2023-04-21

Depression in the Times of COVID-19: A Machine Learning Analysis Based on the Profile of Mood States

PALOMINO, MARCO

10.14746/amup.9788323241775
Adam Mickiewicz University Press

Depression in the Times of COVID-19: A Machine Learning Analysis Based on the Profile of Mood States

Marco A. Palomino*, **Rohan Allen***, **Aditya Padmanabhan Varma**†

*School of Engineering, Computing and Mathematics (SECaM), University of Plymouth
marco.palomino@plymouth.ac.uk; rohan.allen-13@students.plymouth.ac.uk
†Department of Computer Science and Engineering, Chalmers University of Technology
vaditya@chalmers.se

Abstract

As the COVID-19 pandemic continues to unfold, a parallel outbreak of fear and depression is also spreading around, impacting negatively on the well-being of the general public and health care workers alike. In an attempt to develop tools to expedite mental health diagnosis, we have looked into emotion analysis and recognition, as this has become indispensable to understand and mine opinions. We have produced a machine learning classifier capable of identifying one of the moods most commonly associated with COVID-19: depression. To analyse how moods and emotions conveyed about COVID-19 have changed in the public discourse over time, we have gathered two Twitter collections—one from 2020 and one from 2022. Our initial findings indicate that fear and depression remain attached to the COVID-19 discourse over the span of two years. Our insights can aid the design of strategic choices concerning the well-being of people in the UK and worldwide.

Keywords: Sentiment analysis, opinion mining, Twitter, machine learning, COVID-19, profile of mood states

1. Introduction

An outbreak of pneumonia reported in Wuhan, China, in December 2019, quickly spread worldwide, and the thousands of deaths caused by it led the *World Health Organization* (WHO) to declare a pandemic on 11 March 2020 (Ciotti et al., 2020). While adults, particularly men, were at greater risk of developing a serious illness as a consequence of such a *coronavirus* disease (*COVID-19*), studies have shown that the pandemic has affected women, young adults, and the unemployed the hardest in terms of mental health (Imran et al., 2020; Pfefferbaum and North, 2020; Benjamin et al., 2021; O'Connor et al., 2021). Regrettably, these groups also developed frequently both physiological and behavioural symptoms associated with distress (O'Connor et al., 2021; Shader, 2020).

Although a decrease in psychological well-being has been observed in the general public due to COVID-19, and higher levels of psychiatric symptoms have been found among health care workers (Vindegaard and Benros, 2020), those who have sought help have experienced serious delays in being treated (Papautsky et al., 2021). A major goal of our work in the long run is to develop tools to expedite mental health diagnosis. Thus, we would like to delve deeper into the subject of *emotion recognition* (Koolagudi and Rao, 2012) and *sentiment analysis* (Liu, 2012), as they have become indispensable to understand and mine opinions (Sun et al., 2017).

In April 2020, approximately a month after the first nationwide COVID-19 lockdown in the UK was announced, we launched an investigation on the emotions expressed on social media to understand the feelings of the general public. We concentrated on *Twitter* (Murthy, 2018), the micro-blogging platform. To assess the evolution of the emotions expressed about COVID-19 over time, we gathered a second collection of tweets in March 2022.

Then, we processed our two collections of tweets to extract insights into the feelings and emotions expressed by Twitter users. We expect the insights derived from our study to aid in the decision-making of strategic choices concerning the mental health of the population—especially, as a considerable amount of fear, sadness, and depression was conveyed on the tweets that we retrieved.

The reminder of this paper is organised as follows. We will summarise the related work in Section 2. Afterwards, we will describe the corpora that we used for our experiments in Section 3. We will also employ Section 3. to report on the implementation of our machine learning classifier. Section 4. will present our results and, finally, Section 5. will outline our conclusions.

2. Related Work

The availability of large language-based datasets has allowed us to improve the identification and understanding of mental health issues through the study of words (Tausczik and Pennebaker, 2010; Pennebaker et al., 2003). A great deal of research has demonstrated that word use is a reliable indicator of a person's psychological state (Chung and Pennebaker, 2011).

Recognising the emotions expressed by words in pieces of text has earned significance, as an alternative to assess the well-being of people—for example, when attempting to prevent suicide (Desmet and Hoste, 2013). Two of the most notable works on this field, which deserve careful consideration, are *Ekman's* basic emotion model (Ekman, 1992) and *Plutchik's* bipolar emotion model (Plutchik and Kellerman, 2013). Although Ekman's and Plutchik's are well-regarded models, and we would like to look into them in the future, we will not pursue them in our current investigation. Ekman studied facial expressions, but facial recognition is beyond the scope of our project, as we do not have the equipment to pursue it.

