World's Emotional Response to Pandemic: Sentiment Analysis of Twitter Data During the COVID-19 Pandemic

Mahdi Abolfazli Esfahani (✉ mahdi001@e.ntu.edu.sg)

Nanyang Technological University (NTU)

Messiah Abolfazli Esfahani

Ferdowsi University of Mashhad (FUM)

Han Wang

Nanyang Technological University (NTU)

Research Article

Keywords: twitter, COVID-19, trigeminal neuralgia, emotional response

Posted Date: April 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-400235/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

World's Emotional Response to Pandemic: Sentiment Analysis of Twitter Data During the COVID-19 Pandemic

Mahdi Abolfazli Esfahani¹, Messiah Abolfazli Esfahani^{2,3}, and Han Wang¹

¹Nanyang Technological University (NTU), Department of Electrical and Electronics Engineering, Singapore, 639798, Singapore

²Ferdowsi University of Mashhad (FUM), Department of Computer Engineering, Mashhad, 9177948974, Iran

³Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 138671, Singapore

ABSTRACT

The 2019 novel coronavirus disease, shortly named COVID-19, has dramatically changed people's daily life and a new type of life has been forced on people, called new normal life. While the new normal life helped get COVID-19 growth under control, the side effects are still preserved in the societies, such as stress and fear. The existence of anxiety for an extended period results in various illnesses such as trigeminal neuralgia, and if the period gets much longer, it will increase the suicide rate. Once the number of COVID-19 cases spikes in a country, the government imposes strict policies such as complete lock down and quarantine orders, which helps control the disease's spread significantly. However, it substantially increases people's stress and fear. Hence, understanding the people's emotional response to pandemics over time is essential and needs to be considered for analyzing the possibility of increasing restrictions and being alarmed once easing restrictions is required. Artificial Intelligence can help countries analyze people's emotional responses in different states and recommend governments on the appropriate level of measurement control rules to prevent disease spread and keep people well psychologically. This paper aims to study people's emotional responses to the COVID-19 pandemic by performing a sentiment analysis on social media. Twitter, which is a popular social media, is considered in this research to perform analysis. The results show a significant drop in people's positiveness during 2020, in which COVID-19 spiked, compared to 2019.

Introduction

The 2019 novel coronavirus disease, named COVID-19, started to spread out in December 2019 as an acute disease with a high transmission rate. The majority of patients have reported fever, cough, and shortness of breath symptoms¹. The virus can be transmitted through the air and physical interaction easily². Hence, various measurement controls are introduced by governments to take the spread of the virus under control; for instance, using masks, as the most significant protective equipment, has been made mandatory in most countries³, and countries were required to impose full lockdown and quarantine orders periodically⁴. Various arrangements assist in preventing widespread of COVID-19 in communities, and the essential ones can be mentioned as:

- Fast detection of infected people and moving them to dedicated facilities to serve quarantine and get treatment.
- Social responsibility (e.g., monitoring temperatures at the shops' entrance and arranging enough space between customers).
- Imposing social lockdown and quarantine orders to keep the community safe from undetected COVID-19 cases.

While the mentioned resolutions can help control the virus's spread, they would require assistance from recent advances in artificial intelligence to be safe, effective, robust, accurate, and be done at the right time. For instance, detecting COVID-19 from computed tomography (CT)⁵, and X-ray⁶ images is complex, and physicians have a limited speed to check and validate all accurately. Various deep learning-based techniques have been proposed since last year for fast and robust detection of infected patients from X-ray⁷ and CT images⁸, which have significantly assisted in the rapid detection of COVID-19 cases. Various smart devices have also been produced to detect high-temperature people in the environment⁹⁻¹¹ and prevent large groups' formation through surveillance cameras¹².

Despite utilizing advanced technologies (e.g., for temperature screening and monitoring social distancing) to control the spread of the virus, sudden spikes have been seen in the number of COVID-19 cases worldwide due to the presence of

undetected COVID-19 cases that have no symptoms. To flatten the curve of COVID-19 case numbers, governments usually apply strict control measurements requiring people to stay in their household, and no large gatherings are allowed, which significantly impacts flattening the curve⁴. However, as discussed in recent researches^{13,14}, full lockdown for making the number of COVID-19 cases zero in a one long period will result in substantial spreading waves of COVID-19. Governments need to gradually release restrictions once the number of new daily COVID-19 cases reduced and reached a steady-state, and they need to update policies periodically.

