# Twitter Sentiment Analysis of COVID-19 pandemic during lockdown phases in India

**Rashmi R[1], Shobana T S[2], Monisha H M[3] Assistant Professor, Department of ISE, BMS College of Engineering, Bangalore**

## Abstract

With the advancement of web innovation and its development, there is an immense volume of information present in the web. Social media like Instagram, Facebook and Quora provide the platform to share and express their views. Twitter is one of the powerful social media platforms. Tweets from twitter are studied to measure the sentiments of Indians towards the lockdown. This paper emphasis mainly on sentiment analysis of COVID19 twitter data during different stages of lockdown in India. The sentimental quotients from the gathered tweets have been inferred by evaluating the positive, neutral and negative polarities.

**Keywords:** *COVID19, Lockdown, twitter, sentimental analysis, India, Emotions*.

## I. Introduction

The world has witnessed various kinds of pandemic such as SARS, MERS in the past. Now the mankind is going through a very difficult time in fighting against deadly coronavirus (COVID-19). The COVID-19 is a new virus which belongs to the same family as Severe Acute Respiratory Syndrome (SARS) and some types of common cold. First coronavirus case had been witnessed in Wuhan, China in December 2019 and now it has spread across almost all the countries. This is an invisible enemy which is spreading very fast around the globe. The novel coronavirus has created an immense chaos around the globe which has become a threat to the human's lives. This has caused number of deaths around the world. Most of the countries are striving hard to find the solution to get rid of this pandemic as soon as possible. In many nations, government presented numerous guidelines planned for hindering the spread of COVID-19.

World Health Organization (WHO) announced Coronavirus (COVID-19) as pandemic on March 11th, 2020 due to its drastic spread across many nations like Italy, Spain, USA, Iran and other European countries. This pandemic has continued to affect many lives due to which lockdown was embraced in many countries and India was one among them.

Though it was a tough decision, it was inevitable as India being a second largest populated country in the world, with a population of 1.3 Billion, Government of India took a major decision of lockdown to avoid the community spread. There were different phases of lockdown adopted in India as shown table1.

| Phase | Hashtags | Period |
|-------|----------|--------|
| 1 | #Lockdown1.0 | 26$^{th}$ March - 14$^{th}$ April |
| 2 | #Lockdown2.0 | 15$^{th}$ April - 3$^{rd}$ May |
| 3 | #Lockdown3.0 | 4$^{th}$ May - 17$^{th}$ May |
| 4 | #Lockdown4.0 | 18$^{th}$ May – 31$^{st}$ May |

Table1: Different Lockdown phases in India

Lockdown in India has helped in avoiding the rapid increase in coronavirus cases. The figure1 shows the COVID-19 cases across India during lockdown phases.
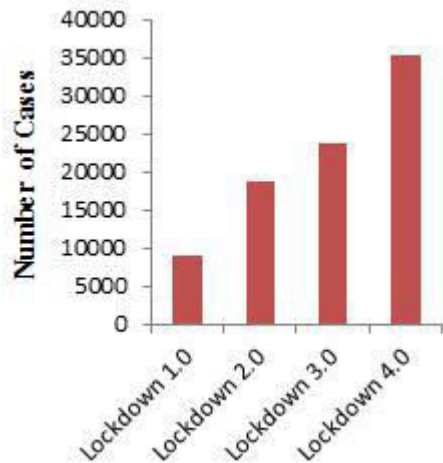


Figure 1: COVID19 Cases reported during different Lockdown phases

Social media is a platform where people can make their opinions by posting their thoughts. Twitter is one of the robust social media platforms to make public opinions about various ongoing topics. It plays an important role in exchanging useful information and in expressing their views. This paper uses twitter data for the sentimental analysis of the people of India.

Sentiment analysis is a means to measure the sentimental content of writing. It can help to analyze the perceptions in a text associated to a specific subject. Sentiment analysis is an extremely fascinating field of research that can support in gathering the emotional content of a text.

This paper is structured in following manner: summary of literature survey is presented in section II. Two stages of Data analysis and methodology are explained in section III. Results are covered in section IV. Section V concludes the work.