Plutchik, on the other hand, considered eight basic, pairwise, contrasting emotions: joy vs. sadness, trust vs. disgust, fear vs. anger, and surprise vs. anticipation (Plutchik, 1980). Even though we plan to widen the range of emotions analysed by our classifiers as our research progresses, the lack of annotated training collections complicates any attempts to implement Plutchik's model.

As depression is a state of mood often associated with COVID-19 (Renaud-Charest et al., 2021; Johns et al., 2022), we wanted to have a classifier capable of detecting depression. Hence, we opted for the *Profile of Mood States* (POMS) (Norcross et al., 1984), which is a psychological test for assessing an individual's mood state (Berger and Motl, 2000). POMS was formulated by McNair *et al.* (McNair et al., 1971) and it contemplates depression. This made it particularly relevant to our work. We will elaborate below on the details of our implementation of POMS.

We were keen on testing approaches that depart from the traditional methods followed by sentiment analysis, which use lexicons and bag-of-words models (Rudkowsky et al., 2018). We are aware of the improvements reported by researchers who have worked with sequences of characters, without pre-processing the text that becomes the input of a *recurrent neural network* (RNN). For instance, Colnerič and Demšar (Colnerič and Demšar, 2018) implemented one of such approaches and used it to classify tweets into emotional categories.

Following Colnerič and Demšar's example, we have implemented our own POMS classifier. However, we only used characters that occurred in the training set 25 times or more, and we removed emoticons and other symbols that were not part of the tweets in our corpus.

3. Materials and Methods

As a testbed for our experiments, we gathered 409,761 tweets about COVID-19 on 22 April 2020, and we will refer to this corpus hereafter as the *2020 Corpus*. We chose 22 April 2020, because it was when the then UK Foreign Secretary, Dominic Raab, delivered a press briefing to address the Government's response to COVID-19, and he highlighted the use of a vaccine for the first time.

The British press started to cover news about a COVID-19 vaccine at the start of April 2020, when the first human trials began in Europe (Walsh, 2020). A significant investment was made on these trials; therefore, we assumed that the briefing on 22 April 2020 would spark off the discussion on Twitter. We thought this would be an ideal moment to capture tweets with a strong sentiment attached to them, either in the form of Government's criticism or concern for the prevailing situation. We expected the briefing to begin at around 16:30; hence, we started the retrieval of tweets a couple of hours prior to the beginning of the briefing, and kept it going for a couple of hours after the end of the briefing. To be precise, we captured our first tweet at 14:24:39, and the last one at 18:56:27.

To ensure that we were capturing information about COVID-19, we looked specifically for tweets comprising the hashtags listed in Table 1. Note that Table 1 also displays the number of tweets retrieved for each hashtag.

Hashtag	Number of tweets
`#covid19`	`238,432`
`#coronavirus`	`116,557`
`#stayhome`	`31,820`
`#covid_19`	`11,068`
`#socialdistancing`	`6,510`
`#covid-19`	`4,636`
`#covid2019`	`2,341`
`#flattenthecurve`	`2,124`
`#coronavirusoutbreak`	`2,058`
`#sarscov2`	`1,861`
`#virus`	`1,211`

Table 1: Hashtags used to retrieve the 2020 Corpus.

Figure 1 shows the number of tweets that we retrieved every 30 minutes. On average, we retrieved 81,952 tweets per hour between 14:24 and 19:24; yet, we retrieved more than 90,000 tweets per hour for the first three hours.

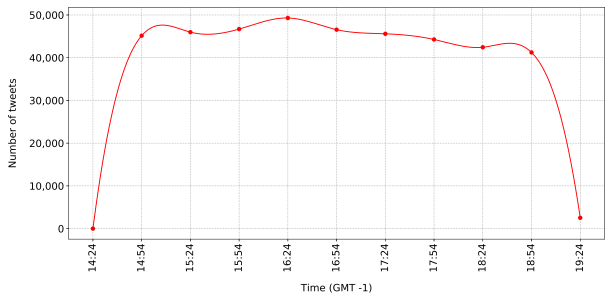

Figure 1: Distribution of tweets in the 2020 Corpus.

