

Communication

Not peer-reviewed version

Analysis of Public Discourse on Twitter involving COVID-19 and MPox: Findings from Sentiment Analysis and Text Analysis

[Nirmalya Thakur](#) *

Posted Date: 27 March 2023

doi: 10.20944/preprints202303.0453.v1

Keywords: COVID-19; MPox; Twitter; Big Data; Data Mining; Data Analysis; Sentiment Analysis; Data Science; Social Media; Monkeypox



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Communication

Analysis of Public Discourse on Twitter involving COVID-19 and MPox: Findings from Sentiment Analysis and Text Analysis

Nirmalya Thakur

Department of Computer Science, Emory University, Atlanta, Georgia 30322, USA;
nirmalya.thakur@emory.edu

Abstract: Mining and analysis of the Big Data of Twitter conversations have been of significant interest to the scientific community in the fields of healthcare, epidemiology, big data, data science, computer science, and their related areas, as can be seen from several works in the last few years that focused on sentiment analysis and other forms of text analysis of Tweets related to Ebola, E-Coli, Dengue, Human papillomavirus (HPV), Middle East Respiratory Syndrome (MERS), Measles, Zika virus, H1N1, influenza-like illness, swine flu, flu, Cholera, Listeriosis, cancer, Liver Disease, Inflammatory Bowel Disease, kidney disease, lupus, Parkinson's, Diphtheria, and West Nile virus. The recent outbreaks of COVID-19 and MPox have served as "catalysts" for Twitter usage related to seeking and sharing information, views, opinions, and sentiments involving both these viruses. While there have been a few works published in the last few months that focused on performing sentiment analysis of Tweets related to either COVID-19 or MPox, none of the prior works in this field thus far involved analysis of Tweets focusing on both COVID-19 and MPox at the same time. With an aim to address this research gap, a total of 61,862 Tweets that focused on Mpox and COVID-19 simultaneously, posted between May 7, 2022, to March 3, 2023, were studied to perform sentiment analysis and text analysis. The findings of this study are manifold. First, the results of sentiment analysis show that almost half the Tweets (the actual percentage is 46.88%) had a negative sentiment. It was followed by Tweets that had a positive sentiment (31.97%) and Tweets that had a neutral sentiment (21.14%). Second, this paper presents the top 50 hashtags that were used in these Tweets. Third, it presents the top 100 most frequently used words that are featured in these Tweets. The findings of text analysis show that some of the commonly used words involved directly referring to either or both viruses. In addition to this, the presence of words such as "Polio", "Biden", "Ukraine", "HIV", "climate", and "Ebola" in the list of the top 100 most frequent words indicate that topics of conversations on Twitter in the context of COVID-19 and MPox also included a high level of interest related to other viruses, President Biden, and Ukraine. Finally, a comprehensive comparative study that involves a comparison of this work with 49 prior works in this field is presented to uphold the scientific contributions and relevance of the same.

Keywords: COVID-19; MPox; Twitter; Big Data; Data Mining; Data Analysis; Sentiment Analysis; Data Science; Social Media; Monkeypox

1. Introduction

In today's world, social media serves as an "integral vehicle" [1] and as an "online community" [2] for seeking and seeking of information, news, views, opinions, perspectives, ideas, awareness, comments, and experiences on various topics such as pandemics, global affairs, current technologies, recent events, politics, family, relationships, and career opportunities, just to name a few [3]. Out of multiple social media platforms, Twitter is highly popular amongst all age groups. As of December 2022, Twitter's audience accounted for over 368 million monthly active users worldwide [4]. Twitter is the most used social media platform amongst journalists [5] and ranks amongst the most popular social media platforms on a global scale [6].

Twitter has been highly popular amongst healthcare researchers, epidemiologists, medical practitioners, data scientists, and computer science researchers for studying, analyzing, modeling, and interpreting social media communications related to pandemics, epidemics, viruses, and diseases

such as Ebola [7], E-Coli [8], Dengue [9], Human papillomavirus (HPV) [10], Middle East Respiratory Syndrome (MERS) [11], Measles [12], Zika virus [13], H1N1 [14], influenza-like illness [15], swine flu [16], flu [17], Cholera [18], COVID 2022, 2 1028 Listeriosis [19], cancer [20], Liver Disease [21], Inflammatory Bowel Disease [22], kidney disease [23], lupus [24], Parkinson's [25], Diphtheria [26], and West Nile virus [27]. The recent outbreaks of COVID-19 and MPox have served as "catalysts," leading to the usage of Twitter for sharing and exchange of information on diverse topics related to these respective viruses leading to the generation of tremendous amounts of Big Data. No prior work in this field has focused on studying and analyzing Tweets that focused on both these viruses simultaneously to understand and interpret the underlying paradigms of conversations. Therefore, this serves as the main motivation for this work.

In December 2019, there was an outbreak of an unknown respiratory disease in the seafood market in Wuhan, China. This outbreak affected about 66% of the people in the market. A prompt investigation from the healthcare and medical sectors revealed that a novel coronavirus was responsible for this disease, and this virus was named severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2, 2019-nCoV) as it was observed to have a high homology of about 80% with SARS-CoV [28]. The disease humans suffer from after getting infected by this virus is known as COVID-19 [29]. Despite best efforts by the Chinese Government to contain the spread of this virus, it soon spread to other parts of the world while undergoing multiple mutations and several variants such as Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Epsilon (B.1.427 and B.1.429), Eta (B.1.525), Iota (B.1.526), Kappa (B.1.617.1), Zeta (P.2), Mu (B.1.621 and B.1.621.1), and Omicron (B.1.1.529, BA.1, BA.1.1, BA.2, BA.3, BA.4, and BA.5) [34]. At present, there have been a total of 681,518,412 cases and 6,811,869 deaths on account of COVID-19 on a global scale [35]. Respiratory systems are the primary target of the SARS-CoV-2 virus, although infections in other organs of the body have been reported in some cases. The symptoms of COVID-19 usually include fever, dry cough, dyspnea, headache, dizziness, exhaustion, vomiting, and diarrhea [36]. However, studies have shown that symptoms can vary from person to person based on diversity characteristics such as age group, preexisting conditions, disabilities, etc. [37,38].

MPox (Monkeypox) is a re-emerging zoonotic disease. It is caused by the Mpox (monkeypox) virus, which belongs to the Poxviridae family, Chordopoxvirinae subfamily, and Orthopoxvirus genus [39]. This virus was originally identified in monkeys in 1958 [40], and the first case of this virus in humans was recorded in 1970. The Mpox virus is closely related to the variola virus and causes a smallpox-like disease in humans. The common symptoms of Mpox include fever, headache, and myalgia. A distinguishing feature of Mpox is the presence of swelling at the maxillary, cervical, or inguinal lymph nodes [41,42]. The Mpox virus was endemic in the Democratic Republic of the Congo (DRC) and a few African countries for a long time, and a few cases outside these geographic regions were recorded only twice—first in 2003 [43] and then in 2017–2018 [44,45]. However, since May 2022, the world is also experiencing an outbreak of the Mpox virus. At present, there have been a total of 86,231 cases of Mpox, with 84,858 of these cases being recorded in regions that have not historically reported Mpox [46].

In the context of recent works related to Twitter data mining and analysis, a number of works have focused on sentiment analysis of Tweets. Sentiment Analysis is the computational analysis of people's attitudes, views, and sentiments regarding an entity that may represent an individual, concept, topic, event, or scenario. Sentiment Analysis can be considered a classification process. The three primary classification levels in Sentiment Analysis are document level, sentence level, and aspect level. The goal of document-level Sentiment Analysis is to categorize an opinion document as expressing a positive or negative sentiment. The entire document is viewed as a single fundamental informational unit in this process. Sentence-level Sentiment Analysis seeks to categorize the sentiment that each sentence expresses. In order to categorize the sentiment in relation to particular features of entities, aspect-level Sentiment Analysis is used. While the prior works in this field focused on performing sentiment analysis, the works were focused on studying either Tweets about COVID-19 or Tweets about MPox and did not include Tweets that focused on both these viruses simultaneously. The outbreak of MPox during the ongoing outbreak of COVID-19 has resulted in

several Tweets involving the views, opinions, concerns, and perspectives of the public regarding both these viruses. Examples of a few such Tweets (obtained by using the Advanced Search feature of Twitter) are shown in Table 1.

Table 1. A random collection of 10 Tweets that focused on COVID-19 and MPox simultaneously.

Tweets related to COVID-19 and MPox	
Tweet #1	They cant figure out how Monkey Pox got here without traveling, and why people are susceptible to it after Covid? Try looking at your immune system after taking the vaccines. Every disease that ever was, is now something for you to fear. Your immune system has been compromised.
Tweet #2	Thanks all you biden fans letting in all these illegal immigrants that have been coming every day since Biden took office. Now we have to worry even more about a new virus coming into this country Monkey Pox forget Covid welcome MONKEY POX
Tweet #3	So, I've got my rainbow sticker, Thank you NHS on my window, 'I've had my covid vaccine' on my fb page, Ukraine flag in the garden. It still isn't enough to show how nice I am! Just need a monkey pox sticker. Deffo going to heaven. Stay safe everyone
Tweet #4	MONKEY POX, I am so not ready for you to show up anywhere. Can you imagine the dilemma of future docs, Now with long COVID , long monkey, monkey heart, monkey lungs, monkey brain might emerge. a monkey mask might help. If lived long enough, might have COVID docs, monkey docs etc
Tweet #5	Are you kidding me, now Monkey Pox?! I've spent 3 years caring for my ill wife, fighting against Covid, and trying to survive...now this?! Some days...
Tweet #6	I sure hope the Government doesn't plan to try to force everyone to get monkey pox vaccines. I'd hate to see where that goes so shortly after covid.
Tweet #7	Another lockdown is incoming. They are trying to make monkey pox look like a pandemic. Their media tools are ready, their vaccines were ready before the pox was introduced. These were the same people that played the COVID19 play. They just changed the name of the movie. Failure!
Tweet #8	Monkey Pox new Covid. Election is coming. Coincidence? No
Tweet #9	First it was maga. Then there came covid. Now, it's Monkey Pox. When will these horrors end!?
Tweet #10	No longer scared of disease be it Covid or Monkey pox; I'm scared of loosing more years of my life ...

