined using Textblob. If the polarity is less than 0, texts are classified as negative. If the polarity is greater than 0, texts are classified as positive. If the polarity is equal to 0, texts are classified as neutral. After calculating the polarity, the revised data for April 16, 2020, and April 30, 2020, is as follows (Figs. 5a and 5b).

3.2.2 Displaying data categorically

Based on the analysis, we find that on April 16, 2020, the number of positive tweets is 140,084, negative tweets are 57,274, and neutral tweets are 100,855. For April 30, 2020, the number of positive tweets is 96,812, negative tweets are 36,562, and neutral tweets are 65,205. In order to depict the data categorically, count plots have been used.

	created_at	text	favourites_count	retweet_count	country_code	sentiment	polarity
0	0.0	G20 agrees debt relief poorest countries amid ...	1075	0	NaN	neutral	0
1	0.0	B.C. health officials urging people kind jump ...	501	0	NaN	positive	0.6
2	0.0	The Lovely Lisa Ann Joins The Show LIVE From N...	13891	0	NaN	positive	0.356534
3	0.0	@allisonnamias @ArriolaLR As @ethanbdrm referen...	327	0	NaN	neutral	0
4	0.0	Study suggests higher rates co-infection #SARS...	1357	130	NaN	positive	0.0416667

a

	created_at	text	favourites_count	retweet_count	country_code	sentiment	polarity
0	0.0	attention seattle shoppers grocery stores work...	935	1	NaN	positive	0.0597619
1	0.0	microsoft sees digital reboot pandemic profits...	5448	1	NaN	neutral	0
2	0.0	despite covid19 impacts victoria street west p...	507	0	NaN	neutral	0
3	0.0	right now nightside china expert bu prof 8 you...	21	1	NaN	positive	0.285714
4	0.0	s ideal person provide insight effects covid19...	19559	2	NaN	positive	0.45

b

Fig. 5 a: Data depicting polarity for April 16, 2020 tweets. b: Data depicting polarity for April 30, 2020 tweets

From Fig. 6, it is evident that there are more positive tweets in both the cases, i.e., for April 16, 2020, and April 30, 2020, respectively. This is followed by the number of neural tweets and negative tweets in both cases respectively.

3.2.3 Sentiment distribution of data

A sentiment distribution plot can identify how polarity is distributed. The range varies from -1 to +1. The Sentiment distribution for April 16, 2020, and April 30, 2020, are as follows (Fig. 7).

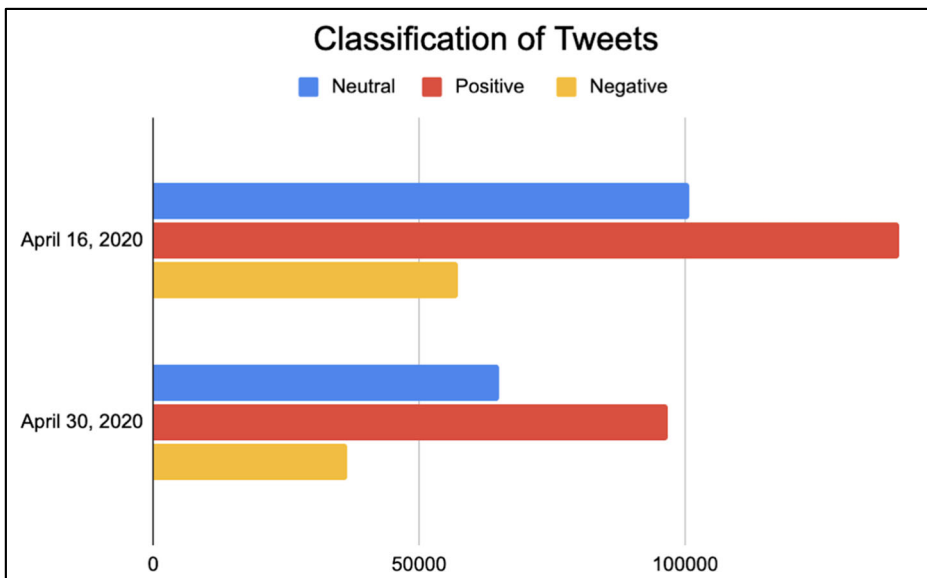


Fig. 6 Classification of tweets for April 16, 2020, and April 30, 2020

	word	frequency		word	frequency
0	covid19	192513	0	covid19	145850
1	coronavirus	105578	1	coronavirus	69637
2	covid	69919	2	s	42820
3	19	63414	3	covid	31033
4	s	63233	4	amp	27718

