ORIGINAL RESEARCH PAPER

Media

COVID-19 PANDEMIC: A STUDY ON SENTIMENTS

KEY WORDS: Coronavirus, COVID-19, #Covid19, social media, twitter

Dr. Sarla Bhardwaj*	Associate Professor, Department of Mathematics, Bhim Rao Ambedkar college, University of Delhi. *Corresponding Author
Neha Singh	Department of Information Technology, Maharaja Agrasen Institute of Technology, GGSIP University.
Dr. Bishnu Mohan Dash	Associate Professor, Department of Social Work, Bhim Rao Ambedkar college University of Delhi, Delhi-110094.

ABSTRACT

COVID-19 global pandemic has created an unprecedented crisis and the entire global community passing through turbulent times. With the number of coronavirus (COVID-19) cases growing exponentially, the entire world has come to a standstill. This Covid-19 global pandemic has created stressful unsettling circumstances for the whole world. It has completely disrupted the normal socio economic activities in the whole world. Most of the countries in the world have enforced a complete lockdown and taking several pro-active measures and necessary precautions to ensure health and safety of its citizens. In this critical juncture people use technology and other electronic means not only to stay connected, work from home but also use it as a major sources of sharing their opinions through social media. People have flooded their social media accounts with their opinion on Covid-19 and lockdown. In this paper, 9 lacs tweets from all over the world are extracted and analysed. Tweets are extracted for the month of March 2020 and April 2020. The findings show that almost half of the people are optimistic about the lockdown and have a positive opinion that they will overcome the situation. Around 2.76 lac people have negative opinions about lockdown and Covid-19 and around 2.1 lac people are neutral in their opinion. A detailed comparison of tweets for months of March 2020 and April 2020 is done. The comparison proves that there is an increase in Negative opinion in April 2020 as compared to March 2020. A comparison of word-cloud of March 2020 and April 2020 concludes a shift of frequency of words from Europe to USA.

INTRODUCTION

World Health Organization has declared Coronavirus as global pandemic on 11th March 2020. (WHO, 2020). Almost one third of the world's total population in under lockdown due to pandemic. Many countries have closed their borders to prevent international travelers in order to halt the spread of the virus. The fear and panic among the people across the globe is spreading at a much faster rate. The entire business houses in the world have come to a standstill and the economy of the country is in jeopardy and has affected negatively in each and every sectors of the economy and has caused massive disruption in the socio-economic life everywhere in the world. It has taken the lives of more than two lakh people the world over. Work from home has become the work culture globally leading to the exponential growth in the use of internet services. Internet is not only seen as an important medium of work including the teaching and learning process but also as an important source of entertainment. Twitters and other social media platforms have become prime outlet to express the views and sentiments particularly on issues related to coronavirus.

In this study, the researchers have extracted a dataset of 9 lakhs of tweets with the help of tweepy between 4th March 2020 to 27th April, 2020 related to coronavirus. The sentiments of the tweets were checked and analyzed by Vader[1] (Valence Aware Dictionary and Sentiment Reasoner)

The paper is divided into following sections.

Literature Review

In this section, the researchers have listed some of the relevant literature related for the study. Silva &Hui (2003) study discusses the implementation details of a real time facial feature extraction and emotion reaction system by using edge counting and image-correlation optical flow techniques to calculate the local motion vectors of facial feature. The study also determined emotional state of a subject using a neural network. The main objective of the paper is the real-time implementation of a facial emotion recognition system and the study relied on conventional off-line calculations which were modified to make real-time implementation possible. Besides that the study had also used a variety of improved techniques to the general algorithms such as edge focusing, global motion cancellation are proposed and implemented