¹ These Tweets are presented here in "as is" form after obtaining the same from the Advanced Search feature of Twitter. These Tweets do not reflect the views or opinions or beliefs or political stance of the author of this paper.

As can be seen from these Tweets, these two ongoing virus outbreaks prompted sharing and exchange of views, information, concerns, and perspectives on a wide range of topics that reflect various sentiments regarding these viruses and how these viruses may affect aspects of people's lives. No prior work in this field thus far has focused on studying and analyzing Tweets that involved conversations about both COVID-19 and MPox. This work aims to address this research gap in this field. The work of this paper involved performing sentiment analysis and text analysis on 61,862 Tweets that focused on Mpox and COVID-19 simultaneously, posted between May 7, 2022, to March 3, 2023. The findings are summarized as follows:

- The results of sentiment analysis using the VADER (for Valence Aware Dictionary for sEntiment Reasoning) approach shows that nearly half the Tweets (actual percentage being 46.88%) had a negative sentiment. It was followed by Tweets that had a positive sentiment (31.97%) and Tweets that had a neutral sentiment (21.14%).
- Using concepts of text analysis, the top 50 hashtags associated with these Tweets were obtained. These hashtags are presented in the paper.
- Using concepts of text analysis, the top 100 most frequently used words featured in these Tweets were obtained. The findings show that some of the commonly used words involved Twitter users directly referring to either or both viruses. In addition to this, the presence of words such as "Polio", "Biden", "Ukraine", "HIV", "climate", and "Ebola" in the list of the top 100 most frequent words indicate that topics of conversations on Twitter in the context of COVID-19 and MPox also included a high level of interest related to other viruses, President Biden, and Ukraine.

In addition to the above, a comprehensive comparative study that compares the contributions of this paper with 49 prior works in this field to uphold its relevance and novelty is also presented in this paper. This paper is organized as follows. In Section 2, an overview of recent works related to sentiment analysis in the context of COVID-19 and MPox is presented. Section 3 outlines the detailed

methodology and the specific steps that were followed for this work. In Section 4, the results are presented and discussed. Section 5 presents the conclusion and scope for future work along these lines. It is followed by references.

2. Literature Review

This section is divided into two parts. Section 2.1 presents a review of recent works in this field that focused on sentiment analysis of Tweets about COVID-19. Section 2.2 presents a review of the recent works in this field that focused on sentiment analysis of Tweets about MPox.

2.1. Recent Works that Focused on Sentiment analysis of Tweets about COVID-19

The study by Vijay et al. [44] examined the impact of COVID-19-related tweets from November 2019 to May 2020 in India. Three categories were created for all tweets (Positive, Negative, and Neutral). To assess how people would react to the COVID-19 lockdown in June 2020, the authors also created many datasets, which were organized by state and month and pooled across all states. The findings showed that most individuals started off tweeting negatively, but as time went on, more and more people began to post positively and neutrally. The work by Mansoor et al. [45] examined the global sentiment analysis of tweets on the coronavirus and the evolution of global sentiment over time. The authors also studied tweets focusing on Work From Home (WFH) and Online Learning to gauge the effect of COVID-19 on daily areas of life. They used different machine learning models, such as Long Short Term Memory (LSTM) and Artificial Neural Networks (ANN), for performing the sentiment analysis and reported the accuracy of these models.

Pokharel [46] used Google Colab for text mining and for performing sentiment analysis on Tweets focusing on COVID-19. The study involved collecting Tweets posted between May 21, 2020, and May 31, 2020, by Twitter users who shared Nepal as their location. According to the study's findings, while the majority of people had a positive attitude towards COVID-19, there were also situations where fear, grief, and disdain were expressed in the Tweets.

In the study done by Chakraborty et al. [47], two categories of tweets related to COVID-19 were studied. In one instance, the top 23,000 retweeted tweets over the period of January 1, 2020, to March 23, 2020, were studied. According to the findings presented by the authors, the majority of the tweets express neutral or negative emotions. The paper also reports the findings from the analysis of a dataset encompassing 226,668 tweets from the period between December 2019 and May 2020. The findings show that there were a disproportionately high number of neutral and positive tweets posted by internet users. The study also showed that despite the majority of COVID-19-related tweets being positive, internet users were preoccupied with retweeting the negative ones.

The objective of the work done by Shofiya et al. [48] was to comprehend and examine popular perceptions of social distancing in the context of COVID-19, as expressed on Twitter. The study focused on analyzing Tweets emerging from Canada and containing social distancing-related keywords. The authors used the SentiStrength tool to determine the sentiment polarity of Tweets. Basiri et al. [49] proposed a methodology based on the fusion of four deep learning and one classical supervised machine learning model for sentiment analysis of COVID-19 tweets. They applied their methodology to Tweets originating from eight countries. Cheeti et al. [50] used a Naïve Bayes Classifier to perform sentiment analysis of Tweets focusing on COVID-19 with a specific focus on Tweets related to education and learning. In their study, Ridhwan et al. [51] performed sentiment analysis of Tweets about COVID-19 posted between February 1, 2020, to August 31, 2020, with a specific focus on Tweets that originated from Singapore. The findings showed that a majority of the Tweets had a positive emotion. Tripathi [52] and Situala et al. [53] used multiple machine-learning approaches to perform sentiment analysis of COVID-19-focused Tweets that were posted by people who stated their location as Nepal on Twitter. The purpose of the work done by Gupta et al. [54] was to examine perceptions of Indians, as expressed on Twitter, towards the Indian Government's countrywide lockdown, which was implemented to slow the spread of COVID-19. In this context, the authors used the LinearSVC classifier to perform sentiment analysis, and their classifier reported a performance accuracy of 84.4%. The work done by Alanezi et al. [55] focused on performing

sentiment analysis of Tweets originating from multiple countries. The results of the study showed that most Tweets originating from the USA, Australia, Nigeria, Canada, and England, had a neutral sentiment. A similar study that focused on performing sentiment analysis of Tweets originating from multiple countries was performed by Dubey [56]. In addition to the above, several studies focused on performing sentiment analysis of Tweets about COVID-19 originating from different countries such as the United Kingdom [57-62], United States [59, 60, 63-67], Canada [68-72], India [73-77], Australia [78-80], and Brazil [81-83].

2.2. Recent Works that Focused on Sentiment Analysis of Tweets about MPox

Iparraguirre-Villanueva et al. [84] aimed to examine people's emotions, including positive, negative, and neutral sentiments, towards the Mpox outbreak by analyzing Tweets containing the hashtag #Monkeypox. The findings of the study indicated that 45.42% of individuals did not express any discernible positive or negative opinions, whereas 19.45% conveyed negative and apprehensive sentiments related to the outbreak. The objective of the study by Mohbey et al. [85] was to infer the range of reactions of the general public in response to the MPox outbreak. The methodology was based on using CNN and LSTM to study relevant Tweets to infer these specific characteristics. Farahat et al. [86] conducted a study involving sentiment analysis and topic modeling of Tweets associated with MPox. The Tweets that were analyzed in this study were posted on Twitter between May 22, 2022, and August 5, 2022. The authors utilized keyword search to mine Tweets containing the keywords "monkeypox", "Monkeypox cases", and "Monkeypox virus". The findings of the sentiment analysis indicated that 48% of the Tweets were neutral, 37% were positive, and 15% were negative. The authors used LDA to extract 12 topics that were present in these Tweets. Sv et al. [87] focused on understanding the attitude of the common public towards Mpox, as expressed on Twitter. They performed sentiment analysis on Tweets containing the keyword monkeypox, which were posted between June 1, 2022, to June 25, 2022. The results of sentiment analysis showed that the percentage of positive Tweets was more as compared to the percentage of negative Tweets. The results of topic detection revealed multiple subject matters associated with both positive and negative Tweets. The work by Bengesi et al. [88] was performed primarily in two steps. The first step of their work involved mining relevant Tweets related to MPox. Thereafter in the next step, they developed and used multiple categorization models to perform sentiment analysis of these Tweets. Dsouza et al.'s work [89] focused on performing sentiment analysis of specific Tweets related to Mpox to detect any stigmatization of the LGBTQ+ community on Twitter. They retrieved Tweets posted between May 1, 2022, and September 7, 2022, comprising the hashtags "#monkeypox", "#MPVS", "stigma," and "#LGBTQ+." The study involved the analysis of a total of 70,832 tweets.

Zuhanda et al. [90] performed sentiment analysis on 5000 Tweets about MPox posted on August 5, 2022. The study showed that the terms "health," "emergency," "public," "covid," and "declares" were often used by Twitter users in the context of tweeting about MPox. The NRC lexicon comparison categorization revealed that fear was the most often expressed emotion, with a presentation rate of 19.73%. This was followed by sorrow at 14.77%, trust at 13.90%, anger at 9.99%, shock at 9.14%, disgust at 8.12%, and happiness at 7.90%. In the work done by Cooper et al. [91], Tweets about MPox posted between May 1, 2022, to July 23, 2022, were studied. The results showed that LGBTQ+ advocates or allies posted a total of 48,330 tweets, and the average sentiment score for all the tweets was -0.413 on a scale of -4 to +4. Ng et al. [92] collected Tweets that contained the phrases "monkeypox," "monkey pox," or "monkey pox" posted on Twitter between May 6, 2022, to July 23, 2022. They used concepts of topic modeling and sentiment analysis to infer characteristics of communication expressed in these Tweets. The authors identified five issues which they divided into three main themes. These included worries about safety, the stigmatization of minority populations, and a general loss of confidence in governmental institutions. The public sentiments highlighted increasing and existing partisanship, personal health concerns related to the changing situation, as well as worries about how the media portrayed minority and LGBTQ communities.