## II. Related Work

[1] Gopalkrishna Barkur, Vibha, Giridhar B. Kamath have described about COVID19 pandemic and national wide lockdown in India. Twitter platform was used for analysing the sentiments of Indian citizen with respect to hashtag #IndiaLockdown and #IndiafightsCorona with the lockdown announcements. R tool was used for generating the results, tweets were represented using wordcloud and graph was used for representing the sentiments.

[2] Akhila Sri Manasa Venigallaa, Dheeraj Vagavolub and Sridhar Chimalakondac have considered twitter is one the social media application that acts as a bridge between the citizen and government in various countries. Based on twitter data a web portal was presented by them to categorize the mood of citizen of India with respect to COVID19. The 7 categories were classified by them that include 6 emotions and 1 neutral category.

[3] Palash Ghosh, Rik Ghosh and Bibhas Chakraborty have analysed the spread of coronavirus disease 2019 in various states of India. Maharashtra and West Bengal states were on major focus for predicting the infected people. Logistics model and Susceptible Infectious Susceptible (SIS) Growth model were considered for predicting the infectious people.

[4] Aliza Sarlan, Chayanit Nadam, Shuib Basri have used one of the social media application i.e. twitter application for sentimental analysis. They have designed and analysed the customer perception towards success in marketplace. They analysed the positive and negative tweets by representing the result in pie chart and html page.

[5] Xing Fang and Justin Zhan have stated about different versions of API that are available in twitter. They have taken the NLP as the major task in opinion mining or sentimental analysis to tackle one the fundamental problem of sentimental polarity categorization in sentimental analysis.

## III. Data Analysis and Methodology

People expressed their feelings and view about the current pandemic and their sentiments during the lockdown in various social media platforms. One such platform where people exchange their opinions and feelings is twitter. In this paper, twitter data have been gathered to analyze the sentiments of the people during various lockdowns in India.

The methods for this study consist of two distinct stages: data collection and data analysis (Stage 1) and visualization (Stage 2).

### Stage1: Data Collection and Data Analysis

In order to gather the tweets from twitter API, one needs to create an app through their twitter account. After authentication consumer key, secret key, access key and access token were collected from twitter. 1000 tweets were fetched for each of the hashtags #Lockdown1.0, #Lockdown2.0, #Lockdown3.0 and #Lockdown4.0 during the period as shown in table1. Data collected from the twitter are unstructured. Data preparation is a way of collecting and organizing the data in a structured manner. Data cleaning plays a major role in preparing the data for sentiment analysis because data obtained from real-time have raw, noisy ,duplicate entries, blank spaces and emoji's. This is a prominent step because the quality of the data will fetch

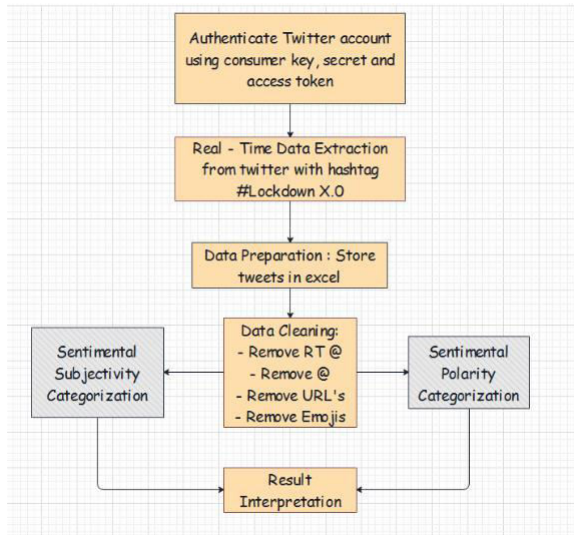more reliable results as shown in figure 2 and algorithm.



