Pandemic Perspectives: A Review of Sentiment Analysis Approaches on COVID-19

Satvika
Computer Engineering Department
TIT&S Bhiwani
Haryana, India

Akhil Kaushik
Computer Engineering Department
TIT&S Bhiwani
Haryana, India

ABSTRACT

The social media has enabled the mundane man to raise his awareness, concern and voice for and against any issue. The social networking platforms empowered by cellphones have given tremendous rise in the involvement of people in countless affairs. During the lockdown and later periods, mostly persons have given their opinions regarding myriad topics related to COVID-19 like vaccination, masks, social distancing, etc. Many articles have been published on sentiment analysis related to this deadly virus. A standard systematic literature review is conducted in this paper, where multiple research articles on given subject are collected and analyzed. This paper discusses the chief findings in the examined papers, highlights the key points and conducts a critical assessment.

Keywords
COVID-19, Coronavirus, Disease, Pandemic, Sentiment Analysis

1. INTRODUCTION

In the late 2019, a pandemic hit planet Earth and took the entire population by a storm. This deadly disease called the "Coronavirus Disease 2019" or "COVID-19" in short is a highly contagious disease caused by the novel coronavirus popularly known as "SARS-COV-2". The outbreak of this fatal disease is believed to come from a seafood market in Wuhan, China, from where it spread to all parts of the globe due to tourism chiefly. The favourite tourist destination of the Chinese people in that ear was Europe especially Italy and this caused blowout of COVID-19 to other countries. This travel of people across the globe caused the disease to disperse to most of the continents and infected millions of people in numerous countries. Hence COVID-19 that initially started as an epidemic was declared as a pandemic by World Health Organization (WHO) on March 11th, 2020 [1].

The coronavirus disease is a viral disease that entered the human population through natural zoonosis and most of the people who are infected through this disease experience respiratory disorders. It primarily spreads through the respiratory droplets that are generated by an infected person when he/ she sneezes, coughs, talks, etc. Some secondary reasons for COVID-19 spread are aerosol infection through air, direct contact to infected person and fomites i.e. contaminated objects [2]. It is vital to know that some people can easily recover from this disease without any proper medication, but the older generation persons may face serious sickness and command vital medical attention. The symptoms of COVID-19 are myriad ranging from breathing difficulties, fever, cough, tiredness, loss of smell and taste, headache, diarrhea, sore threat, chest pain to shortness of breath. This symptoms could surface from day 1 to day 14 after the exposure of this lethal virus, though the average is 5-6 days. However, it is significant to understand that approximate one-third of the infected persons may not show any noticeable symptoms [3]. However, some people become sternly ill and their respiratory systems especially lungs are severely affected, which may lead to lack to oxygen and breathlessness. The effects of COVID-19 are chiefly notable in the persons with higher age and greater weight, as they are more gravely affected by this deadly disease. Also, some people that have a history of medical issues face profound consequences upon exposing to the novel coronavirus. This virus got its name from the peculiar crown like shape or corona of the spike proteins. It also contains protrusions i.e. bits which stick out of the membrane protein and are responsible for infecting the hosts. The following diagram shows the wholesome structure of COVID-19 virus:
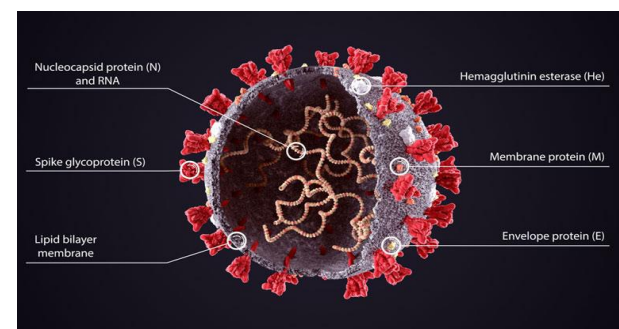

Fig 1. Internal Structure of COVID-19 [4]

Before the vaccines came to the rescue, there were some common precautions that were popular to avoid this ghastly virus like wearing masks, social distancing, regular handwashing, personal hygiene, surface cleaning, self-isolation, proper ventilation and avoiding crowded public places. Nevertheless, development of vaccination was being done simultaneously by gathering information regarding Severe Acute Respiratory Syndrome (SARS) and Middle East respiratory syndrome (MERS). In the year 2020, the first ever vaccination was developed to avoid severe illness and were made available to the public. Later, myriad vaccines were developed that could deliver safety in one dose except of two-doses offered by the rival vaccination companies. It is crucial to understand that the effect of the vaccination doses diminished with time and an extra booster dose was required to improve the immunity [5].

1.1 Role of Social Media
Social media started as a way to connect school and college friends that have moved away in search of job, but later it became a tool for influencing others and spreading the political beliefs. Nevertheless, social media has revolutionized the way humans remain connected, communicate, consume information, enjoy in spare time and even disperse ideology,

whether of a particular religion, sect or political party. Hence, it can be said that the social media is now playing a multifaceted role in modern society. It serves as a commanding platform for enabling people to share their thoughts, ideas, and experiences instantly across geographical confines, thus providing global connectivity.