As can be seen from these works that focused on sentiment analysis of Tweets related to MPox and COVID-19, none of these works focused on mining and analyzing Tweets that focused on

COVID-19 and MPox at the same time to infer the underlying patterns of sentiments. The work presented in this paper aims to address this research gap. The methodology that was followed is presented and discussed in Section 3, and the results are presented in Section 4.

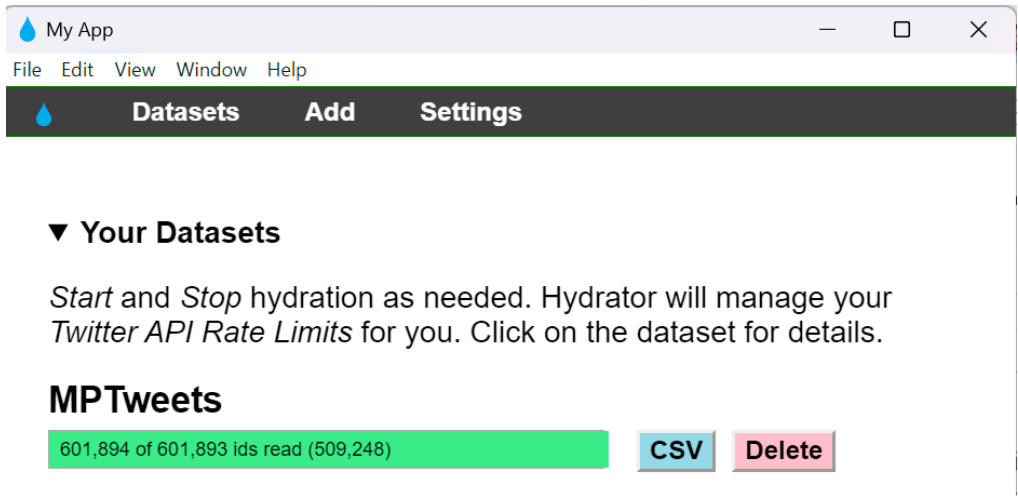
3. Materials and Method

This section outlines the methodology that was followed for the development and implementation of the proposed framework for performing sentiment analysis and text analysis of Tweets that focused on COVID-19 and MPox simultaneously.

First of all, a relevant Twitter dataset had to be selected. The dataset that was selected for this study is MonkeyPox2022Tweets [107]. This dataset presents more than 600,000 Tweet IDs of Tweets about the 2022 outbreak of MPox. These Tweets were posted between May 7, 2022, to March 3, 2023. The dataset comprises Tweets in 34 languages, with English being the most common language in which the Tweets are available. The Tweets in the dataset include 5470 distinct hashtags related to MPox, out of which #monkeypox is the most frequent hashtag. As this dataset comprises only Tweet IDs, the Hydrator App [108] was used to hydrate this dataset. The process of hydration refers to the process of obtaining the Tweets and related information corresponding to each of the Tweet IDs. The Hydrator App works by complying with the policies of accessing the Twitter API as well as the specific rate limits in terms of accessing the Twitter API. The following steps were followed for hydrating the Tweet IDs present in this dataset:

1. The desktop version of Hydrator was downloaded and installed on a Computer with Microsoft Windows 10 Pro operating system (Version 10.0.19043 Build 19043) comprising of Intel(R) Core(TM) i7-7600U CPU @ 2.80GHz, 2904 Mhz, 2 Core(s), and 4 Logical Processor(s)
2. The Hydrator app was then connected to the Twitter API by clicking on the "Link Twitter Account" button on the app's interface.
3. This next step involved uploading a dataset file to the Hydrator app for hydration. As the Hydrator App allows only one file to be uploaded at a time, so all the dataset files were merged to create one .txt file, which was uploaded to the app.
4. Then specific information about the uploaded dataset file (such as Title, Creator, Publisher, and URL) was entered in the Hydrator app, and then the "Add Dataset" button was clicked to complete the process of dataset upload
5. Thereafter, in the "Datasets" tab of the Hydrator App, the "Start" button was clicked to initiate the process of hydration.

Figure 1 is a screenshot from the Hydrator App obtained after the completion of this hydration task.



The output of the Hydrator app provided 509,248 Tweets about MPox. Upon obtaining these Tweets, it was crucial to perform text filtering to obtain those Tweets that contained keywords related to COVID-19. The specific keywords that were selected for the text filtering were "COVID", "COVID19", "coronavirus", "coronavirus pandemic", "COVID-19", "corona", "corona outbreak",

"omicron variant", "SARS-CoV-2", "corona virus", and "Omicron". These keywords were selected based on the findings of [93]. The text filtering task produced a set of 61,862 Tweets, i.e., each of these Tweets focused on MPox and COVID-19 at the same time. This set of 61,862 Tweets was selected for performing sentiment analysis and text analysis.

There are various approaches for performing sentiment analysis, such as manual classification, Linguistic Inquiry and Word Count (LIWC), Affective Norms for English Words (ANEW), the General Inquirer (GI), SentiWordNet, and machine learning-oriented techniques relying on Naive Bayes, Maximum Entropy, and Support Vector Machine (SVM) algorithms. However, the specific approach that was used for this study was VADER (Valence Aware Dictionary for sEntiment Reasoning). VADER was used as it has been reported to outperform manual classification as well as it addresses the limitations in similar approaches for sentiment analysis as outlined below [94]:

- a) VADER distinguishes itself from LIWC as it is more sensitive to sentiment expressions in social media contexts.
- b) The General Inquirer suffers from a lack of coverage of sentiment-relevant lexical features common to social text.
- c) The ANEW lexicon is also insensitive to common sentiment-relevant lexical features in social text.
- d) The SentiWordNet lexicon is very noisy; a large majority of synsets have no positive or negative polarity.
- e) The Naïve Bayes classifier involves the naive assumption that feature probabilities are independent of one another.
- f) The Maximum Entropy approach makes no conditional independence assumption between features and thereby accounts for information entropy (feature weightings).
- g) In general, machine learning classifiers require (often extensive) training data, which are, as with validated sentiment lexicons, sometimes troublesome to acquire.
- h) In general, machine learning classifiers also depend on the training set to represent as many features as possible.

VADER uses sparse rule-based modeling to build a computational sentiment analysis engine that performs well on social media style text while easily generalizing to multiple domains, needs no training data but is built from a generalizable, valence-based, human-curated gold standard sentiment lexicon, is quick enough to utilize online with streaming data, does not suffer significantly from a speed-performance tradeoff, has a time complexity of $O(N)$, and is freely available without any subscription or purchase costs. In addition to detecting the polarity (positive, negative, and neutral), VADER is also able to detect the intensity of the sentiment expressed in the texts. For developing the system architecture for sentiment analysis and text analysis, RapidMiner was used. RapidMiner, formerly known as Yet Another Learning Environment (YALE) [95], is a data science platform that enables the development, implementation, and utilization of several algorithms and models related to Machine Learning, Data Science, Artificial Intelligence, and Big Data. RapidMiner is utilized for both academic research and the creation of business-related applications and solutions. RapidMiner is available as an integrated development environment that consists of—(1) RapidMiner Studio, (2) RapidMiner Auto Model, (3) RapidMiner Turbo Prep, (4) RapidMiner Go, (5) RapidMiner Server, and (6) RapidMiner Radoop. For all the work related to the methodologies proposed in this paper, RapidMiner Studio was used. For the remainder of this paper, wherever the phrase "RapidMiner" has been used, it refers to "RapidMiner Studio" and not any of the other development environments associated with this software tool. RapidMiner is created as an open-core model with a powerful Graphical User Interface (GUI) that enables developers to create numerous applications and workflows and develop and implement algorithms. In the RapidMiner development environment, specific operations or functions are referred to as "operators," and a collection of "operators" (connected linearly or hierarchically or a combination of both) to achieve a desired task or goal is referred to as a "process". For the creation of a particular "process," RapidMiner offers a variety of built-in "operators" that may be utilized straight away with or without any changes. A particular class of "operators" can also be utilized to change the distinguishing qualities of other "operators". Moreover, the development environment also allows developers to construct their own

"operators," which can then be shared and made accessible to all other RapidMiner users via the RapidMiner Marketplace.

The VADER approach for performing sentiment analysis is available as an "operator" in RapidMiner, which can be directly used in a "process." This "operator" calculates and then outputs the sum of all sentiment word scores in a given text(s) by following the VADER approach. If the advanced output option of this "operator" is selected, then it also outputs a nominal attribute with all words taking part in the scoring, the sum of positive components, the sum of negative components, and the number of used and unused tokens. The "process" that was developed in RapidMiner involving the use of this "operator" and other "operators" connected to it is shown in Figure 2.

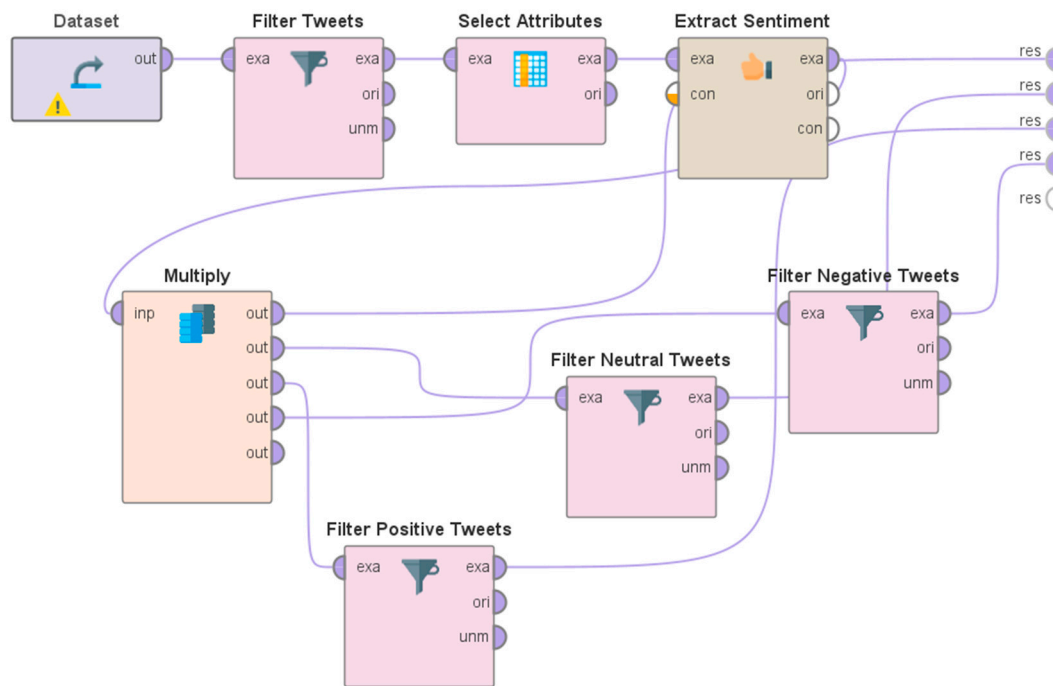


Figure 2. The RapidMiner process developed for performing sentiment analysis.