To assess how the emotion related to COVID-19 has changed from 2020 to 2022, we retrieved a second corpus, and we will refer to it as the *2022 Corpus*. We started the retrieval of this Corpus on 24 March 2022, because this date marked the second anniversary of the announcement of the first nationwide lockdown in the UK (Johnson, 2022). There was potential for the anniversary to spike the volume of the COVID-19 discourse on Twitter.

We planned to gather as many tweets as we did in 2020. However, COVID-19 seems to have lost popularity as a Twitter topic recently. Hence, we were unable to gather as many tweets as we did before. We collected tweets for five consecutive days, from 24 March 2022 to 29 March 2022—the first tweet was collected on 24 March 2022 at 23:30:05, and the last one on 29 March 2022 at 00:30:07. While the retrieval of the 2020 Corpus lasted 4 hours, 31 minutes, and 48 seconds, and gathered 427,639 tweets, the second one spanned over 4 days, 13 hours, and 2 seconds, and gathered only 265,108 tweets. The size of the 2022 Corpus is only 62% of the size of the 2020 Corpus. However, it provides enough material to assess the evolution of the subject between 2020 and 2022.

To retrieve the 2022 Corpus, we looked for tweets comprising the hashtags listed in Table 2. We considered hashtags that were not part of the discourse in 2020 but have now emerged—for example, `#longcovid`, `#omicron`, `#vaccinated`, and `#covidiots`.

Hashtag	Number of tweets
#covid	182,686
#covid19	125,434
#longcovid	34,565
#covidisnotover	25,010
#omicron	15,608
#coronavirus	9,487
#covid-19	8,443
#pandemic	8,429
#mask	6,627
#sarscov2	2,950
#stayhome	2,118
#virus	2,032
#vaccinated	1,320
#covidiots	1,038
#deltacron	283
#cases	273

Table 2: Hashtags used to gather the 2022 Corpus.

3.1. Profile of Mood States (POMS)

POMS is a test to measure an individual's mood (Curran et al., 1995). POMS is relevant to clinical and social psychology. POMS specifies 65 adjectives that are rated by the individual on a five-point scale. Each adjective contributes to one of seven categories: *anger*, *confusion*, *depression*, *fatigue*, *friendliness*, *tension*, and *vigour*.

Given that POMS can recognise depression, it became ideal for our work, as depression is commonly associated with COVID-19 (Renaud-Charest et al., 2021). Adjectives such as *unworthy*, *miserable* or *gloomy* used to describe a person's feelings contribute to classify her mood within the *depression* category (Mackenzie, B, 2022). Also, we removed *friendliness*, as Norcross *et al.* have found that the adjectives corresponding to it are too weak to ensure a valid classification (Norcross et al., 1984). We complemented the model with other adjectives suggested by the *BrianMac Sports Coach* website (Mackenzie, B, 2022). Table 3 shows the full list of adjectives that we employed to identify each of the mood states under consideration.

For the implementation of our classifier, we experimented with the following options: SVM (Noble, 2006), Naïve Bayes (Berrar, 2018), logistic regression (Kleinbaum et al., 2002), random forests (Biau, 2012)—the number of trees was selected using linear search—and *long short term memory* (LSTM) (Nowak et al., 2017).

To train our classifier, we used Colnerič and Demšar's training set, which is based on a corpus comprising 73 billion tweets annotated using distant supervision (Colnerič and Demšar, 2018). The corpus was collected between August 2008 and May 2015, and it is split into training (60%), validation (20%) and test (20%) sets. Colnerič and Demšar's corpus is considerably larger than other options, such as Mohammad and Kiritchenko's corpus (Mohammad and Kiritchenko, 2015). Unfortunately, the random forest was so slow that we were only able to build forests with a maximum of 100 trees. Training 100 trees using bi-grams took longer than a day on *Google Colab*.

Mood state	Adjectives
anger	angry, peeved, grouchy, spiteful, annoyed, resentful, bitter, ready to fight, deceived, furious, bad tempered, rebellious
confusion	forgetful, unable to concentrate, muddled, confused, bewildered, uncertain about things
depression	sorry for things done, unworthy, guilty, worthless, desperate, hopeless, helpless, lonely, terrified, discouraged, miserable, gloomy, sad, unhappy
fatigue	fatigued, exhausted, bushed, sluggish, worn out, weary, listless
tension	tense, panicky, anxious, shaky, on edge, uneasy, restless, nervous
vigour	active, energetic, full of pep, lively, vigorous, cheerful, carefree, alert

Table 3: Chosen adjectives for POMS.