Fig. 11 Most common words for April 16, 2020, and April 30, 2020

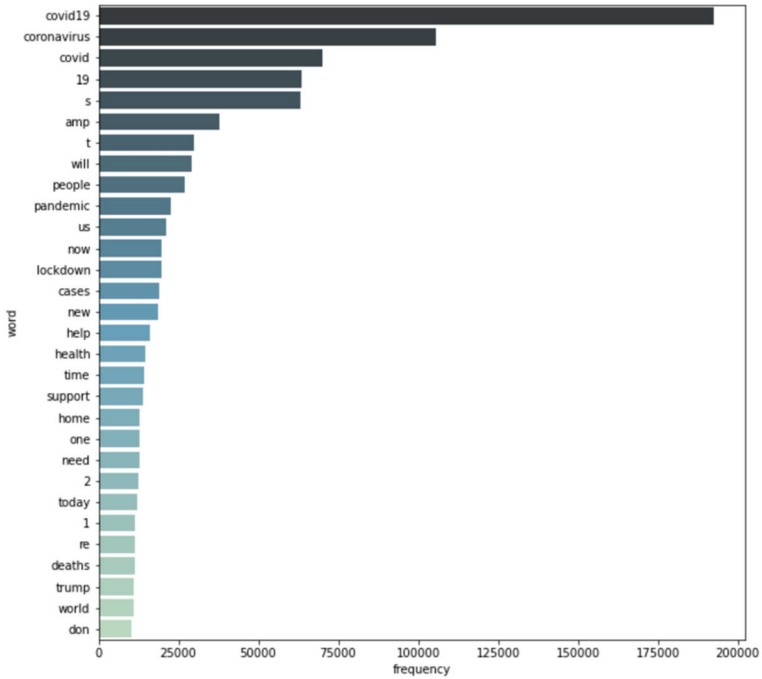
that for Dataset 3, the number of positive tweets is the highest followed by the number of neutral tweets and negative tweets respectively. In both the datasets, i.e., Dataset 2 and Dataset 3, the number of positive tweets is the highest. This coincides with the initial study performed in which the number of positive tweets is the highest with respect to negative tweets and neutral tweets for April 16, 2020, and April 30, 2020, respectively. Thus, the study is validated.