The description of all the "operators" used in this "process" is presented next. The "Dataset" "operator" was used to import the original dataset of 509,248 Tweets about MPox (obtained from the output of the Hydrator app). The "Filter Tweets" "operator" was used to perform text filtering on the text of the Tweets. Specifically, the Tweets that contained the keywords "COVID", "COVID19", "coronavirus", "coronavirus pandemic", "COVID-19", "corona", "corona outbreak", "omicron variant", "SARS-CoV-2", "corona virus", and "Omicron" were filtered. Thereafter, the "Select Attributes" "operator" was used to select only that specific attribute from the dataset that would be used for sentiment analysis. The specific attribute in this context was the text of the Tweets. The output of this "operator" was provided as an input to the "Extract Sentiment" "operator" which performed sentiment analysis according to the VADER approach. The output of this "operator" comprised a score associated with each Tweet classifying it into a positive, neutral, or negative Tweet. To compute the number of positive, neutral, or negative Tweets, additional data filters were used. However, this required creating multiple copies of the output. To achieve the same, the "Multiply" "operator" was used. Specifically, three copies of the output from the VADER "operator" were created by using this operator. Each of these copies of the output was passed to data filters which were set up to filter out the positive, neutral, and negative Tweets based on specific rules based on the working of the VADER approach – a Tweet with a score greater than 0 was filtered as a positive Tweet, a Tweet with a score equal to 0 was filtered as a neutral Tweet and a Tweet with a score less than 0 was filtered as a negative Tweet. Thereafter, an analysis of the number of Tweets from these respective data filters

was performed to infer the percentages of positive, neutral, and negative Tweets. These results are discussed in Section 4.

In addition to performing sentiment analysis, this study also involved the detection of some of the commonly used hashtags and words in the 61,862 Tweets that were considered for this study. The RapidMiner "process" that was developed to implement the same is shown in Figure 3.

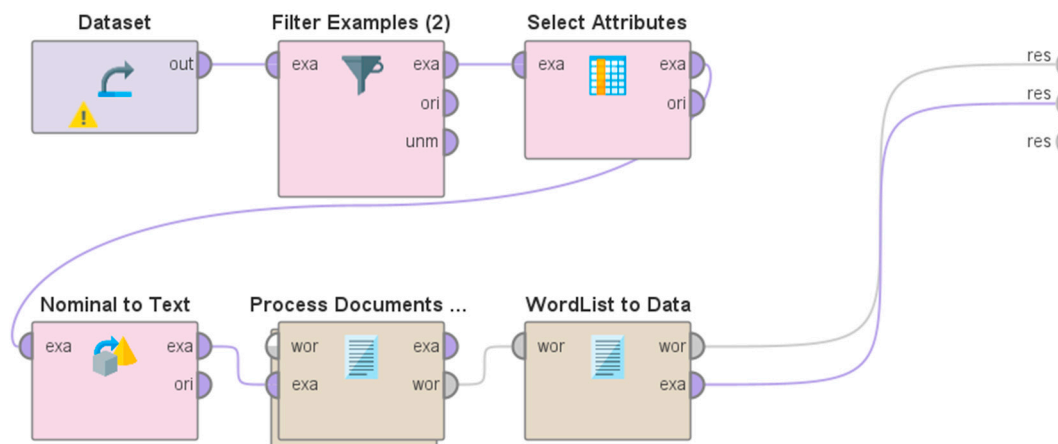


Figure 3. RapidMiner process for performing text analysis of the Tweets.

The description of all the "operators" used in this "process" is presented next. The "Dataset" "operator" was used to import the original dataset of 509,248 Tweets about MPox (obtained from the output of the Hydrator app). The "Filter Tweets" "operator" was used to perform text filtering on the text of the Tweets. Specifically, the Tweets that contained the keywords "COVID", "COVID19", "coronavirus", "coronavirus pandemic", "COVID-19", "corona", "corona outbreak", "omicron variant", "SARS-CoV-2", "corona virus", and "Omicron" were filtered. Thereafter, the "Select Attributes" "operator" was used to select only that specific attribute from the dataset that would be used for sentiment analysis. The specific attribute in this context was the text of the Tweets. The output of this "operator" was provided as an input to the "Nominal to Text" operator. Thereafter the "sub-process" "Process Documents" was used. This "sub-process" comprised specific operators to perform tokenization and elimination of stop words. The output of this "operator" was provided as an input to the "WordList to Data" operator to display the results for detection and analysis of the commonly used hashtags and words in these Tweets. The results of this "process" are discussed in Section 4. It is worth mentioning here that the VADER "operator" performs tokenization and elimination of stop words automatically, so the "sub-process" "Process Documents" was not used in the RapidMiner "process" for performing sentiment analysis shown in Figure 3.

4. Results and Discussions

The section presents the results and novel findings of this study. This section is divided into three parts. Section 4.1 presents the results of sentiment analysis of 61,862 Tweets that focused on Mpox and COVID-19 at the same time. In Section 4.2, the results of text analysis of the Tweets are presented. Specifically, this section reports some of the commonly used hashtags and words that were present in these Tweets. Section 4.3 presents a comprehensive comparative study with all the prior works in this field (reviewed in Section 2) to further uphold the scientific contributions of this paper.

4.1. Results of Sentiment Analysis

The sentiment analysis on this set of 61,862 Tweets that focused on Mpox and COVID-19 at the same time was performed using the VADER approach. The output of the VADER "operator" presented multiple new attributes, and each attribute provided specific information related to the sentiment associated with the Tweets that were analyzed. Figure 4 shows the output that was

produced by the RapidMiner "process" shown in Figure 2. To avoid presenting an image with 61,862 rows, Figure 4 shows a random selection of 19 rows from the output table.

Row No.	Score	Scoring String	Negativity	Positivity	Uncovered Tokens	Total Tokens	text
19	0.128	top (0.21) comedy (0.38) lies (-0.46)	0.462	0.590	41	44	You know how I know this cou...
20	0.179	growing (0.18)	0	0.179	27	28	Monkey pox is the COVID versi...
21	-0.564	infected (-0.56)	0.564	0	37	38	Bill Gates intentionally infected...
22	-1.410	die (-0.74) dumb (-0.59) super (0.74) pissed (-...	2.154	0.744	23	27	If I survive leukemia twice and ...
23	0.026	sure (0.33) worried (-0.31)	0.308	0.333	62	64	@PaganTri @iva_question @...
24	-0.282	want (0.08) super (0.74) strain (-0.05) definite...	2.282	2	57	67	If they want me to stay inside &...
25	0.359	like (0.38) increased (0.28) destruction (-0.69)...	0.692	1.051	47	51	@debrakidd T cell Depletion d...
26	-2	wrong (-0.54) gun (-0.36) violence (-0.79) no ...	2	0	17	21	@shawnouten3 And You ainâ...
27	0.308	ok (0.31)	0	0.308	25	26	@BoushahriAliaa OK, I quit
28	-0.026	yes (0.44) injury (-0.46)	0.462	0.436	21	23	Yes it does. Monkey Pox is the...
29	-0.872	stop (-0.31) fear (-0.56)	0.872	0	23	25	@GBNEWS Covid was a lie a...
30	-0.846	no (-0.31) no (-0.31) forget (-0.23)	0.846	0	43	46	@amlivemon No equivalency, ...
31	0		0	0	17	17	@Sally718068071 Covid inve...
32	-2.615	attack (-0.54) recession (-0.46) recession (-0.4...	3.128	0.513	43	51	â"2001 - 9/11 attack
33	-0.538	frustration (-0.54)	0.538	0	16	17	@DataDrivenMD Covid is tech...
34	0.462	lol (0.46)	0	0.462	20	21	Lol. Itâ"s coming. Part two of ...
35	0.821	please (0.33) thanks (0.49)	0	0.821	20	22	Covid numbers are up again, ...
36	0.103	fright (-0.41) welcome (0.51)	0.410	0.513	22	24	Covid was really unseen scare...
37	0		0	0	18	18	@MidnightRyderX Monkey pox...

Figure 4. A random selection of 19 rows from the output table generated by RapidMiner.

In this Figure, the columns marked in yellow were introduced by the VADER "operator" and were not originally present in the dataset. For each Tweet, the VADER approach performed tokenization at first. This is represented in the attributes "Total Tokens" and "Uncovered Tokens" in Figure 4. Thereafter, it captured those tokens from the Tweets which expressed either a positive or negative sentiment and then assigned a sentiment score to these respective tokens. This score was assigned on a scale of -4 to +4, where -4 meant highly negative, and +4 meant highly positive. These sets of tokens and their respective sentiment scores comprised the value of the "scoring string" (as shown in Figure 4) for each Tweet. Thereafter, for each Tweet, the VADER approach grouped all those tokens that had a positive sentiment and computed the sum of the sentiment scores for those tokens. This comprised the "Positivity" value of that Tweet. In a similar manner, the VADER approach grouped all those tokens that had a negative sentiment and computed the sum of the sentiment scores for those tokens. This comprised the "Negativity" value of that Tweet. The difference between the "Positivity" value and the "Negativity" value was thereafter computed by the VADER approach to display the overall score of the Tweet. If this score was negative, the Tweet was considered to have an overall negative sentiment. If this score was positive, the Tweet was considered to have an overall positive sentiment. Finally, if this score was zero, the Tweet was considered to have a neutral sentiment. Based on this analysis, the number of Tweets with a positive sentiment was observed to be 29000, the number of Tweets with a negative sentiment was observed to be 19780, and the number of Tweets with a neutral sentiment was observed to be 13082. This is illustrated in Figure 5. According to these findings, it can be concluded that almost half the Tweets (actual percentage being 46.88%) involving COVID-19 and MPox had a negative sentiment. It was followed by Tweets that had a positive sentiment (31.97%) and Tweets that had a neutral sentiment (21.14%).