Using strict measurement controls, such as lockdown and not allowing people to have social visits, for an extended period can result in secondary effects (e.g., poverty¹⁵, mental illnesses¹⁶). Hence, one important parameter that needs to be considered by governments while they are updating restrictions is the people's psychological well-being. This paper aims to study this parameter by analyzing the data generated by people online in social media over time and obtain the world's emotional responses to the pandemic. In this regard, Twitter, a popular social media, is considered to perform this study. Compared to previous studies¹⁷⁻¹⁹, this study collects and compares all tweets rather than only comparing tweets related to COVID-19 as most people do not use hashtags or mention pandemics in their tweets (text messages in Twitter). Pandemic contains side effects on economics and the whole life, and this research aims to analyze emotional response during a pandemic, so all tweets need to be considered. Furthermore, the analysis is done between the years 2019 and 2020 for the first time in this paper among five populated cities of five countries.

Methods

United States of America (USA), Canada, Australia, and the United Kingdom (UK) are considered to perform this study. These countries are mainly selected due to the popularity of Twitter in those areas and using English as their primary language. The top 5 populated cities in each of the countries mentioned above are selected due to sufficient data in those areas. The populated cities of each of the mentioned countries studied in this research can be named as:

- **USA:** New York City, Los Angeles, Chicago, Houston, and Phoenix
- **Canada:** Toronto, Montreal, Vancouver, Calgary, and Edmonton
- **UK:** London, Birmingham, Manchester, Leeds, and Glasgow
- **Australia:** Sydney, Melbourne, Brisbane, Perth, and Adelaide

Twitter's original application programming interface (API) has various limitations on collecting historical data. Hence, the historical data from Twitter is scraped via the open-source Twitter Intelligence Tool²⁰. Seventy thousand daily tweets limit is set, and historical data is gathered daily. In this manner, data is not biased towards specific dates (e.g., the beginning of a month), and data is gathered uniformly. After analyzing the amount of collected data over different years, we have observed a sudden increase in the amount of gathered data since April 2019, and a low amount of data is scraped before that; it is mainly due to the limitations on scraping Twitter data. Hence, the data collected since April 2019 is considered to perform research and do analysis. To have a better sense of the total amount of data, the USA's total gathered data, as an instance, during 2019 and 2020, is more than 44 million tweets. A Tweet is a microblog message which is limited to 280 characters and may contain images. After gathering the data, the dataset is cleaned by removing URLs and retweets (re-posting of tweets).

The valence aware dictionary for sentiment reasoning (VADER)²¹ approach is the state-of-the-art approach for sentiment analysis of microblog data and is utilized in this research to perform analysis. VADER is a fast lexicon and rule-based sentiment analysis tool that relies on a dictionary that maps words and other lexical features common to microblogs' sentiment expression. VADER even understands and considers emojis and non-conventional texts (e.g., the difference between sad and SAD, and ? and ???). Vader outputs the compound score, which is computed by summing the valence scores of each word in the lexicon, adjusted to the rules, and then normalized in the range of [-1,+1], where -1 stands for the most extreme negative sentence, and +1 stands for the most extreme positive sentence. Each of the countries' response to the COVID-19 pandemic is analyzed by comparing the averaged compound score between 2019 and 2020.

Results

In order to perform analysis, the rolling average of the compound score is computed over seven days and visualized in the Figures 1-4. Going deeper into the Figures and mapping measurements with the real world news, we observed that the minimum compound score is related to the kill of George Floyd in the USA, which gained lots of attention globally²². Besides, analyzing the results show that the weekly average of the compound score in 2020 is usually less than the 2019 average value. The average compound value in each year is computed to have a better sense, and the result can be seen in both Figure 5 and Table 1. In general, the compound score's value is reduced, and the most significant drop is related to the USA by %-14.89. While

UK is one of the worst countries in controlling COVID-19 and huge spikes have been seen there, it can be seen that the drop is negligible, which shows that UK's people are more dispassionate and have less stress during the pandemic. The countries can be ordered based on their emotional level, from positive to negative, as the UK, Australia, Canada, and the USA, which is aligned with Gallup group emotional report²³ which has studied the emotions of people in various countries through different factors.