4 Discussions

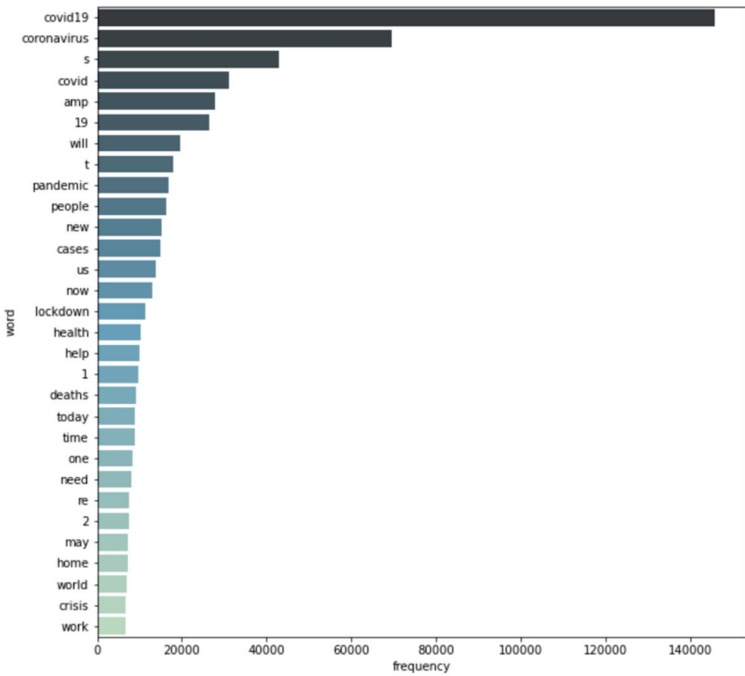
Based on the experimental analysis and the results obtained we observed the twitter data for dates April 16, 2020, and April 30, 2020. While the number of tweets with respect to the dataset considered are more for April 16, 2020, with respect to April 30, 2020, it is observed that in both the cases positive tweets take the lead. For April 16, 2020, a total of 140,084 positive tweets were observed with respect to 96,812 positive tweets on April 30, 2020. For both, the dates, negative tweets are low with respect to positive tweets. For April 16, 2020, the number of negative tweets is 57,274 whereas, for April 30, 2020, the number of negative tweets is 36,562. Therefore, observing the huge difference in the number of positive and negative tweets of both the dates it is appropriate to say that people have a positive sentiment towards the pandemic in the situation of the lockdown. The same is verified on Dataset 2 and Dataset 3, respectively. Although the number of positive tweets is lower in April 30, 2020, there is a significant decrease in the number of negative tweets. The negative sentiments that exist in such a situation may be due to several causes ranging from inability to socialize to unemployment during the lockdown. Hence a greater number of positive sentiments are indicative of a healthy state of mind of the public. We provide a comparative analysis of our proposed work with some existing works (Table 2).

Based on the comparative analysis, we assert that

- While many research works target prediction and cure as their objectives, little is mentioned about human psychology and sentiment analysis. In this paper, we attempt to understand human psychology during the lockdown using sentiment analysis.



a



b

Fig. 12 a: Barplot depicting the frequency of words for April 16, 2020. **b:** Barplot depicting the frequency of words for April 30, 2020

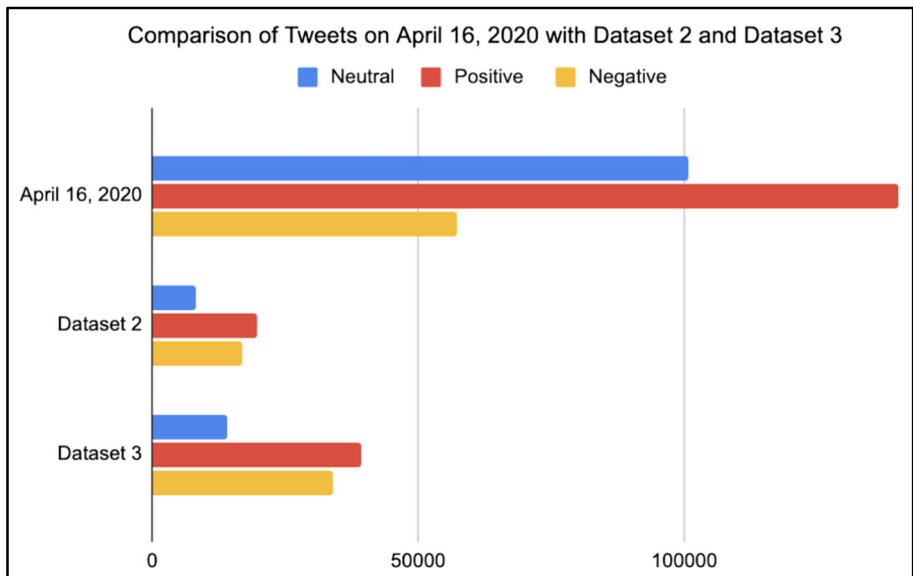
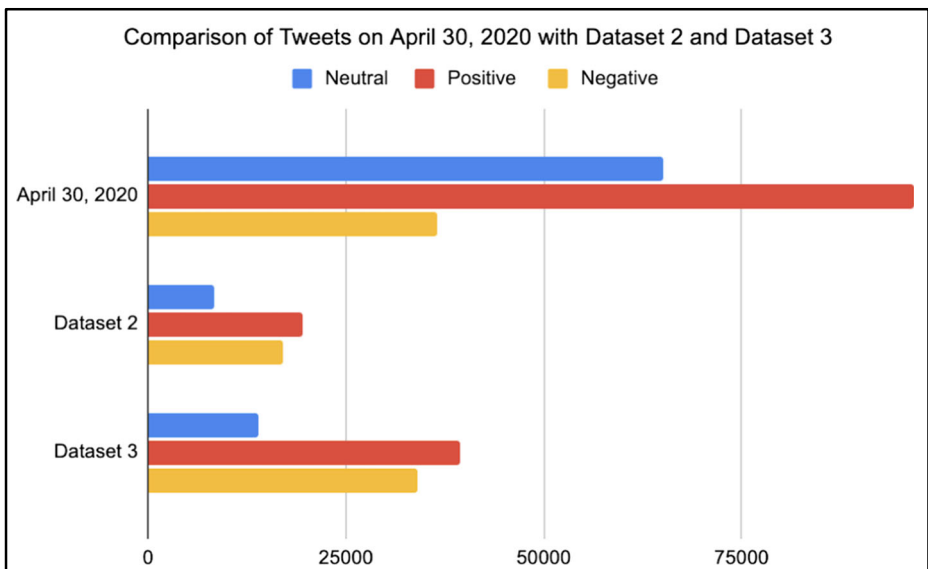
**a****b**