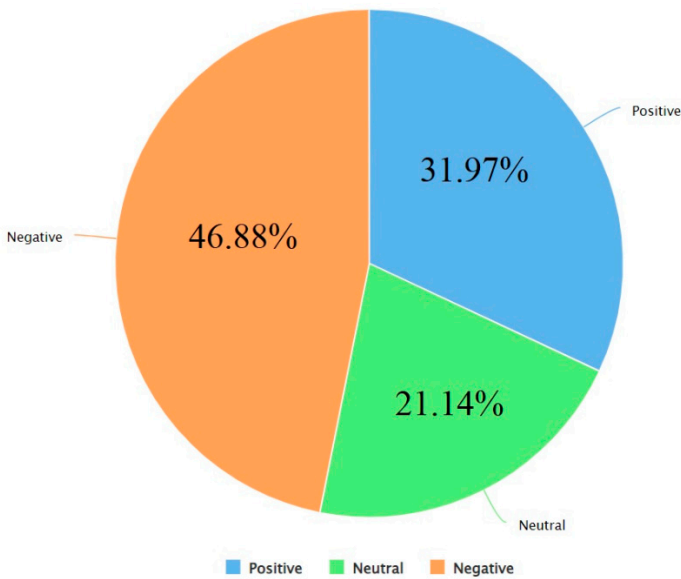


Figure 5. Representation of the percentage of positive, negative, and neutral Tweets present in this dataset.

4.2. Results of Text Analysis

The results of text analysis of the Tweets are presented in this section. The specific set of analyses that were performed included hashtag analysis and tokenization for detecting the most frequent words featured in these Tweets. Table 2 shows the list of the top 50 hashtags and their frequencies. Here, frequency refers to the number of times each of these hashtags were present in the total number of Tweets.

Table 2. The list of top 50 hashtags and their frequencies in the given Tweets.

Hashtag.	Frequency
monkeypox	350
COVID19	97
Monkeypox	88
monkeypox COVID19	77
COVID19 monkeypox	64
COVID	31
MonkeyPox	29
SchlongCovid	27
monkeypox COVID	24
CovidIsNotOver	21
covid monkeypox	21
COVID monkeypox	19
MonkeypoxVirus	18
monkeypox Covid_19	17
covid19	16
COVIDisAirborne	15
moneypox	15
monkeypox covid	15
schlongcovid	15
auspol	14
COVID19 Monkeypox	13
CovidIsNotOver monkeypox	12
Monkeypox COVID19	12

Covid monkeypox	11
Covid19	11
LongCovid	11
covid	11
covid19 monkeypox	11
Covid_19	9
Covid_19 monkeypox	9
LoveIslandUSA	9
MoneyPox	9
monkeypox covid19	9
Monkeypox COVID	8
PrimeMorning	8
monkeypox monkeypox	8
COVID19aus COVID19vic Wearamask	7
Covid	7
Covid19 monkeypox	7
LoveIsland	7
MedTwitter	7
MonkeyPox COVID19	7
monkeypox CovidIsNotOver	7
rogerbezanis LetsGoBrandon	7
Covid Monkeypox	6
FJB	6
Russia Ukraine StopTheTreaty monkeypox BuildBackBetter NWO COVID19	
GreatReset arrestBillGates FireFauci	6
SmartNews	6
cdnpoli	6
covid Monkeypox	6

The results of tokenization are presented next. In view of the large number of tokens obtained from the set of Tweets, this analysis was performed by including the top 100 tokens in terms of their respective frequencies. Table 3 shows these tokens, and a visual representation of the same in the form of a word cloud is shown in Figure 6. As can be seen from this table and Figure 6, several words directly related to these respective viruses are in the list of the top 100 used words (tokens). This was expected as this study focuses on Tweets about COVID-19 and MPox. At the same time the fact that this analysis shows several words which are not directly related to any of these viruses, such as "Polio", "Biden", "Ukraine", "HIV", "climate" and "Ebola" in the list of top 100 most frequent words that featured in these Tweets, underlines the fact that topics of conversations on Twitter in the context of COVID-19 and MPox also included a high level of interest related to other viruses, President Biden, and Ukraine.

Table 3. The list of the top 100 words and their frequencies in the given Tweets.

Word.	Frequency
pox	40154
monkey	34485
Covid	25992
covid	21385
Monkey	15963
COVID	15078
Pox	10051
monkeypox	6578
people	6223
get	5968
going	3763
vaccine	4040

Monkeypox	3247
got	3004
time	2744
know	2579
shit	2565
virus	2540
go	2331
think	2286
pandemic	2226
flu	2096
want	2008
polio	1939
getting	1985
health	2005
cases	2036
spread	2006
see	1895
world	1823
vaccines	1808
thing	1614
why	1586
mask	1559
years	1518
make	1393
disease	1365
said	1373
work	1403
say	1237
keep	1167
Polio	1128
POX	1133
scared	1216
fear	1155
outbreak	1125
Biden	1131
Ukraine	1064
year	1127
emergency	1146
stop	1119
come	1033
gay	1092
change	1017
spreading	1010
good	1006
coming	985
masks	987
global	973
bad	954
HIV	943
climate	925
trying	897
Why	940

day	898
MONKEY	862
news	903
vaccinated	893
cause	862
stay	827
vax	1001
government	820
care	844
safe	810
else	769
CDC	822
made	785
days	802
country	765
shot	979
Flu	755
sick	765
believe	750
case	758
risk	791
start	717
corona	727
catch	736
control	753
thought	711
saying	725
look	706
diseases	720
Ebola	714
moneypox	689
kids	744
life	699
sex	756
give	695
Lol	691

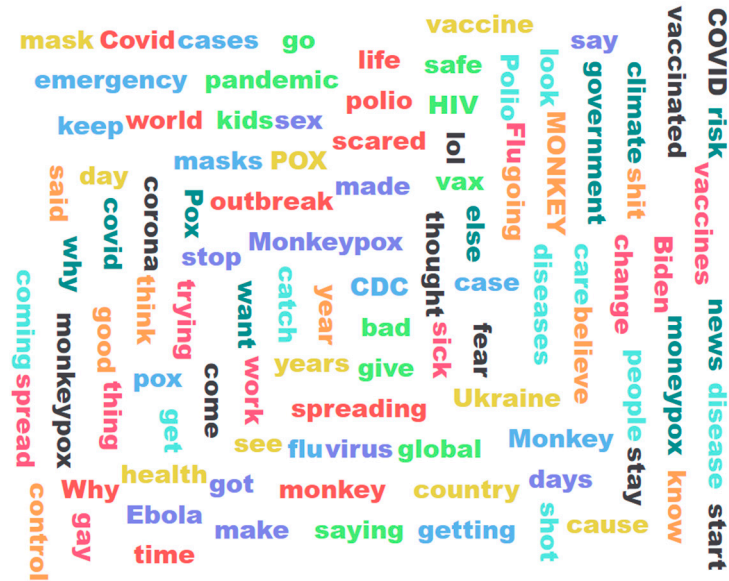


Figure 6. Representation of some of the most frequently used words in these Tweets in the form of a word cloud.

4.3. Comparative Study with Prior Works in this Field

This section presents a comparative study with prior works in this field which were reviewed in Section 2. This comparative study is represented in Table 4. As can be seen from Table 3, the work presented in this paper is the first work in this area that focuses on sentiment analysis of Tweets that focused on COVID-19 and MPox at the same time.

Table 4. Comparative study with prior works in this field.

Work	SA of Tweets about COVID-19	SA of Tweets about MPox
Vijay et al. [44]	✓	
Mansoor et al. [45]	✓	
Pokharel [46]	✓	
Chakraborty et al. [47]	✓	
Shofiya et al. [48]	✓	
Basiri et al. [49]	✓	
Cheeti et al. [50]	✓	
Ridhwan et al. [51]	✓	
Tripathi [52]	✓	
Situala et al. [53]	✓	
Gupta et al. [54]	✓	
Alanezi et al. [55]	✓	
Dubey [56]	✓	
Rahman et al. [57]	✓	
Ainlet et al. [58]	✓	
Slobodin et al. [59]	✓	
Zou et al. [60]	✓	
Alhuzali et al. [61]	✓	
Hussain et al. [62]	✓	
Liu et al. [63]	✓	
Hu et al. [64]	✓	
Khan et al. [65]	✓	
Ahmed et al. [66]	✓	
Lin et al. [67]	✓	
Jang et al. [68]	✓	
Tsao et al. [69]	✓	

Griffith et al. [70]	✓	
Chum et al. [71]	✓	
Kothari et al. [72]	✓	
Barkur et al. [73]	✓	
Afroz et al. [74]	✓	
Hota et al. [75]	✓	
Venigalla et al. [76]	✓	
Paliwal et al. [77]	✓	
Zhou et al. [78]	✓	
Lamsal et al. [79]	✓	
Zhou et al. [80]	✓	
de Melo et al. [81]	✓	
Brum et al. [82]	✓	
de Sousa et al. [83]	✓	
Iparraguirre-Villanueva et al. [84]		✓
Mohbey et al. [85]		✓
Farahat et al. [86]		✓
Sv et al. [87]		✓
Bengesi et al. [88]		✓
Dsouza et al. [89]		✓
Zuhanda et al. [90]		✓
Cooper et al. [91]		✓
Ng et al. [92]		✓
Thakur [this work]	✓	✓