Instead of pre-processing the tweets, we treated each of them as a sequence of characters, and pass such characters one by one into the RNN. The network's task was to combine the characters into a suitable representation and predict the moods expressed on it. The RNN had to learn which sequences of characters form words, since space was not treated differently from other characters. The benefit of this approach is that it does not require any preprocessing. If we were working with words, we would need a tokenizer first and then we would have to decide which morphological variations of the words are similar enough to consider them equivalent, which is what stemming and lemmatization do (Palomino and Aider, 2022). However, in our character setting approach, all those decisions were left to the neural network to figure out.

4. Results

A common metric to estimate the overall sentiment expressed towards a topic on social media is the *net sentiment rate* (NSR). While the NSR was developed for digital marketing, it has been successfully applied to other fields, for example, Palomino *et al.* (Palomino et al., 2016) have applied it to public health studies. The NSR is defined as the difference between the number of positive conversations—positive tweets—and the number of negative conversations—negative tweets—divided by the total number of conversations—total number of tweets:

$$NSR = \frac{Positive\ tweets - Negative\ tweets}{Total\ number\ of\ tweets}.$$

We selected *SentiStrength* (Thelwall et al., 2010) to determine the sentiment expressed on the tweets constituting our experimental corpus. SentiStrength estimates the strength of positive and negative sentiment in short texts, such as tweets, using methods to exploit the *de-facto* grammars and spelling styles of the informal communication that regularly takes place in social media, blogs and discussion forums (Thelwall et al., 2012).

The NSR values for both the 2020 Corpus and the 2022 Corpus are negative: -0.32% and -0.37%, respectively, reflecting the negative nature of the corpus as a whole. Figure 2 displays the percentages of positive, negative and neutral tweets in the 2020 Corpus and the 2022 Corpus, according to SentiStrength. Because the size of the 2020 Corpus differs from the size of the 2022 Corpus, Figure 2 displays percentages, as opposed to absolute values. As shown in Figure 2, the distribution of the polarities in the 2020 Corpus and the 2022 Corpus is very similar.

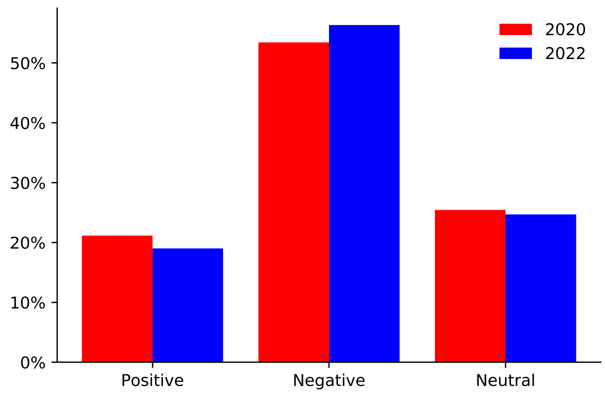

Figure 2: Polarity of the experimental corpus.

Using POMS, we assigned each tweet in the corpus a probability associated with each of the moods under consideration. Figure 3 displays the addition of the probabilities for each mood to occur in each of the tweets. Because the size of the 2020 Corpus differs from the size of the 2022 Corpus, Figure 3 displays percentages, instead of absolute values. Clearly, depression dominates the experimental corpora—the presence of the rest of the emotions appears minimal. Additionally, the percentages of all moods are utterly similar, despite of the year.

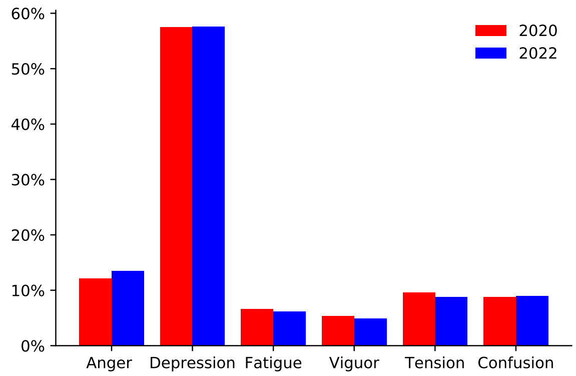

Figure 3: Moods according to POMS.