Fig. 13 a: Comparison of tweets on April 16, 2020, with Dataset 2 and Dataset 3. **13b:** Comparison of Tweets on April 30, 2020, with Dataset 2 and Dataset 3

- The analysis has been conducted by acquiring tweets from the global social networking platform, Twitter. Since the most common language used over this platform is English, the tweets we are analyzing are also in English.

Table 2 Comparative analysis of our proposed work with existing works

Author and year	Proposed work	Methodology/ parameters	Results
[8]	Sentiment analysis and Emotion detection during COVID-19 pandemic in Spain	Analyzing application programming interface (API) and web scraping techniques	The study identified emotions like disgust, sadness, fear, anger over different points in time during the pandemic, however, the study is restricted to one country, i.e., Spain
[43]	Analyzed negative sentiment on social media in China	BERT (Bidirectional Encoder Representations from Transformers) model for sentence classification	“Gamey Food” and “Bat” are primary assumptions, however the study is restricted to only China
[29]	Sentiment analysis of Online Delivery of Instructions during COVID-19 Quarantine in the Philippines	Machine Learning Techniques	66.55% had negative sentiments, 4.05% has positive sentiments, the study is confined to the Philippines
[2]	Sentiment Analysis of nationwide lockdown for India	Natural language Processing	The study identified several emotions like anger, disgust, fear, joy, etc., however, it is limited to a specific country (India) whereas, COVID-19 is a global issue.
[21]	Analyzing COVID-19 Sentiment impact on US Stock market	Time series regression models on Daily News Sentiment Index (DNSI) and Google Trends	There is a correlation between COVID-19 sentiment and eleven industries (Healthcare, Materials, Estate, etc), study is confined to the US stock market
[28]	Sentiment Analysis of Egyptian Students during COVID-19 pandemic	Word2vec and Machine Learning techniques	Naive Bayes is the most accurate at 87%. The decision tree is the least accurate at 76%.
[35]	Tweets Classification for Coronavirus Tweets of varying lengths	Naive Bayes and Logistic Regression	91% accuracy for Naive Bayes and 74% accuracy for Logistic Regression
Our Proposed Work	Sentiment Analysis after two and four weeks of worldwide lockdown restriction	Natural Language Processing, Textblob	For both the days, people all over the world have positive sentiments followed by neutral and negative sentiments respectively.