5. Conclusions

The Big Data of Twitter conversations holds the potential for inference of the views, opinions, perspectives, mindset, sentiment, and feedback of the general public towards pandemics, epidemics, viruses, and diseases. This has attracted the attention of researchers in the fields of computer science, big data, data science, epidemiology, healthcare, medicine, and their interrelated areas in the last few years. Various forms of analysis of this Big Data, such as sentiment analysis, hashtag analysis, and frequent keyword analysis, can be seen in prior works in this field that focused on studying Tweets involving some of the virus outbreaks of the past, such as Ebola, E-Coli, Dengue, Human papillomavirus (HPV), Middle East Respiratory Syndrome (MERS), Measles, Zika virus, H1N1, influenza-like illness, swine flu, flu, Cholera, COVID, Listeriosis, cancer, Liver Disease, Inflammatory Bowel Disease, kidney disease, lupus, Parkinson's, Diphtheria, and West Nile virus. The recent outbreaks of COVID-19 and MPox have escalated the use of Twitter for conversations related to these respective viruses. While there have been a few works published in the last few months that focused on performing sentiment analysis of Tweets related to either COVID-19 or MPox, none of the prior works in this field thus far focused on the analysis of Tweets focusing on both COVID-19 and MPox and performing sentiment analysis of the same. To address this challenge, this study presents the findings from a comprehensive sentiment analysis study involving 61,862 Tweets that focused on Mpox and COVID-19 at the same time. The VADER approach was used for performing the sentiment analysis. The results show that almost half the Tweets (actual percentage being 46.88%) involving COVID-19 and MPox had a negative sentiment. It was followed by Tweets that had a positive sentiment (31.97%) and Tweets that had a neutral sentiment (21.14%). This study also presents the findings from hashtag analysis and keyword analysis of these Tweets. The top 50 hashtags featured in all these Tweets are detected and presented in this paper. The top 100 most frequently used words that featured in all these Tweets were also detected using concepts of tokenization and are presented. The findings of frequent word analysis show that some of the commonly used words involved directly refer to either of these viruses. In addition to this, the presence of words such as "Polio", "Biden", "Ukraine", "HIV", "climate" and "Ebola" in the list of the top 100 most frequent words indicate that topics of conversations on Twitter in the context of COVID-19 and MPox also included a high level of interest related to other viruses, President Biden, and Ukraine. Future work in this area would

involve collecting more Tweets over the next months and repeating this study to infer any potential evolutions of public sentiment related to these viruses over the course of time.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The data analyzed in this study are publicly available at <https://doi.org/10.7910/DVN/CR7T5E>

Conflicts of Interest: The author declares no conflict of interest.

References

1. Katz, M.; Nandi, N. Social Media and Medical Education in the Context of the COVID-19 Pandemic: Scoping Review. *JMIR Med. Educ.* 2021, 7, e25892. <https://doi.org/10.2196/25892>.
2. Lee, H.E.; Cho, J. Social Media Use and Well-Being in People with Physical Disabilities: Influence of SNS and Online Community Uses on Social Support, Depression, and Psychological Disposition. *Health Commun.* 2019, 34, 1043–1052. <https://doi.org/10.1080/10410236.2018.1455138>.
3. Kavada, A. Social Media as Conversation: A Manifesto. *Soc. Media Soc.* 2015, 1, 205630511558079. <https://doi.org/10.1177/2056305115580793>.
4. Twitter: Number of Users Worldwide 2024 Available online: <https://www.statista.com/statistics/303681/twitter-users-worldwide/> (accessed on March 26 2023).
5. Hutchinson, A. New Study Shows Twitter Is the Most Used Social Media Platform among Journalists Available online: <https://www.socialmediatoday.com/news/new-study-shows-twitter-is-the-most-used-social-media-platform-among-journa/626245/> (accessed on March 26 2023).
6. Biggest Social Media Platforms 2023 Available online: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/> (accessed on March 26 2023).
7. Lazard, A.J.; Scheinfeld, E.; Bernhardt, J.M.; Wilcox, G.B.; Suran, M. Detecting Themes of Public Concern: A Text Mining Analysis of the Centers for Disease Control and Prevention's Ebola Live Twitter Chat. *Am. J. Infect. Control* 2015, 43, 1109–1111. <https://doi.org/10.1016/j.ajic.2015.05.025>.
8. Bolotova, Y.V.; Lou, J.; Safro, I. Detecting and Monitoring Foodborne Illness Outbreaks: Twitter Communications and the 2015 U.S. Salmonella Outbreak Linked to Imported Cucumbers. *arXiv [stat.AP]* 2017.
9. Gomide, J.; Veloso, A.; Meira, W., Jr; Almeida, V.; Benevenuto, F.; Ferraz, F.; Teixeira, M. Dengue Surveillance Based on a Computational Model of Spatio-Temporal Locality of Twitter. In *Proceedings of the Proceedings of the 3rd International Web Science Conference*; ACM: New York, NY, USA, 2011.
10. Tomaszewski, T.; Morales, A.; Lourentzou, I.; Caskey, R.; Liu, B.; Schwartz, A.; Chin, J. Identifying False Human Papillomavirus (HPV) Vaccine Information and Corresponding Risk Perceptions from Twitter: Advanced Predictive Models. *J. Med. Internet Res.* 2021, 23, e30451. <https://doi.org/10.2196/30451>.
11. Lee, S.Y.; Khang, Y.H.; Lim, H.K. Impact of the 2015 Middle East Respiratory Syndrome Outbreak on Emergency Care Utilization and Mortality in South Korea. *Yonsei Med. J.* 2019, 60, 796–803. <https://doi.org/10.3349/ymj.2019.60.8.796>.
12. Radzikowski, J.; Stefanidis, A.; Jacobsen, K.H.; Croitoru, A.; Crooks, A.; Delamater, P.L. The Measles Vaccination Narrative in Twitter: A Quantitative Analysis. *JMIR Public Health Surveill.* 2016, 2, e1. <https://doi.org/10.2196/publichealth.5059>.
13. Fu, K.-W.; Liang, H.; Saroha, N.; Tse, Z.T.H.; Ip, P.; Fung, I.C.-H. How People React to Zika Virus Outbreaks on Twitter? A Computational Content Analysis. *Am. J. Infect. Control* 2016, 44, 1700–1702. <https://doi.org/10.1016/j.ajic.2016.04.253>.
14. Signorini, A.; Segre, A.M.; Polgreen, P.M. The Use of Twitter to Track Levels of Disease Activity and Public Concern in the US during the Influenza A H1N1 Pandemic. *PLoS One* 2011, 6, e19467. <https://doi.org/10.1371/journal.pone.0019467>.
15. Gesualdo, F.; Stilo, G.; Agricola, E.; Gonfiantini, M.V.; Pandolfi, E.; Velardi, P.; Tozzi, A.E. Influenza-like Illness Surveillance on Twitter through Automated Learning of Naïve Language. *PLoS One* 2013, 8, e82489. <https://doi.org/10.1371/journal.pone.0082489>.
16. Szomszor, M.; Kostkova, P.; de Quincey, E. #swineflu: Twitter Predicts Swine Flu Outbreak in 2009. In *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp. 18–26 ISBN 9783642236341.
17. Alessa, A.; Faezipour, M. Flu Outbreak Prediction Using Twitter Posts Classification and Linear Regression with Historical Centers for Disease Control and Prevention Reports: Prediction Framework Study. *JMIR Public Health Surveill.* 2019, 5, e12383. <https://doi.org/10.2196/12383>.