Kaur et.al(2010) has tried to locate a new idea for detecting an unknown human face in input imagery and recognizing his/her facial expression. The study has used Principal Component Analysis (PCA) with Singular value decomposition (SVD) for Feature Extraction to determine principal emotions. The experiments show that the proposed facial expression recognition framework yields relatively little degradation in recognition rate due to facial images wearing glasses or loss of feature points during tracking. Sharma et.al(2019) finds the sentiments of the tweets on cybersecurity in different countries. The paper of Waterloo et.al(2018) describes how people on the basis of different demographic profiles such as ages, sex, and other demographic variables present themselves with different emotions on different social media sites viz. Twitter, Instagram, Facebook and WhatsApp. In another study, the authors have shown operating process of Face API of Microsoft Azure. (Verma et.al,2019). Jung et.al(2016) reported advantages of using different face-APIs which include Face++, IBM Bluemix Visual Recognition, and various other face. In another study You, et.al (2016) have built a dataset which is the biggest dataset even bigger than the dataset used at that time with 3+ million weakly labelled data which consists of images of different emotions which were used for image classification while the study of Khanal et.al,2016 checks how Google and Microsoft face API works for different picture emotion datasets. The study of Kang et.al (2019) shows a method for finding images with different emotions concerning different geographic places. Pardàs and Bonafonte(2002) finds the emotion of human beings in videos. In their paper Boy and Uitermark, 2017 analysed the dataset of 400000 posts on Instagram from Amsterdam and presents how the city is reassembled on the platform. Another study shows a neural network which helps in discovering the face of a person in the picture (Rowley, 1998). Bhardwaj et.al(2018) in their study analyses the sentiments of tweets that are found with the hashtag of TripAdvisor from twitter to understand the opinion of people in different places. In another study Bhardwaj et.al(2016) reported how fast the data is being created and how tough it is to store. A study on sentiments analysis of twitters shows the reliability of the website out of all travel websites(Bhardwaj et.al,2017). Hutto and Gilbert(2014) in their study.

Methodology and Experimental Setup

1.1 Data-Extraction and Cleaning

1. Extracting tweets from 4th March 2020 to 27th April 2020. Almost 9 lac tweets were extracted with the help of tweepy.
2. Cleaning dataset of every day in March 2020 and April 2020 by deleting the tweets that were not in English and did not have country code with it or did not have a proper place mentioned.
3. Concatenating all the excel files of each day to one single file so that we have one single file.
4. Cleaning the text of tweets by Lemmatisation and Stemming of words in text
5. Removing stop-words from the text.

1.2 Dataset

The final dataset for the study consists of 9 lac tweets bearing country-code and country name as well. This dataset was gathered between 4th march to 27th April. This dataset included all the tweets available in English language with the name of the country given.

Figure 1: World Map of Countries considered for research in March 2020

Figure 2: World Map of Countries considered for research in April 2020

Figure 1 and 2 represents the world map of countries on which the research has been conducted. All the green coloured countries were included in the dataset for the purpose of study. The countries in grey were not considered as the researchers could not find any tweets written in English language during the period of collection of data

1.3 Sentiment Analysis

Sentiment analysis is done on VADER. It is a rule-based tool that is used for analysis of the sentiments. It is very accurate in predicting sentiments available in social media. **VADER** uses a list of features which are lexical and are labelled with respect to their orientation that can be positive or negative. **VADER** also accurately presents the percentage of positivity and negativity in the sentence.

1.4 Graphical Representation

All the world-maps were created from software called Tableau and the study has used python libraries such as matplotlib and seaborn to perform data visualisation on the dataset. Matplotlib is used to create bar-charts and word cloud library is used to create wordcloud of each type of sentiment.

Results Analysis and Comparison

The work is executed in Spyder IDE for python. Python is an interpreted, high-level, general-purpose programming language. The first phase of work is to extract the data and pre-process it. After pre-processing, a cleaned dataset is prepared for sentiment analysis. This cleaned dataset is used for sentiment analysis using Vader. Bar-charts and world density maps are made to depict the analysis. Word-cloud of most frequently used words in each sentiment is also depicted

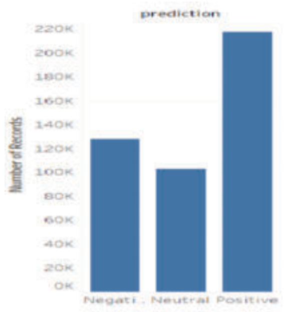

Fig. 3: Sentiment Analysis of Tweets in March 2020

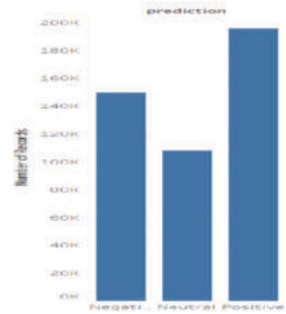

Fig. 4: Sentiment Analysis of Tweets in April 2020

Figure 3 shows sentiment analysis of tweets. A total of 4.46 lac tweets were considered for evaluation out of which, 2.17 lac tweets were positive, 1.27 lac tweets were negative and 1.02 lac tweets were neutral. A majority of positive tweets reflects that people are optimistic with fight against #Covid19 and believe that sooner or later they will over the global pandemic .