Doing a case study over the seven-day rolling average of the compound score in Manchester UK (Figure 3c), a huge drop in the compound score can be seen in November 2020, which is due to the entrance of Manchester to the lockdown at the end of October²⁴. However, it recovered soon, as 2021 was on the way, and people got to celebrating it over social media, and people had hope of a better year ahead. Similar behavior can be seen in Calgary, Canada (Figure 2d), where the number of COVID-19 cases has spiked during October and November²⁵.

Further analysis is applied by studying the 30 most frequent adjectives over 2019 and 2020 in different countries. The investigation has shown that in

- USA, positive words such as "amazing" and "beautiful" are not in the top 30 most frequent adjectives of 2020 while they existed in 2019. Overall, there is no change in the number of positive or negative adjectives, but the frequency of positive words has been reduced, and some have changed their place with other adjectives with a lower level of positiveness.
- Canada, the words "amazing" and "beautiful" are among the top 30 most frequent adjectives of 2019, while they didn't exist in the top 30 most frequent adjectives of 2020. However, they are mainly changed with other positive adjectives. Overall, the number of positive adjectives that exist in the top 30 most frequent adjectives is reduced by one in 2020, and the number of negative adjectives has been increased by one.
- UK, the frequency of positive adjectives has been reduced, and the top 30 most frequent adjectives did not change.
- Australia, the number of positive adjectives in the top 30 most frequent adjectives has been decreased by one, and their frequency has reduced significantly.

The word cloud related to the top 100 most frequent adjectives and nouns in different countries can be seen in the Figure 6 which demonstrates important information; For instance, in USA people had more focus on "money" and "party" which cannot be seen in 2020.

Overall, the result shows a significant drop in emotion in terms of compound score, and change in the words and adjectives people utilize during a pandemic. Based on the results, governments can use Twitter to extract different states' emotional responses to the pandemic by looking into the seven-day rolling average compound score and most frequent adjectives utilized. Considering the extracted information can help to simultaneously control the growth of pandemics and keep people healthy psychologically.

(a) New York City

(b) Los Angeles

(c) Chicago

(d) Houston

(e) Phoenix

Figure 1. Seven-Day Rolling Average of Compound Score in USA's Cities

(a) Toronto

(b) Montreal

(c) Vancouver

(d) Calgary

(e) Edmonton

Figure 2. Seven-Day Rolling Average of Compound Score in Canada's Cities

(a) London

(b) Birmingham

(c) Manchester

(d) Leeds

(e) Glasgow

Figure 3. Seven-Day Rolling Average of Compound Score in UK's Cities

(a) Sydney

(b) Melbourne

(c) Brisbane

(d) Perth

(e) Adelaide

Figure 4. Seven-Day Rolling Average of Compound Score in Australia's Cities

Figure 5. Yearly Average of Compound Score in Different Countries Cities, and The Entire Country.

	2019 Compound Score	2020 Compound Score	% Percentage of Change
US	0.159906	0.136091	-14.89
Canada	0.179872	0.156473	-13.00
Australia	0.192896	0.174852	-9.35
UK	0.219931	0.212914	-3.19

Table 1. Yearly Average of Compound Score, and Relative Change.