Figure 2: Flow diagram of twitter sentiment analysis

Data is analyzed for diverse feelings for different hashtags. Sentimental analysis refers to the general method to extract polarity and subjectivity from semantic orientation which refers to the strength of words and polarity text or phrases. Polarity and subjectivity estimates are given by the Textblob. The estimates are provided for the parsed documents using dictionaries given by the pattern library. Polarity scores ranges from -1 to 1, where 1 means positive text and -1 means a negative text. Subjectivity scores ranges from 0 to 1 which refers to personal opinion, emotions and judgement. Polarity and subjectivity scores were generated and calculated using (1)

$$\text{Score} = \textit{No. of positive words} - \textit{No. of negative words.} \quad \text{--------- (1)}$$

Once the score are calculated for each hashtags, sentiments are classified as positive, negative and neutral sentiments

---------------------------------------------------

Algorithm: Sentimental Analysis for #Lockdown X.0

---------------------------------------------------

```
Input: Twitter Data

Output: Sentiment analysis graph (number of tweets versus Sentiment classifier)

1: consumerKey ← Consumer key of twitter account
2: consumerSecret ← Consumer secret key of twitter account
3: accessToken ← Access token of twitter account
4: accessTokenSecret ← Access token secret from twitter
5: For each lockdown stages:
6: do
7:      search_word= "#Lockdown X.0"
8:      date_since= "YYYY-MM-DD"
9:      Extract Twitter Data: tweets =tweepy.Cursor(api.search,
10:                                      q=search_words,
11:                                      lang="en",
12:                                      since=date_since).items(10000)
13:     Create the data frame:
14:        tweet_df=pd.DataFrame(data=tweet_details,columns=['geo','text','user','location'])
15:     For each tweet:
16:     do
17:        Data Cleaning: text = re.sub("RT@[\w]*:","",text)
18:                       text = re.sub("@[\w]*","",text)
17:                       text = re.sub("https?://[A-Za-z)-9./]*","",text)
18:                       text = re.sub("\n","",text)
19:     end
19:     tweet_df['text']=tweet_df['text'].apply(lambda x: clean_tweets(x))
20:     Save to csv: tweet_df.to_csv('filename.csv')
21:     def getSubjectivity(text):
22:          return TextBlob(text).sentiment.subjectivity
23:     def getPolarity(text):
24:          return  TextBlob(text).sentiment.polarity
25:     Generate wordCloud = WordCloud(width=500, height=300, random_state=21,
        max_font_size=110). generate(allWords)
26:     def getAnalysis(score):
27:     if score < 0:
28:        return 'Negative'
29:      elif score == 0:
30:        return 'Neutral'
31:     else:
32:        return 'Positive'
33:     tweet_df['Analysis']  = tweet_df['Polarity'].apply(getAnalysis)
26:     Generate Scatter plot for sentiment classifier
27: end
```

**Stage 2: Visualization of word frequency**
Word level sentiment analysis is done using Word cloud. The word clouds were generated depicting the frequency of all words observed in the text of tweets. It is a perception wherein the most consistent words show up in large size and the less continuous words show up in tinier sizes.

Figure 3 shows the word frequency of #lockdown1.0. It also consists of some negativity expressed by the people and prominent opinion of the people was positive.
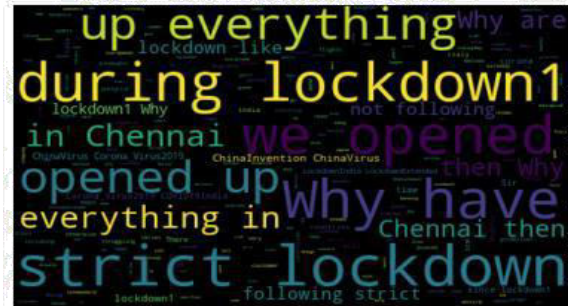