Research has found that the rates of depression have doubled since the pandemic began (Khubchandani et al., 2021; Seal et al., 2022). An increase in depression commonly accompanies large-scale disasters, whether natural or environmental (Galea et al., 2020).

5. Conclusions

We have applied machine learning to analyse tweets about COVID-19 recorded at two different points in time. Our analysis provides practical insights to aid in the decision-making of strategic choices concerning the well-being of the population. The advantages of using POMS to identify depression have been discussed, and we expect to start the search for other models that are applicable to expedite mental health diagnosis.

6. References

Benjamin, Asaf, Yael Kuperman, Noa Eren, Ron Rotkopf, Maya Amitai, Hagai Rossman, Smadar Shilo, Tomer Meir, Ayya Keshet, Orit Nuttman-Shwartz, et al., 2021. Stress-related Emotional and Behavioural Impact following the First COVID-19 Outbreak Peak. *Molecular psychiatry*, 26(11):6149–6158.

Berger, Bonnie G and Robert W Motl, 2000. Exercise and mood: A selective review and synthesis of research employing the profile of mood states. *Journal of Applied Sport Psychology*, 12(1):69–92.

Berrar, Daniel, 2018. Bayes' theorem and naive bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 403.

Biau, Gérard, 2012. Analysis of a random forests model. *The Journal of Machine Learning Research*, 13(1):1063–1095.

Chung, Cindy K and James W Pennebaker, 2011. Using computerized text analysis to assess threatening communications and behavior. *Threatening Communications and Behavior: Perspectives on the Pursuit of Public Figures*:3–32.

Ciotti, Marco, Massimo Ciccozzi, Alessandro Terrinoni, Wen-Can Jiang, Cheng-Bin Wang, and Sergio Bernardini, 2020. The COVID-19 Pandemic. *Critical Reviews in Clinical Laboratory Sciences*, 57(6):365–388.

Colnerič, Niko and Janez Demšar, 2018. Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. *IEEE Transactions on Affective Computing*, 11(3):433–446.

Curran, Shelly L, Michael A Andrykowski, and Jamie L Studts, 1995. Short Form of the Profile of Mood States (POMS-SF): Psychometric Information. *Psychological assessment*, 7(1):80.

Desmet, Bart and VéRonique Hoste, 2013. Emotion Detection in Suicide Notes. *Expert Systems with Applications*, 40(16):6351–6358.

Ekman, Paul, 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.

Galea, Sandro, Raina M Merchant, and Nicole Lurie, 2020. The Mental Health Consequences of COVID-19 and Physical Distancing: The Need for Prevention and Early Intervention. *JAMA Internal Medicine*, 180(6):817–818.

Imran, Nazish, Muhammad Zeshan, and Zainab Pervaiz, 2020. Mental Health Considerations for Children & Adolescents in COVID-19 Pandemic. *Pakistan Journal of Medical Sciences*, 36(COVID19-S4):S67.

Johns, G, V Samuel, L Freemantle, J Lewis, and L Waddington, 2022. The Global Prevalence of Depression and Anxiety Among Doctors during the Covid-19 Pandemic: Systematic Review and Meta-Analysis. *Journal of Affective Disorders*, 298:431–441.

Johnson, Boris, 2022. Prime Minister's Statement on Coronavirus (COVID-19): 23 March 2020.

Khubchandani, Jagdish, Sushil Sharma, Fern J Webb, Michael J Wiblishauser, and Sharon L Bowman, 2021. Post-Lockdown Depression and Anxiety in the USA during the COVID-19 Pandemic. *Journal of Public Health*, 43(2):246–253.

Kleinbaum, David G, K Dietz, M Gail, Mitchel Klein, and Mitchell Klein, 2002. *Logistic Regression*. Springer.

Koolagudi, Shashidhar G and K Sreenivasa Rao, 2012. Emotion Recognition from Speech: A Review. *International journal of speech technology*, 15(2):99–117.

Liu, Bing, 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Mackenzie, B, 2022. Scoring for POMS. `https://www.brianmac.co.uk/pomscoring.htm`.

McNair, Douglas M, Maurice Lorr, and Leo F Droppleman, 1971. *Manual Profile of Mood States*. Educational & Industrial Testing Service.

Mohammad, Saif M and Svetlana Kiritchenko, 2015. Using Hashtags to Capture Fine Emotion Categories from Tweets. *Computational Intelligence*, 31(2):301–326.

Murthy, Dhiraj, 2018. *Twitter*. Polity Press Cambridge.