References

1. Tian, S. *et al.* Pulmonary pathology of early-phase 2019 novel coronavirus (covid-19) pneumonia in two patients with lung cancer. *J. thoracic oncology* **15**, 700–704 (2020).
2. Peng, X. *et al.* Transmission routes of 2019-ncov and controls in dental practice. *Int. journal oral science* **12**, 1–6 (2020).
3. Organization, W. H. *et al.* Rational use of personal protective equipment (ppe) for coronavirus disease (covid-19): interim guidance, 19 march 2020. Tech. Rep., World Health Organization (2020).
4. Tang, K. H. D. Movement control as an effective measure against covid-19 spread in malaysia: an overview. *J. Public Heal.* 1–4 (2020).
5. Jacobi, A., Chung, M., Bernheim, A. & Eber, C. Portable chest x-ray in coronavirus disease-19 (covid-19): A pictorial review. *Clin. imaging* (2020).
6. Chettri, S., Debnath, D. & Devi, P. Leveraging digital tools and technologies to alleviate covid-19 pandemic. *Available at SSRN 3626092* (2020).
7. Ozturk, T. *et al.* Automated detection of covid-19 cases using deep neural networks with x-ray images. *Comput. biology medicine* **121**, 103792 (2020).
8. Zhao, J., Zhang, Y., He, X. & Xie, P. Covid-ct-dataset: a ct scan dataset about covid-19. *arXiv preprint arXiv:2003.13865* (2020).
9. Yang, G.-Z. *et al.* Combating covid-19—the role of robotics in managing public health and infectious diseases (2020).
10. Benjauthrit, B., Benjauthrit, S. B., Benjauthrit, V. L. & Benjauthrit, K. J. Covid-19 symptoms alert machine (csam) scanners (2021). US Patent 10,888,283.
11. Mohammed, M., Syamsudin, H., Al-Zubaidi, S., AKS, R. R. & Yusuf, E. Novel covid-19 detection and diagnosis system using iot based smart helmet. *Int. J. Psychosoc. Rehabil.* **24**, 2296–2303 (2020).
12. Rezaei, M. & Azarmi, M. Deepsocial: Social distancing monitoring and infection risk assessment in covid-19 pandemic. *Appl. Sci.* **10**, 7514 (2020).
13. Di Lauro, F., Kiss, I. Z., Rus, D. & Della Santina, C. Covid-19 and flattening the curve: A feedback control perspective. *IEEE Control. Syst. Lett.* **5**, 1435–1440 (2020).
14. Köhler, J. *et al.* Robust and optimal predictive control of the covid-19 outbreak. *Annu. Rev. Control.* (2020).
15. Goolsbee, A. & Syverson, C. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *J. public economics* **193**, 104311 (2020).
16. Bhuiyan, A. I., Sakib, N., Pakpour, A. H., Griffiths, M. D. & Mamun, M. A. Covid-19-related suicides in bangladesh due to lockdown and economic factors: case study evidence from media reports. *Int. J. Mental Heal. Addict.* 1–6 (2020).
17. Valdez, D., Ten Thij, M., Bathina, K., Rutter, L. A. & Bollen, J. Social media insights into us mental health during the covid-19 pandemic: Longitudinal analysis of twitter data. *J. medical Internet research* **22**, e21418 (2020).
18. Stella, M., Restocchi, V. & De Deyne, S. # lockdown: Network-enhanced emotional profiling in the time of covid-19. *Big Data Cogn. Comput.* **4**, 14 (2020).
19. Yin, H., Yang, S. & Li, J. Detecting topic and sentiment dynamics due to covid-19 pandemic using social media. In *International Conference on Advanced Data Mining and Applications*, 610–623 (Springer, 2020).
20. Twitter intelligence tool. <https://github.com/twintproject/twint>.
21. Hutto, C. & Gilbert, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8 (2014).
22. Dreyer, B. P. *et al.* The death of george floyd: bending the arc of history toward justice for generations of children. *Pediatrics* **146** (2020).
23. Gallup group emotional report. <https://www.gallup.com/analytics/248906/gallup-global-emotions-report-2019.aspx>.
24. Manchester lockdown from october 2020. <https://www.lexisnexis.co.uk/blog/covid-19/coronavirus-covid-19-manchester-to-enter-local-lockdown-from-23-october>.
25. Alberta covid-19 pandemic. https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Alberta.

Author contributions statement

Mahdi Abolfazli Esfahani got the initial idea and started data collection and performing experiments and analysis. Messiah Abolfazli Esfahani helped conduct experiments, and Han Wang helped in reviewing the study and manuscript.

Figures

Figure 1

Seven-Day Rolling Average of Compound Score in USA's Cities

(a) Toronto

(b) Montreal

(c) Vancouver

(d) Calgary

(e) Edmonton

Figure 2

Seven-Day Rolling Average of Compound Score in Canada's Cities

Figure 3

Seven-Day Rolling Average of Compound Score in UK's Cities

Figure 4

Seven-Day Rolling Average of Compound Score in Australia's Cities

(a) USA

(b) Canada

(c) UK

(d) Australia

Figure 5

Yearly Average of Compound Score in Different Countries Cities, and The Entire Country.