Figure3: Word Cloud for Lockdown 1

Figure 4 & 5 shows word frequency of #Lockdown2.0 and #Lockdown3.0 respectively. Understanding the necessity of social distancing from previous phase, positive sentiment stood out during the Lockdown 2 and 3.
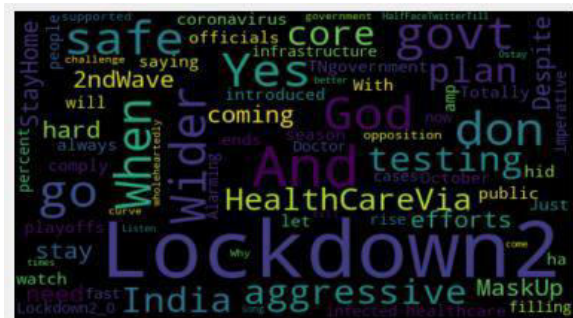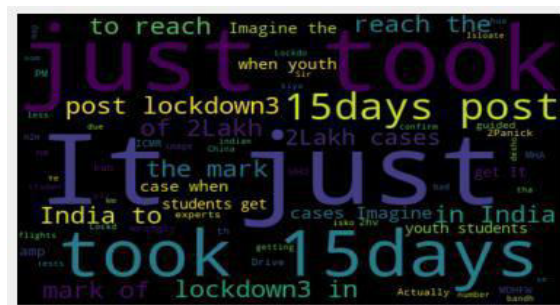

Figure4: Word Cloud for Lockdown 2


Figure5: Word Cloud for Lockdown 3

Figure 6 shows word frequency of #Lockdown4.0. Even though people faced many problems during previous lockdowns they expressed positive thoughts during this phase. People raised negative sentiments due to the relaxation in restrictions, as they realized strict lockdown was the best solution to fight against coronavirus.
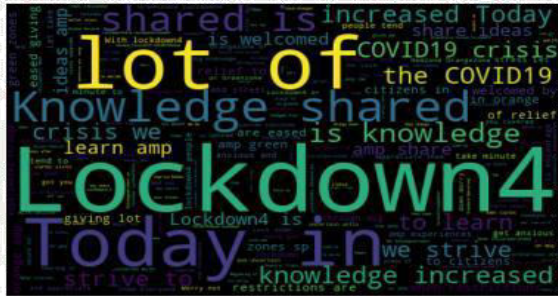
Figure6: Word Cloud for Lockdown 4

## IV. Results

For each hashtags, 10000 tweets were collected and separate datasets were generated. Results of sentimental analysis scores are depicted using pie chart which represents positive, negative and neutral emotions of people of India during Lockdown phases. The pie chart is divided into 3 parts- blue represents the maximum sentiment scores, red denotes the medium sentiment scores and yellow signifies the minimum sentiment scores.

Figure 7 shows the pie chart for phase 1 Lockdown where the maximum sentiments were shown up to be positive.
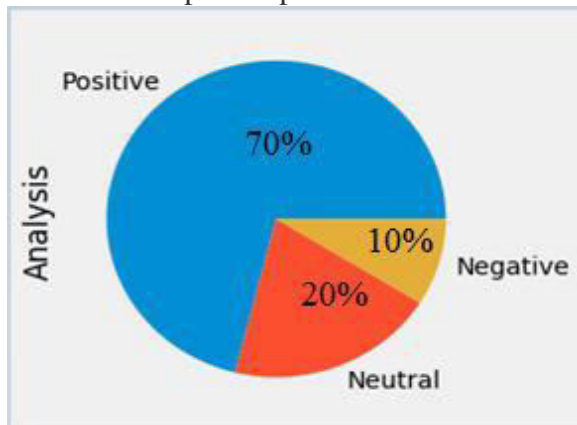

Figure7: Pie chart for Lockdown 1

Figure 8 shows the pie chart for phase 2 Lockdown during this phase sentiments among the people shifted to neutral.
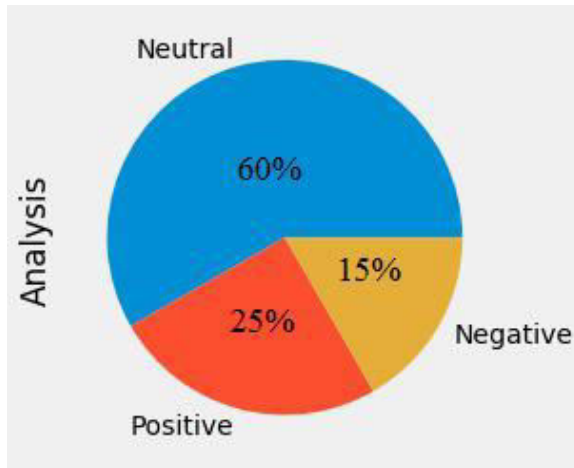