Figure 4 shows sentiment analysis of tweets. A total of 4.53 lac tweets were considered for evaluation out of which, 1.95 lac tweets were positive ,1.49 lac tweets were negative and 1.08 lac tweets were neutral. A majority of positive tweets proves that people are optimistic with fight against #Covid19 but as the days are passing by the negativity in tweets is also increasing as it can be seen from the comparison between tweets in the month of March and April, 2020.

Fig.5: Distribution of tweets for various countries in March 2020

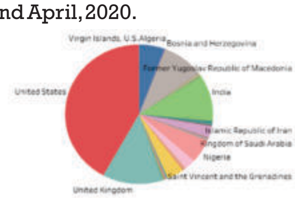

Fig.6: Distribution of tweets for various countries in April 2020

Fig. 5 depicts the distribution of tweets from various countries in the month of March, 2020. The findings shows that almost half of the total tweets, (2,13,956 tweets), were from USA which is one of the most affected country in the world. The second most number of tweets were from UK (62,869) followed by India and Canada that is 35,272 and 25,854 tweets respectively.

Fig. 6 depicts the distribution of tweets from various countries in the month of April, 2020. About fifty percent of the tweets, (1,90,117) were from USA one of the worst affected country of Covid-19 in the world: USA. The second most number of tweets(59,998) were reported from UK followed by India(47,993) and Canada(47,993) respectively.

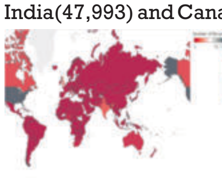

Fig. 7: Density of Negative tweets in March 2020

Fig. 8: Density of Negative tweets in April 2020

Fig. 7 presents the density of negative tweets for different countries in world. USA reported maximum number of negative tweets (62,717). The findings also shows that India, UK and Canada which are coloured in pink reported less number of negative tweets as compared to USA. The number of negative tweets of India, UK and Canada are 9576, 17697, 6968 respectively. that the countries Nigeria and Australia coloured in light red colour reported less number of tweets i.e. 3170 and 3193 tweets respectively.

Fig. 8 presents the density of negative tweets for different countries in world. USA again reported the maximum number of negative tweets which is 67419. Here we can see that India, UK and Canada are pink which seen lesser number of tweets 14197 , 19719 , 12210 respectively. It can also be seen that Nigeria and Australia earmarked in lighter versions of red seen 7876 and 8692 tweets respectively.

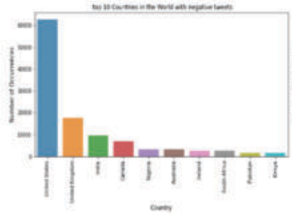

Fig. 9: Top 10 countries with Negative Tweets in March 2020

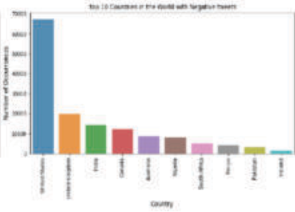

Fig. 10: Top 10 countries with Negative Tweets in April 2020

Fig. 9 shows top 10 countries in the world with negative tweets. It can be seen that USA is at the top with 62717 tweets followed by UK, India, Canada and Nigeria with 17697, 9576, 6968, 3193 negative tweets respectively.

Fig. 10 shows top 10 countries with negative tweets. It can be seen that USA is at the top with 67419 tweets followed by UK, India, Canada and Australia with 19719, 14197, 12210, 8,692 negative tweets respectively.

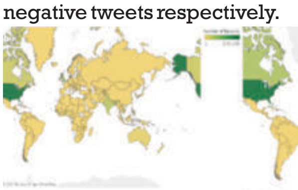

Fig. 11: Density of Positive tweets in March 2020

Fig. 12: Density of Positive tweets in April 2020

Fig. 11 presents the density of positive tweets for different countries in world. USA reported to have the maximum number of positive tweets followed by India, UK and Canada with 18179, 33044, 13115 positive tweets respectively. The study shows that Nigeria and Ireland are lighter versions of yellow as they have 6256 and 5222 tweets respectively.

Fig. 12 elaborates on the density of positive tweets for different countries in world. USA again reported has maximum number of positive tweets followed by India, UK and Canada with 22354, 26930, 18237 tweets respectively. The study also reveals that Nigeria and Australia shown in lighter versions of yellow as they have witnessed 10331 and 11357 tweets respectively.