Noble, William S, 2006. What is A Support Vector Machine? *Nature Biotechnology*, 24(12):1565–1567.

Norcross, John C, Edward Guadagnoli, and James O Prochaska, 1984. Factor Structure of the Profile of Mood States (POMS): Two Partial Replications. *Journal of Clinical Psychology*, 40(5):1270–1277.

Nowak, Jakub, Ahmet Taspinar, and Rafał Scherer, 2017. LSTM recurrent neural networks for short text and sentiment classification. In *International Conference on Artificial Intelligence and Soft Computing*. Springer.

O'Connor, Rory C, Karen Wetherall, Seonaid Cleare, Heather McClelland, Ambrose J Melson, Claire L Niedzwiedz, Ronan E O'Carroll, Daryl B O'Connor, Steve Platt, Elizabeth Scowcroft, et al., 2021. Mental health and well-being during the covid-19 pandemic: Longitudinal analyses of adults in the uk covid-19 mental health & wellbeing study. *The British Journal of Psychiatry*, 218(6):326–333.

Palomino, Marco, Tim Taylor, Ayse Göker, John Isaacs, and Sara Warber, 2016. The online dissemination of nature–health concepts: Lessons from sentiment analysis of social media relating to "nature-deficit disorder". *International Journal of Environmental Research and Public Health*, 13(1):142.

Palomino, Marco A. and Farida Aider, 2022. Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis. *Applied Sciences*, 12(17).

Papautsky, Elizabeth Lerner, Dylan R Rice, Hana Ghoneima, Anna Laura W McKowen, Nicholas Anderson, Angie R Wootton, and Cindy Veldhuis, 2021. Characterizing Health Care Delays and Interruptions in the United States During the COVID-19 Pandemic: Internet-Based, Cross-sectional Survey Study. *Journal of Medical Internet Research*, 23(5):e25446.

Pennebaker, James W, Matthias R Mehl, and Kate G Niederhoffer, 2003. Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual review of psychology*, 54(1):547–577.

Pfefferbaum, Betty and Carol S North, 2020. Mental Health and the COVID-19 Pandemic. *New England Journal of Medicine*, 383(6):510–512.

Plutchik, Robert, 1980. A general psychoevolutionary theory of emotion. In *Theories of Emotion*. Elsevier.

Plutchik, Robert and Henry Kellerman, 2013. *Theories of Emotion*, volume 1. Academic Press.

Renaud-Charest, Olivier, Leanna MW Lui, Sherry Eskander, Felicia Ceban, Roger Ho, Joshua D Di Vincenzo, Joshua D Rosenblat, Yena Lee, Mehala Subramaniapillai, and Roger S McIntyre, 2021. Onset and Frequency of Depression in Post-Covid-19 Syndrome: A Systematic Review. *Journal of Psychiatric Research*, 144.

Rudkowsky, Elena, Martin Haselmayer, Matthias Wastian, Marcelo Jenny, Štefan Emrich, and Michael Sedlmair, 2018. More than Bags of Words: Sentiment Analysis with Word Embeddings. *Communication Methods and Measures*, 12(2-3):140–157.

Seal, Adam, Andrew Schaffner, Suzanne Phelan, Hannah Brunner-Gaydos, Marilyn Tseng, Sarah Keadle, Julia Alber, Isabelle Kiteck, and Todd Hagobian, 2022. COVID-19 Pandemic and Stay-At-Home Mandates Promote Weight Gain in US Adults. *Obesity*, 30(1):240–248.

Shader, Richard I, 2020. COVID-19 and Depression. *Clinical Therapeutics*, 42(6):962–963.

Sun, Shiliang, Chen Luo, and Junyu Chen, 2017. A Review of Natural Language Processing Techniques for Opinion Mining Systems. *Information Fusion*, 36.

Tausczik, Yla R and James W Pennebaker, 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of language and social psychology*, 29(1):24–54.

Thelwall, Mike, Kevan Buckley, and Georgios Paltoglou, 2012. Sentiment Strength Detection for the Social Web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173.

Thelwall, Mike, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas, 2010. Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Vindegaard, Nina and Michael Eriksen Benros, 2020. COVID-19 Pandemic and Mental Health Consequences: Systematic Review of the Current Evidence. *Brain, Behavior, and Immunity*, 89:531–542.

Walsh, Fergus, 2020. Coronavirus: First Patients Injected in UK Vaccine Trial. BBC News Services.