Figure8: Pie chart for Lockdown 2

Figure 9 shows the pie chart for phase 3 Lockdown. During this phase sentiments among the people shifted to neutral. Drastic increase in negative emotion where shown up due relaxation of rules.
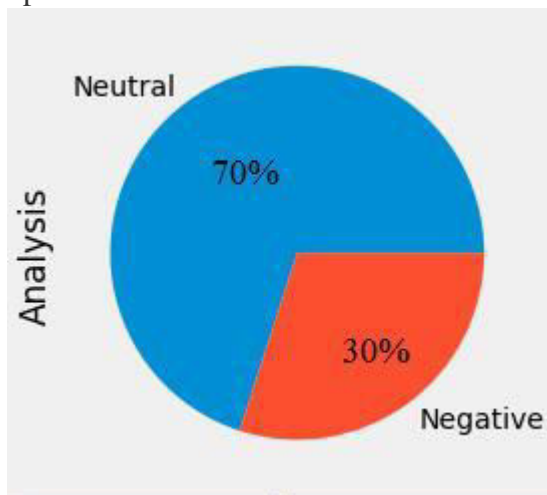

Figure9: Pie chart for Lockdown 3

Figure 10 shows the pie chart for phase 4 Lockdown. India's economy was hindered due to constant lockdown. To improve the economy and other aspects, Government of India decided to ease the lockdown rules. So, people expressed mixed feeling during this phase.
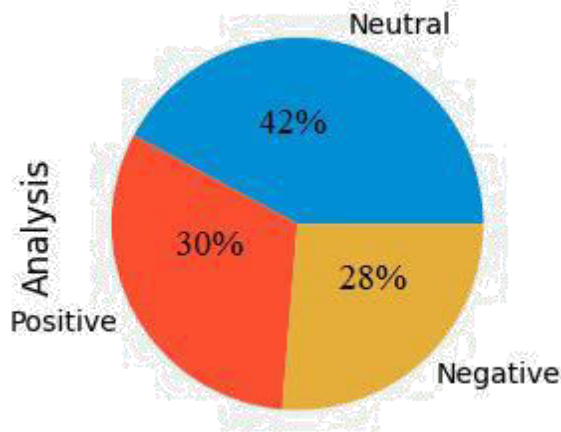
Figure10: Pie chart for Lockdown 4

## V. Conclusion

In general, people of India have shown the positive response towards Lockdown's in fighting against COVID-19 to flatten the curve. It could be seen from the tweets, few people have expressed their negative opinion regarding the relaxation of Lockdown rules. Overall, majority of the people have comprehended the significance of the Lockdown to battle against this pandemic, since it is one of the best solutions for reducing the fast spread of coronavirus.

## Bibliography

[1]     G. Barkur and G. B. Kamath, "Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information website . Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories , such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source . These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active . Sentiment analysis of nationwide lockdown due to COVID 19 outbreak : Evidence from India," no. January, 2020.

[2]     A. S. M. Venigalla, D. Vagavolu, and S. Chimalakonda, "Mood of India During Covid-19 -- An Interactive Web Portal Based on Emotion Analysis of Twitter Data," *arXiv*, 2020, [Online]. Available: http://arxiv.org/abs/2005.02955.

[3]     P. Ghosh, R. Ghosh, and B. Chakraborty, "COVID-19 in India: State-wise Analysis and Prediction," *medRxiv*,no.April,p. 2020.04.24.20077792, 2020, doi: 10.1101/2020.04.24.20077792.

[4]     A. Sarlan, C. Nadam, and S. Basri, "Twitter sentiment analysis," *Conf. Proc. - 6th Int. Conf. Inf. Technol. Multimed. UNITEN Cultiv. Creat. Enabling Technol. Through Internet Things, ICIMU 2014*, no. November 2016, pp. 212–216, 2015, doi: 10.1109/ICIMU.2014.7066632.

[5]     X. Fang and J. Zhan, "Sentiment analysis using product review data," *J. Big Data*, vol. 2, no. 1, 2015, doi: 10.1186/s40537-015-0015-2.