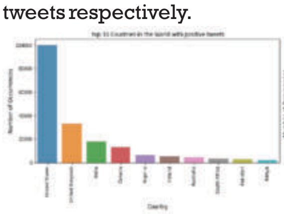

Fig. 13: Top 10 countries with Positive Tweets in March 2020

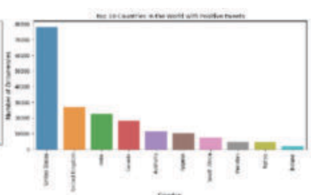

Fig. 14: Top 10 countries with Positive Tweets in April 2020

Fig. 13 depicts top 10 countries with Positive Tweets. The figure reveals that USA is at the top with 100238 tweets. It is followed by UK, India, Canada and Nigeria with 33044, 18179, 13115, 6256 positive tweets respectively.

Fig. 14 depicts top 10 countries with Positive Tweets which again reveals that USA is at the top with 78058 tweets. It is followed by UK, India, Canada and Australia with 26930, 22354, 18237, 11375 positive tweets respectively.

Fig. 15: Density of Neutral tweets in March 2020

Fig. 16: Density of Neutral tweets in April 2020

Similarly the study has attempted to examine the density of neutral tweets for different countries. Fig. 15 and 16 depicts the density of neutral tweets for different countries in world. The findings show that USA has witnessed the maximum density of tweets. It is also shown that India, UK and Canada are coloured light purple because they have low density of neutral tweet. It is very important to mention that Nigeria and Australia which are presented in lighter versions of blue have reported lesser density of tweets.

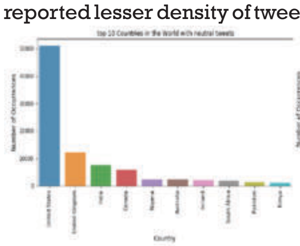

Fig. 17: Top 10 countries with Neutral Tweets in March 2020

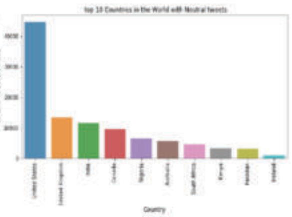

Fig. 18: Top 10 countries with Neutral Tweets in April 2020

Fig. 17 depicts top 10 countries with Neutral Tweets. Again, USA is found to be at the top in the list a highest number of 51001 tweets followed by UK, India, Canada and Nigeria with 12128, 7517, 5771, 2368 neutral tweets respectively .

Fig. 18 depicts top 10 countries with Neutral Tweets. Again, a similar trend was seen. USA is found to be at the top with 44640 tweets followed by UK, India, Canada and Nigeria with 13349, 11442, 9551, 6553 neutral tweets respectively.

Fig. 19: Word-cloud of Neutral tweets in March 2020

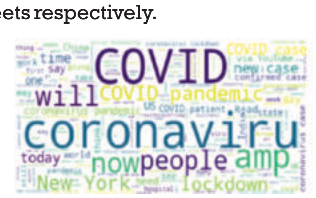

Fig. 20: Word-cloud of Neutral tweets in April 2020

The study also tried to examine the size of the tweets and frequency of words usage in the tweets. Fig. 19 shows words with maximum frequency in Neutral Tweets. The bigger the size of the words, the greater the frequency of words in the tweets. It is observed that words like Covid19, coronavirus, breaking, confirmed, , work, break, US, joke are the words used frequently in neutral tweets .

Fig. 20 shows the words with maximum frequency in Neutral Tweets. It is also observed that words like Covid, coronavirus, COVID, NEW YORK, lockdown, case, patient, US, people are the words used very frequently in neutral tweets.

Fig. 21: Word-cloud of Positive tweets in March 2020

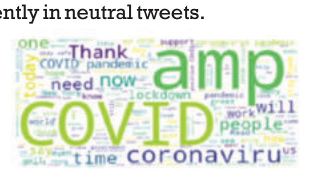

Fig. 22: Word-cloud of Positive tweets in April 2020

The Fig. 21 clearly show words with maximum frequency in case of Positive Tweets. It revealed that the bigger the size of the words, the greater is the frequency of words in the tweets. Words like coronavirus, covid19, Italy, trying, stand, positive, great, friend are the words being used most frequently in positive tweets .

Fig. 22 also show words with maximum frequency in Positive Tweets. It also reported the same nature of the trends. Like the previous figure, it was also found that the bigger the size of the words, the greater the frequency of words in the tweets. The figure finds that the Words like coronavirus, COVID, Thank, support, need, now, US are the words that were being used frequently in positive tweets .