- We have specifically chosen the dates April 16, 2020, and April 30, 2020, which are approximately two and four weeks after the worldwide lockdown was imposed. Thus we were able to compare the sentiments of people using the two dates.
- COVID-19 is a global issue. Most of the research works pertaining to COVID-19 sentiment analysis are confined to specific countries. This paper performs sentiment analysis on global data for the first time to the best of our knowledge.

Based on the analysis, with the majority of positive texts for all the datasets considered, it is evident that people exhibit a positive attitude towards the pandemic during the lockdown. It is not surprising that the effects of the lockdown can be seen in the general public devoting much of their time to multimedia devices. With restrictions on socializing many people spend their time playing games or watching videos during the lockdown. This may plant the thoughts of

loneliness and depression in individuals. However, multimedia also offers various methods of socializing and interacting virtually through gaming platforms, social media sites, and video platforms. Although increased use of multimedia may affect people psychologically, the increased number of positive tweets is indicative of an optimistic attitude that people have exhibited during the lockdown. Some limitations of the study are as follows.

1. The study is restricted to comprehending the psychological state of people who use Twitter. Other social media platforms and networking sites may include the public who are not Twitter users.
2. The study does not consider the psychology and sentiments of other communities and groups like daily wage workers or healthcare staff if they do not use social media.
3. There are other languages like Chinese, Spanish, Arabic, etc., that are frequently used on social media platforms like Twitter. There is a large population of Twitter users relying on these languages. We have analyzed the tweets in English for this study.

5 Conclusion and future works

In this paper, we performed sentiment analysis of Twitter data specified during the lockdown due to the COVID-19 pandemic. We chose two dates, i.e., April 16, 2020, and April 30, 2020, which marks two and four weeks after the worldwide lockdown was imposed. The idea behind choosing the dates two weeks apart was to analyze the sentiments of people all over the world due to the lockdown. The lockdown led to many businesses closing down, economic crises, and suicides among the general public. Moreover, people being dependent on multimedia devices for spending their time during lockdown may also serve as a victim of severe psychological effects like loneliness and depression. Hence there is a need to analyze the psychology of the human mind during a situation like this. Based on the sentiment analysis conducted, we observed that the number of positive tweets was more in both the cases depicting a positive attitude of people towards the pandemic during the lockdown. In the future, we would like to broaden the study and perform emotion detection like anger, disgust, fear, joy, anxiety, panic, sadness, sarcasm, etc. using global data. Moreover, an integrated analysis may be conducted to associate sentiments to a psychological profile, behaviors, demographic characteristics, transactions, events, and/or other data.

References

1. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z (2020) Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. *J Med Internet Res* 22(4):e19016
2. Barkur G, Vibha GBK (2020) Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: evidence from India. *Asian J Psychiatr*
3. Block P, Hoffman M, Raabe IJ, Dowd JB, Rahal C, Kashyap R, Mills MC (2020) Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat Hum Behav*:1–9
4. Boldog P, Tekeli T, Vizi Z, Dénes A, Bartha FA, Röst G (2020) Risk assessment of novel coronavirus COVID-19 outbreaks outside China. *J Clin Med* 9(2):571
5. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, Zheng J (2020) The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry Res* 112934