18. Hirschfeld, D. Twitter Data Accurately Tracked Haiti Cholera Outbreak. *Nature* 2012. <https://doi.org/10.1038/nature.2012.9770>.
19. van der Vyver, A.G. The Listeriosis Outbreak in South Africa: A Twitter Analysis of Public Reaction Available online: <http://www.icmis.net/icmis18/ICMIS18CD/pdf/S198-final.pdf> (accessed on March 26 2023).
20. Thackeray, R.; Burton, S.H.; Giraud-Carrier, C.; Rollins, S.; Draper, C.R. Using Twitter for Breast Cancer Prevention: An Analysis of Breast Cancer Awareness Month. *BMC Cancer* 2013, 13, 508. <https://doi.org/10.1186/1471-2407-13-508>.
21. Da, B.L.; Surana, P.; Schueler, S.A.; Jalaly, N.Y.; Kamal, N.; Taneja, S.; Vittal, A.; Gilman, C.L.; Heller, T.; Koh, C. Twitter as a Noninvasive Bio-Marker for Trends in Liver Disease. *Hepatol. Commun.* 2019, 3, 1271–1280. <https://doi.org/10.1002/hep4.1394>.
22. Szvarca, D.; Tabbara, N.; Masur, J.; Greenfest, A.; Clarke, L.M.; Borum, M.L. P013 Unregulated: Medical Companies Use Social Media to Sell Alternative Treatments for Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* 2019, 25, S8–S8. <https://doi.org/10.1093/ibd/izy393.016>.
23. McLean, R.; Shirazian, S. Women and Kidney Disease: A Twitter Conversation for One and All. *Kidney Int. Rep.* 2018, 3, 767–768. <https://doi.org/10.1016/j.ekir.2018.05.001>.
24. Stens, O.; Weisman, M.H.; Simard, J.; Reuter, K. Insights from Twitter Conversations on Lupus and Reproductive Health: Protocol for a Content Analysis. *JMIR Res. Protoc.* 2020, 9, e15623. <https://doi.org/10.2196/15623>.
25. Cevik, F.; Kilimci, Z.H. Analysis of Parkinson's Disease Using Deep Learning and Word Embedding Models. *acperpro* 2019, 2, 786–797. <https://doi.org/10.33793/acperpro.02.03.86>.
26. Porat, T.; Garaizar, P.; Ferrero, M.; Jones, H.; Ashworth, M.; Vellido, M.A. Content and Source Analysis of Popular Tweets Following a Recent Case of Diphtheria in Spain. *Eur. J. Public Health* 2019, 29, 117–122. <https://doi.org/10.1093/eurpub/cky144>.
27. Sugumaran, R.; Voss, J. Real-Time Spatio-Temporal Analysis of West Nile Virus Using Twitter Data. In *Proceedings of the Proceedings of the 3rd International Conference on Computing for Geospatial Research and Applications*; ACM: New York, NY, USA, 2012.
28. Ksiazek, T.G.; Erdman, D.; Goldsmith, C.S.; Zaki, S.R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J.A.; Lim, W.; et al. A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* 2003, 348, 1953–1966. <https://doi.org/10.1056/NEJMoa030781>.
29. Fauci, A.S.; Lane, H.C.; Redfield, R.R. Covid-19 - Navigating the Uncharted. *N. Engl. J. Med.* 2020, 382, 1268–1269. <https://doi.org/10.1056/NEJMe2002387>.
30. CDC SARS-CoV-2 Variant Classifications and Definitions Available online: <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html> (accessed on March 26 2023).
31. COVID - Coronavirus Statistics - Worldometer Available online: <https://www.worldometers.info/coronavirus/> (accessed on March 26 2023).
32. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *Lancet* 2020, 395, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
33. Iser, B.P.M.; Sliva, I.; Raymundo, V.T.; Poletto, M.B.; Schuelter-Trevisol, F.; Bobinski, F. Suspected COVID-19 Case Definition: A Narrative Review of the Most Frequent Signs and Symptoms among Confirmed Cases. *Epidemiol. Serv. Saude* 2020, 29, e2020233. <https://doi.org/10.5123/S1679-49742020000300018>.
34. Al Khatib, H.A.; Benslimane, F.M.; Elbashir, I.E.; Coyle, P.V.; Al Maslamani, M.A.; Al-Khal, A.; Al Thani, A.A.; Yassine, H.M. Within-Host Diversity of SARS-CoV-2 in COVID-19 Patients with Variable Disease Severities. *Front. Cell. Infect. Microbiol.* 2020, 10, 575613. <https://doi.org/10.3389/fcimb.2020.575613>.
35. McCollum, A.M.; Damon, I.K. Human Monkeypox. *Clin. Infect. Dis.* 2014, 58, 260–267.
36. Magnus, P. von; Andersen, E.K.; Petersen, K.B.; Birch-Andersen, A. A Pox-like Disease in Cynomolgus Monkeys. *Acta Pathol. Microbiol. Scand.* 2009, 46, 156–176. <https://doi.org/10.1111/j.1699-0463.1959.tb00328.x>.
37. Charniga, K.; Masters, N.B.; Slayton, R.B.; Gosdin, L.; Minhaj, F.S.; Philpott, D.; Smith, D.; Gearhart, S.; Alvarado-Ramy, F.; Brown, C.; et al. Estimating the Incubation Period of Monkeypox Virus during the 2022 Multi-National Outbreak. *bioRxiv* 2022.
38. Jezek, Z.; Szczeniowski, M.; Paluku, K.M.; Mutombo, M. Human Monkeypox: Clinical Features of 282 Patients. *J. Infect. Dis.* 1987, 156, 293–298. <https://doi.org/10.1093/infdis/156.2.293>.
39. Centers for Disease Control and Prevention (CDC) Update: Multistate Outbreak of Monkeypox—Illinois, Indiana, Kansas, Missouri, Ohio, and Wisconsin, 2003. *MMWR Morb. Mortal. Wkly. Rep.* 2003, 52, 561–564.
40. Vaughan, A.; Aarons, E.; Astbury, J.; Balasegaram, S.; Beadsworth, M.; Beck, C.R.; Chand, M.; O'Connor, C.; Dunning, J.; Ghebrehewet, S.; et al. Two Cases of Monkeypox Imported to the United Kingdom, September 2018. *Euro Surveill.* 2018, 23. <https://doi.org/10.2807/1560-7917.ES.2018.23.38.1800509>.

41. Yinka-Ogunleye, A.; Aruna, O.; Dalhat, M.; Ogoina, D.; McCollum, A.; Disu, Y.; Mamadu, I.; Akinpelu, A.; Ahmad, A.; Burga, J.; et al. Outbreak of Human Monkeypox in Nigeria in 2017-18: A Clinical and Epidemiological Report. *Lancet Infect. Dis.* 2019, 19, 872–879. [https://doi.org/10.1016/S1473-3099\(19\)30294-4](https://doi.org/10.1016/S1473-3099(19)30294-4).
42. CDC 2022 Mpox Outbreak Global Map Available online: <https://www.cdc.gov/poxvirus/mpox/response/2022/world-map.html> (accessed on March 26 2023).
43. Medhat, W.; Hassan, A.; Korashy, H. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Eng. J.* 2014, 5, 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>.
44. Vijay, T.; Chawla, A.; Dhanka, B.; Karmakar, P. Sentiment Analysis on COVID-19 Twitter Data. In *Proceedings of the 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*; IEEE, 2020; pp. 1–7.
45. Mansoor, M.; Gurumurthy, K.; U, AR; Prasad, V.R.B. Global Sentiment Analysis of COVID-19 Tweets over Time. *arXiv [cs.CL]* 2020.
46. Pokharel, B.P. Twitter Sentiment Analysis during Covid-19 Outbreak in Nepal. *SSRN Electron. J.* 2020. <https://doi.org/10.2139/ssrn.3624719>.
47. Chakraborty, K.; Bhatia, S.; Bhattacharyya, S.; Platos, J.; Bag, R.; Hassanien, A.E. Sentiment Analysis of COVID-19 Tweets by Deep Learning Classifiers-A Study to Show How Popularity Is Affecting Accuracy in Social Media. *Appl. Soft Comput.* 2020, 97, 106754. <https://doi.org/10.1016/j.asoc.2020.106754>.
48. Shofiya, C.; Abidi, S. Sentiment Analysis on COVID-19-Related Social Distancing in Canada Using Twitter Data. *Int. J. Environ. Res. Public Health* 2021, 18, 5993. <https://doi.org/10.3390/ijerph18115993>.
49. Basiri, M.E.; Nemati, S.; Abdar, M.; Asadi, S.; Acharrya, U.R. A Novel Fusion-Based Deep Learning Model for Sentiment Analysis of COVID-19 Tweets. *Knowl. Based Syst.* 2021, 228, 107242. <https://doi.org/10.1016/j.knosys.2021.107242>.
50. Cheeti, S.S.; Li, Y.; Hadaegh, A. Twitter Based Sentiment Analysis of Impact of Covid-19 on Education Globally. *Int. J. Artif. Intell. Appl.* 2021, 12, 15–24. <https://doi.org/10.5121/ijai.2021.12302>.
51. Mohamed Ridhwan, K.; Hargreaves, C.A. Leveraging Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak in Singapore. *International Journal of Information Management Data Insights* 2021, 1, 100021. <https://doi.org/10.1016/j.jjimei.2021.100021>.
52. Tripathi, M. Sentiment Analysis of Nepali COVID19 Tweets Using NB, SVM AND LSTM. September 2021 2021, 3, 151–168. <https://doi.org/10.36548/jaicn.2021.3.001>.
53. Sitaula, C.; Basnet, A.; Mainali, A.; Shahi, T.B. Deep Learning-Based Methods for Sentiment Analysis on Nepali COVID-19-Related Tweets. *Comput. Intell. Neurosci.* 2021, 2021, 2158184. <https://doi.org/10.1155/2021/2158184>.
54. Gupta, P.; Kumar, S.; Suman, R.R.; Kumar, V. Sentiment Analysis of Lockdown in India during COVID-19: A Case Study on Twitter. *IEEE Trans. Comput. Soc. Syst.* 2021, 8, 992–1002. <https://doi.org/10.1109/tcss.2020.3042446>.
55. Alanezi, M.A.; Hewahi, N.M. Tweets Sentiment Analysis during COVID-19 Pandemic. In *Proceedings of the 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*; IEEE, 2020; pp. 1–6.
56. Dubey, A.D. Twitter Sentiment Analysis during COVID19 Outbreak. *SSRN Electron. J.* 2020. <https://doi.org/10.2139/ssrn.3572023>.
57. Rahman, M.M.; Islam, M.N. Exploring the Performance of Ensemble Machine Learning Classifiers for Sentiment Analysis of COVID-19 Tweets. In *Advances in Intelligent Systems and Computing*; Springer Singapore: Singapore, 2022; pp. 383–396.
58. Ainley, E.; Witwicki, C.; Tallett, A.; Graham, C. Using Twitter Comments to Understand People's Experiences of UK. Health Care during the COVID-19 Pandemic: Thematic and Sentiment Analysis. *J. Med. Internet Res.* 2021, 23, e31101. <https://doi.org/10.2196/31101>.
59. Slobodin, O.; Plohotnikov, I.; Cohen, I.-C.; Elyashar, A.; Cohen, O.; Puzis, R. Global and Local Trends Affecting the Experience of US and UK Healthcare Professionals during COVID-19: Twitter Text Analysis. *Int. J. Environ. Res. Public Health* 2022, 19, 6895. <https://doi.org/10.3390/ijerph19116895>.
60. Zou, C.; Wang, X.; Xie, Z.; Li, D. Public Reactions towards the COVID-19 Pandemic on Twitter in the United Kingdom and the United States. *medRxiv* 2020. <https://doi.org/10.1101/2020.07.25.20162024>.
61. Alhuzali, H.; Zhang, T.; Ananiadou, S. Emotions and Topics Expressed on Twitter during the COVID-19 Pandemic in the United Kingdom: Comparative Geolocation and Text Mining Analysis. *J. Med. Internet Res.* 2022, 24, e40323. <https://doi.org/10.2196/40323>.
62. Hussain, Z.; Sheikh, Z.; Tahir, A.; Dashtipour, K.; Gogate, M.; Sheikh, A.; Hussain, A. Artificial Intelligence-Enabled Social Media Analysis for Pharmacovigilance of COVID-19 Vaccinations in the United Kingdom: Observational Study. *JMIR Public Health Surveill.* 2022, 8, e32543. <https://doi.org/10.2196/32543>.
63. Liu, S.; Liu, J. Public Attitudes toward COVID-19 Vaccines on English-Language Twitter: A Sentiment Analysis. *Vaccine* 2021, 39, 5499–5505. <https://doi.org/10.1016/j.vaccine.2021.08.058>.