Fig. 23: Word-cloud of Negative tweets in March 2020

Fig. 24: Word-cloud of Negative tweets in April 2020

Fig 23 presents the word-cloud of Negative Tweets. It also found that similar results. The frequency of the words in the tweets depends upon the size of the words. It found that the bigger the size of the words, the greater the frequency of words in the tweets. The analysis shows that words like gathering, covid19, virus, spread, death, serious, precautionary and Afghanistan are words being used frequently in the tweets .

Likewise, the Fig 24 also shows the word-cloud of Negative Tweets. The same results were seen as above as frequency of words in the tweets was greater due to the bigger size of the words. Bigger the size of the word greater the frequency of word in the tweets. Words like crisis, pandemic, death toll, Trump, death and Covid death are words being used most frequently.

CONCLUSION

The study has tried to clearly depict analysis of sentiments of the people towards Covid-19 from different parts of the world in this unsettling times of global pandemic. On the basis of comparison of word-cloud of March 2020 and April 2020, it shows the shift of frequency of words from Europe to USA. The study also finds that almost half of the people are optimistic about the lockdown and have a positive opinion that they will overcome the crisis and conquer the situation. Despite such odds and challenges, many people are optimistic that such uncertainty, turbulent times will soon pass and the world will be back on track. The study also found that around 2.76 lac people have negative opinions about lockdown and COVID19 and around 2.1 lac people neutral opinion in this regard. The findings of the study also shows an increase in negative opinion in April 2020 as compared to March 2020. As the days are passing by, although as most of the people are optimistic about conquering the situation, the amount of negative tweets are also growing among the people.

REFERENCES

1. Bhardwaj, P., Gautam, S., & Pahwa, P. (2018). A novel approach to analyze the sentiments of tweets related to TripAdvisor. Journal of Information and Optimization Sciences, 39(2), 591-605.
2. Bhardwaj, P., Gautam, S., Pahwa, P., & Singh, N. (2016). Characteristics and challenges of big data. Int. J. Recent Innov. Trends Comput. Commun, 5, 187-190.
3. Bhardwaj, P., Gautam, S., & Pahwa, P. (2017). Opinion Mining and Sentiment Analysis of Travel Websites through Twitter. International Journal of Applied Engineering Research, 12(22), 12431-12439.
4. Boy, J. D., & Uitermark, J. (2017). Reassembling the city through Instagram. *Transactions of the Institute of British Geographers*, *42*(4), 612-624.
5. De Silva, L. C., & Hui, S. C. (2003, December). Real-time facial feature extraction and emotion recognition. In *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint* (Vol. 3, pp. 1310-1314). IEEE.
6. Hutto, C. J., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.
7. Jung, S. G., An, J., Kwak, H., Salminen, J., & Jansen, B. J. (2018, June). Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In *Twelfth International AAAI Conference on Web and Social Media*.
8. Kaur, M., Vashisht, R., & Neeru, N. (2010). Recognition of facial expressions with principal component analysis and singular value decomposition. *International Journal of Computer Applications*, *9*(12), 36-40.
9. Kartikay Sharma, Siddharth Bhasin, Piyush Bharadwaj(2019, September). A Worldwide Analysis of Cyber Security And Cyber Crime using Twitter. In International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6S3.
10. Khanal, S. R., Barroso, J., Lopes, N., Sampaio, J., & Filipe, V. (2018, June). Performance analysis of Microsoft's and Google's Emotion Recognition API using pose-invariant faces. In *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion* (pp. 172-178).
11. Kang, Y., Jia, Q., Gao, S., Zeng, X., Wang, Y., Angsuesser, S., ... & Fei, T. (2019). Extracting human emotions at different places based on facial expressions and spatial clustering analysis. *Transactions in GIS*, *23*(3), 450-480.
12. Pardàs, M., & Bonafonte, A. (2002). Facial animation parameters extraction and expression recognition using Hidden Markov Models. *Signal Processing: Image Communication*, *17*(9), 675-688.
13. Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence*, *20*(1), 23-38.
14. Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2018). Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *new media & society*, *20*(5), 1813-1831.
15. Verma, A., Malla, D., Choudhary, A. K., & Arora, V. (2019, February). A Detailed Study of Azure Platform & Its Cognitive Services. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 129-134). IEEE.
16. You, Q., Luo, J., Jin, H., & Yang, J. (2016, February). Building a large scale dataset for image emotion recognition: The fine print and the benchmark. In *Thirtieth AAAI conference on artificial intelligence*.