6. Cucinotta D, Vanelli M (2020) WHO declares COVID-19 a pandemic. *Acta bio-medica: Atenei Parmensis* 91(1):157–160
7. Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I, Le DN (2020) Forecasting confirmed and fatalities cases of COVID-19 outbreak. *Front Public Health* 8:592
8. de Las Heras-Pedrosa C, Sánchez-Núñez P, Peláez JI (2020) Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. *Int J Environ Res Public Health* 17(15):5542
9. Dhar BK, Ayithey FK, Sarkar SM (2020) Impact of COVID-19 on psychology among the university students. *Global Chall* 4(11):38–44
10. Dsouza DD, Quadros S, Hyderabadwala ZI, Mamun MA (2020) Aggregated COVID-19 suicide incidences in India: fear of COVID-19 infection is the prominent causative factor. *Psychiatry Res*:113145
11. Dubey, Akash Dutt, Twitter Sentiment Analysis during COVID-19 Outbreak (April 9, 2020). Available at SSRN: <https://ssrn.com/abstract=3572023> or <https://doi.org/10.2139/ssrn.3572023>
12. Fanelli D, Piazza F (2020) Analysis and forecast of COVID-19 spreading in China, Italy, and France. *Chaos, Solitons Fractals* 134:109761
13. Gautret P, Lagier JC, Parola P, Meddeb L, Sevestre J, Mailhe M, Hocquart M (2020) Clinical and microbiological effect of a combination of hydroxychloroquine and azithromycin in 80 COVID-19 patients with at least a six-day follow up: a pilot observational study. *Travel Med Infect Dis* 101663
14. Iyengar K, Upadhyaya GK, Vaishya R, Jain V (2020) COVID-19 and applications of smartphone technology in the current pandemic. *Diabetes Metab Syndr Clin Res Rev* 14(5):733–737
15. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z (2020) Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *J Gen Intern Med*:1–5
16. Jung SM, Akhmetzhanov AR, Hayashi K, Linton NM, Yang Y, Yuan B, Nishiura H (2020) Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: inference using exported cases. *J Clin Med* 9(2):523
17. Kazak AE (2020) Psychology is an essential science: American psychologist highlights the role of psychology in understanding and addressing COVID-19. *Am Psychol* 75(5):605–606
18. Ke YY, Peng TT, Yeh TK, Huang WZ, Chang SE, Wu SH, Lin WH (2020) Artificial intelligence approach fighting COVID-19 with repurposing drugs. *Biom J* 43(4):85–96
19. Koh D (2020) COVID-19 lockdowns throughout the world. *Occup Med*:47–54
20. Kumar A, Gupta PK, Srivastava A (2020) A review of modern technologies for tackling COVID-19 pandemic. *Diabetes Metab Syndr Clin Res Rev* 14(4):569–573
21. Lee HS (2020) Exploring the initial impact of COVID-19 sentiment on US stock market using big data. *Sustainability* 12(16):6648
22. Lewnard JA, Lo NC (2020) Scientific and ethical basis for social-distancing interventions against COVID-19. *Lancet Infect Dis* 20(6):631
23. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Xing X (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N Engl J Med* 382:1199–1207
24. Linka K, Peirlinck M, Sahli Costabal F, Kuhl E (2020) Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. *Comput Methods Biomech Biomed Eng* 23(11):710–717
25. Lowenthal P, Borup J, West R, Archambault L (2020) Thinking beyond zoom: using asynchronous video to maintain connection and engagement during the COVID-19 pandemic. *J Technol Teach Educ* 28(2):383–391
26. Mahato S, Pal S, Ghosh KG (2020) Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India. *Sci Total Environ* 139086
27. Meng H, Xu Y, Dai J, Zhang Y, Liu B, Yang H (2020) Analyze the psychological impact of COVID-19 among the elderly population in China and make corresponding suggestions. *Psychiatry Res* 289:112983
28. Mostafa, L. (2020, October). Egyptian student sentiment analysis using Word2vec during the coronavirus (Covid-19) pandemic. In International Conference on Advanced Intelligent Systems and Informatics (pp. 195–203). Springer, Cham.
29. Pastor CKL (2020) Sentiment analysis on synchronous online delivery of instruction due to extreme community quarantine in the Philippines caused by the Covid-19 pandemic. *Asia J Multidiscip Stud* 3(1):1–6
30. Petrosillo N, Viceconte G, Ergonul O, Ippolito G, Petersen E (2020) COVID-19, SARS, and MERS: are they closely related? *Clin Microbiol Infect* 26:729–734
31. Pirouz B, Shaffiee Haghshenas S, Haghshenas S, S.& Piro, P. (2020) Investigating a serious challenge in the sustainable development process: analysis of confirmed cases of COVID-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. *Sustainability* 12:2427
32. Priyadarshini I, Mohanty P, Kumar R, Son LH, Chau HTM, Nhu V-H, Thi Ngo PT, Tien Bui D (2020) Analysis of outbreak and global impacts of the COVID-19. *Healthcare* 8:148