64. Hu, T.; Wang, S.; Luo, W.; Zhang, M.; Huang, X.; Yan, Y.; Liu, R.; Ly, K.; Kacker, V.; She, B.; et al. Revealing Public Opinion towards COVID-19 Vaccines with Twitter Data in the United States: Spatiotemporal Perspective. *J. Med. Internet Res.* 2021, 23, e30854. <https://doi.org/10.2196/30854>.
65. Khan, R.; Rustam, F.; Kanwal, K.; Mehmood, A.; Choi, G.S. US Based COVID-19 Tweets Sentiment Analysis Using TextBlob and Supervised Machine Learning Algorithms. In *Proceedings of the 2021 International Conference on Artificial Intelligence (ICAI)*; IEEE, 2021; pp. 1–8.
66. Ahmed, M.E.; Rabin, M.R.I.; Chowdhury, F.N. COVID-19: Social Media Sentiment Analysis on Reopening. *arXiv [cs.SI]* 2020.
67. Lin, S.J.; Bustos, V.P.; Comer, C.D.; Manstein, S.M.; Laikhter, E.; Shiah, E.; Xun, H.; Lee, B.T. Twitter Voices: Twitter Users' Sentiments and Emotions about COVID-19 Vaccination within the United States. *Eur. J. Environ. Public Health* 2022, 6, em0096. <https://doi.org/10.21601/ejeph/11499>.
68. Jang, H.; Rempel, E.; Roe, I.; Adu, P.; Carenini, G.; Janjua, N.Z. Tracking Public Attitudes toward COVID-19 Vaccination on Tweets in Canada: Using Aspect-Based Sentiment Analysis. *J. Med. Internet Res.* 2022, 24, e35016. <https://doi.org/10.2196/35016>.
69. Tsao, S.-F.; MacLean, A.; Chen, H.; Li, L.; Yang, Y.; Butt, Z.A. Public Attitudes during the Second Lockdown: Sentiment and Topic Analyses Using Tweets from Ontario, Canada. *Int. J. Public Health* 2022, 67, 1604658. <https://doi.org/10.3389/ijph.2022.1604658>.
70. Griffith, J.; Marani, H.; Monkman, H. COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework. *J. Med. Internet Res.* 2021, 23, e26874. <https://doi.org/10.2196/26874>.
71. Chum, A.; Nielsen, A.; Bellows, Z.; Farrell, E.; Durette, P.-N.; Banda, J.M.; Cupchik, G. Changes in Public Response Associated with Various COVID-19 Restrictions in Ontario, Canada: Observational Infoveillance Study Using Social Media Time Series Data. *J. Med. Internet Res.* 2021, 23, e28716. <https://doi.org/10.2196/28716>.
72. Kothari, A.; Foisey, L.; Donelle, L.; Bauer, M. How Do Canadian Public Health Agencies Respond to the COVID-19 Emergency Using Social Media: A Protocol for a Case Study Using Content and Sentiment Analysis. *BMJ Open* 2021, 11, e041818. <https://doi.org/10.1136/bmjopen-2020-041818>.
73. Barkur, G.; Vibha; Kamath, G.B. Sentiment Analysis of Nationwide Lockdown Due to COVID 19 Outbreak: Evidence from India. *Asian J. Psychiatr.* 2020, 51, 102089. <https://doi.org/10.1016/j.ajp.2020.102089>.
74. Afroz, N.; Boral, M.; Sharma, V.; Gupta, M. Sentiment Analysis of COVID-19 Nationwide Lockdown Effect in India. In *Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*; IEEE, 2021; pp. 561–567.
75. Hota, HS; Sharma, D.K.; Verma, N. Lexicon-Based Sentiment Analysis Using Twitter Data. In *Data Science for COVID-19*; Kose, U., Gupta, D., de Albuquerque, V.H.C., Khanna, A., Eds.; Elsevier: San Diego, CA, 2021; pp. 275–295 ISBN 9780128245361.
76. Venigalla, A.S.M.; Chimalakonda, S.; Vagavolu, D. Mood of India during Covid-19 - an Interactive Web Portal Based on Emotion Analysis of Twitter Data. In *Proceedings of the Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing*; ACM: New York, NY, USA, 2020.
77. Paliwal, S.; Parveen, S.; Afshar Alam, M.; Ahmed, J. Sentiment Analysis of COVID-19 Vaccine Rollout in India. In *ICT Systems and Sustainability*; Springer Nature Singapore: Singapore, 2022; pp. 21–33 ISBN 9789811659867.
78. Zhou, J.; Yang, S.; Xiao, C.; Chen, F. Examination of Community Sentiment Dynamics Due to COVID-19 Pandemic: A Case Study from a State in Australia. *SN Comput. Sci.* 2021, 2, 201. <https://doi.org/10.1007/s42979-021-00596-7>.
79. Lamsal, R.; Read, MR; Karunasekera, S. A Twitter Narrative of the COVID-19 Pandemic in Australia. *arXiv [cs.SI]* 2023.
80. Zhou, J.; Zogan, H.; Yang, S.; Jameel, S.; Xu, G.; Chen, F. Detecting Community Depression Dynamics Due to COVID-19 Pandemic in Australia. *IEEE Trans. Comput. Soc. Syst.* 2021, 8, 982–991. <https://doi.org/10.1109/tcss.2020.3047604>.
81. de Melo, T.; Figueiredo, C.M.S. Comparing News Articles and Tweets about COVID-19 in Brazil: Sentiment Analysis and Topic Modeling Approach. *JMIR Public Health Surveill.* 2021, 7, e24585. <https://doi.org/10.2196/24585>.
82. Brum, P.; Cândido Teixeira, M.; Vimieiro, R.; Araújo, E.; Meira, W., Jr; Lobo Pappa, G. Political Polarization on Twitter during the COVID-19 Pandemic: A Case Study in Brazil. *Soc. Netw. Anal. Min.* 2022, 12, 140. <https://doi.org/10.1007/s13278-022-00949-x>.
83. de Sousa, A.M.; Becker, K. Pro/Anti-vaxxers in Brazil: a temporal analysis of COVID vaccination stance in Twitter. In *Proceedings of the Anais do IX Symposium on Knowledge Discovery, Mining and Learning (KDMiLe 2021)*; Sociedade Brasileira de Computação - SBC, 2021; pp. 105–112.
84. Iparaguire-Villanueva, O.; Alvarez-Risco, A.; Herrera Salazar, J.L.; Beltozar-Clemente, S.; Zapata-Paulini, J.; Yáñez, J.A.; Cabanillas-Carbonell, M. The Public Health Contribution of Sentiment Analysis of

- Monkeypox Tweets to Detect Polarities Using the CNN-LSTM Model. *Vaccines* (Basel) 2023, 11, 312. <https://doi.org/10.3390/vaccines11020312>.
85. Mohbey, K.K.; Meena, G.; Kumar, S.; Lokesh, K. A CNN-LSTM-Based Hybrid Deep Learning Approach to Detect Sentiment Polarities on Monkeypox Tweets. *arXiv [cs.CV]* 2022.
 86. Farahat, R.A.; Yassin, M.A.; Al-Tawfiq, J.A.; Bejan, CA; Abdelazeem, B. Public Perspectives of Monkeypox in Twitter: A Social Media Analysis Using Machine Learning. *New Microbes New Infect.* 2022, 49, 101053. <https://doi.org/10.1016/j.nmni.2022.101053>.
 87. Sv, P.; Ittamalla, R. What Concerns the General Public the Most about Monkeypox Virus? - A Text Analytics Study Based on Natural Language Processing (NLP). *Travel Med. Infect. Dis.* 2022, 49, 102404. <https://doi.org/10.1016/j.tmaid.2022.102404>.
 88. Bengesi, S.; Oladunni, T.; Olusegun, R.; Audu, H. A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion from Twitter Tweets. *IEEE Access* 2023, 11, 11811–11826. <https://doi.org/10.1109/access.2023.3242290>.
 89. Dsouza, VS; Rajkhowa, P.; Mallya, BR; Raksha, D.S.; Mrinalini, V.; Cauvery, K.; Raj, R.; Toby, I.; Pattanshetty, S.; Brand, H. A Sentiment and Content Analysis of Tweets on Monkeypox Stigma among the LGBTQ+ Community: A Cue to Risk Communication Plan. *Dialogues Health* 2023, 2, 100095. <https://doi.org/10.1016/j.dialog.2022.100095>.
 90. Analysis of Twitter User Sentiment on the Monkeypox Virus Issue Using the Nrc Lexicon Available online: <https://doi.org/10.35335/mantik.v6i4.3502> (accessed on March 26 2023).
 91. Cooper, L.N.; Radunsky, A.P.; Hanna, J.J.; Most, Z.M.; Perl, T.M.; Lehmann, C.U.; Medford, R.J. Analyzing an Emerging Pandemic on Twitter: Monkeypox. *Open Forum Infect. Dis.* 2023. <https://doi.org/10.1093/ofid/ofad142>.
 92. Ng, QX; Yau, C.E.; Lim, Y.L.; Wong, L.K.T.; Liew, T.M. Public Sentiment on the Global Outbreak of Monkeypox: An Unsupervised Machine Learning Analysis of 352,182 Twitter Posts. *Public Health* 2022, 213, 1–4. <https://doi.org/10.1016/j.puhe.2022.09.008>.
 93. Thakur, N. A Large-Scale Dataset of Twitter Chatter about Online Learning during the Current COVID-19 Omicron Wave. *Data* (Basel) 2022, 7, 109. <https://doi.org/10.3390/data7080109>.
 94. Hutto, C.; Gilbert, E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 2014, 8, 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>.
 95. Mierswa, I.; Wurst, M.; Klinkenberg, R.; Scholz, M.; Euler, T. YALE: Rapid Prototyping for Complex Data Mining Tasks. In *Proceedings of the Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*; ACM: New York, NY, USA, 2006.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.