33. Qianwei LIU, Guozhen ZHAO, Bo JI, Yitian LIU, ZHANG J, Qiujie MOU, Tianyu SHI (2020) Analysis of the influence of the psychology changes of fear induced by the COVID-19 epidemic on the body. *World J Acupunct Moxibustion* 30(2):85–89
34. Rao ASS, Vazquez JA (2020) Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey when cities and towns are under quarantine. *Infect Control Hosp Epidemiol* 41(7):826–830
35. Samuel J, Ali GG, Rahman M, Esawi E, Samuel Y (2020) Covid-19 public sentiment insights and machine learning for tweets classification. *Information* 11(6):314
36. Sarma P, Kaur H, Kumar H, Mahendru D, Avti P, Bhattacharyya A, Singh A (2020) Virological and clinical cure in COVID-19 patients treated with hydroxychloroquine: a systematic review and meta-analysis. *J Med Virol* 92:776–785
37. Serrano G, Kochergina I, Albors A, Diaz E, Oroval M, Hueso G, Serrano JM (2020) Liposomal lactoferrin as potential preventative and cure for COVID-19. *Int. J. Res. Health Sci* 8:8–15
38. Sharif A, Aloui C, Yarovaya L (2020) COVID-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet-based approach. *Int Rev Financ Anal*:101496
39. Suryawanshi R, More V (2020) A study of the effect of Corona virus Covid-19 and lock down on human psychology of the Pune City region. *Studies in Indian Place Names* 40(70):984–994
40. Varalakshmi R, Swetha R (2020) Covid-19 lock down: people psychology due to law enforcement. *Asian J Psychiatr* 51:102102
41. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, Booy R (2020) School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *The Lancet Child & Adolescent Health*
42. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Tien Bui D, Priyadarshini I, Sarkar M, Le T (2020) Crime rate detection using social media of different crime locations and twitter part-of-speech tagger with Brown clustering. *J Intell Fuzzy Syst*, (Preprint):1–13
43. Wang T, Lu K, Chow KP, Zhu Q (2020) COVID-19 sensing: negative sentiment analysis on social media in China via Bert model. *IEEE Access* 8:138162–138169
44. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Yuan ML (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579(7798):265–269
45. Yamaguchi K, Takebayashi Y, Miyamae M, Komazawa A, Yokoyama C, Ito M (2020) Role of focusing on the positive side during COVID-19 outbreak: mental health perspective from positive psychology. *Psychol Trauma Theory Res Pract Policy* 12(S1):S49
46. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Chen HD (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798):270–273

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Ishaani Priyadarshini¹ · Pinaki Mohanty² · Raghvendra Kumar³ · Rohit Sharma⁴ · Vikram Puri⁵ · Pradeep Kumar Singh⁶

Ishaani Priyadarshini
ishaani@udel.edu

Pinaki Mohanty
mohantypinakiranjan@gmail.com

Raghvendra Kumar
raghvendraagrawal7@gmail.com

Rohit Sharma
rohitapece@gmail.com

Vikram Puri
purivikram@duytan.edu.vn

¹ Department of Electrical and Computer Science, University of Delaware, Newark, DE 19716, USA

² Department of Computer Science, Purdue University, 610 Purdue Mall, West Lafayette, IN 47907, USA

³ Department of Computer Science and Engineering, GIET University, Gunupur, India

⁴ Department of Electronics & Communication Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, NCR Campus, Delhi- NCR Campus, Delhi-Meerut Road, Modinagar, Ghaziabad, UP, India

⁵ Center of Visualization and Simulation, Duy Tan University, Da Nang, Vietnam

⁶ ABES Engineering College, Ghaziabad, UP, India