On the crest side, social media fosters community building by providing ample space for like-minded individuals to unite over common causes, interests or ideologies. It also amplifies voices of the unheard or marginalized groups a big platform to advocate regarding their rights or raise awareness about critical issues, which are unheard usually. Also, social media has renovated the educational arena as countless platforms are available to enhance learning, share educational resources and academic discussions. Not to forget that social media is also indispensable instrument for businesses as it helps the corporates to reach the mundane man, understand his/ her requirements and then build the products or provide services accordingly. Hence, it is acting as a vital apparatus customer engagement, brand developing and marketing [6].

On the flip side, social media is considered a substantial addiction and nowadays everyone (whether children, young or old) is clinging on to their cellphones at all the time. It is not uncommon that every member of a modern family is found sitting alone and busy on their smartphones rather than sitting like a family and interacting with each other. Another drawback is cyber bullying that may lead to digital abuse, and ultimately affecting privacy and mental health of netizens [7]. However, despite all the primary quandaries, social media is here to stay and will definitely act as critical catalyst for communication, connectivity and contains colossal potential to shape the socio-economic and socio-cultural fabric of the modern world.

1.2 Working of Sentiment Analysis

Sentiment analysis, also known as opinion mining, is a specialized branch of natural language processing (NLP) that focuses on identifying and categorizing emotions, opinions, or sentiments expressed in textual data. This extreme branch of artificial intelligence is highly effective in recognizing human emotions expressed through texts of various natural languages or even emojis. This attained knowledge can be easily applied to the domains of customer service and marketing of new or available products for gauging public opinions or achieving better customer satisfaction. The input for this highly enigmatic process can be taken from numerous sources like surveys, customer reviews or the comments collected from multiple social networking sites.

After the initial gathering of data, next comes the phase of preprocessing that encompasses tokenization (dividing complex sentences into individual parts called tokens), stemming/ lemmatizing (tracing out the more optimal root word) and eradicating irrelevant contents like stopwords (common but less important words) and punctuation symbols [8]. Subsequently, the data is submitted for sentiment analysis, which is the ultimate goal here and visualized in figure 2 below. To employ sentiment analysis on the processed data, there are variety of options, some notable mentions are lexicon-based approaches, machine learning, deep learning, pre-trained models and hybrid approaches.

Fig 2. Judging of various emotions using Sentiment Analysis [9]

The lexicon-based approaches is based on list of pre-defined list of words with associated sentiment polarities and sometimes even intensity of the sentiments. However, this method is considered quite basic and struggles when the data contains sarcasm or irony [10]. The next method – machine learning utilizes numerous algorithms like Naive Bayes, Support Vector Machines, Logistic Regression, etc., to undertake the labelled data, learn the underlying patterns of it and then find the sentiments of unseen data based on it. A more sophisticated version of machine learning is deep learning which uses the neural networks to do the same. It is actually a subset of machine learning, but far more advanced in terms of efficiency. Then comes the pre-trained models like BERT, RoBERTa, etc., which are fine-tuned on comparatively small and domain-specific datasets. For this, enhanced computational resources are required along with some labelled data. It is noteworthy that a hybrid approach that combines some or all of the above-mentioned methods can also be used for sentiment analysis. These superior sentiment analysis systems are remarkable in considering aspects like context, sarcasm, and negations, which can significantly alter the sentiment of a phrase. For instance, the sentence "The weather is not bad for Cricket" conveys a positive sentiment despite the presence of a negative word. These advanced approaches are also capable of doing Multilingual sentiment analysis i.e., undertaking data in multiple natural languages. Nevertheless, the insights generated through sentiment analysis empower organizations to make informed decisions, improve customer experiences, and adapt strategies based on real-time emotional feedback from their audience [11].

1.3 Sentiment Analysis during COVID Epidemic

As it is quite an established fact that the COVID-19 pandemic was a historic event that traumatized millions across the world, hence sentiment analysis act as an indispensable tool that provided deep insights into public opinions, human emotions and their behavioral trends on an unprecedented scale. When the people were imprisoned inside their homes during lockdown periods, they used social media as their voice. This vastly helped the researchers, social scientists, private organizations and even the governments to study and analyze public sentiments in real time and come up with effective strategies to counter the fear and uncertainty in the minds of the public. The colossal amount of data generated on the social networking platforms made them understand how people are perceiving and reacting to public safety and health measures like wearing masks, social distancing, lockdowns and vaccination [12]. For instance, during initial phases of vaccination, there were negative sentiments regarding it that the vaccination has strong side effects. Hence, the government crafted campaigns to alleviate the misconceptions and unnecessary fears related to vaccination. Additionally, the businesses and corporate organizations also used opinion

mining in this ghastly era to better understand their customers and customize their products in a better way. For example: Designers of women ethnic wear in India started giving matching masks to attract more customers [13]. On the other hand, healthcare professionals used sentiment analysis to better comprehend the psychological influence of the pandemic on common people by studying and tracking loneliness, anxiety, etc., in the comments extracted from social networking sites. This also helped the policy makers to detect and counter the misinformation and rumors by flagging the dangerous content. However, it could be challenging due to the presence of sarcasm and satire used in the social media posts. But overall, it can be said that sentiment analysis acted as a crucial instrument for tracing harmful content, fostering resilience and superior response to the evolving landscape of the globe in dearth situations.

2. REVIEW PROTOCOL

The systematic literature review of this paper includes the following parameters:

• Development of review protocol
• Conducting reviews
• Studying and analyzing the results
• Discussion of the results

The first step in the development of review protocol is framing the vital research questions followed by searching the important question papers from myriad journals and conferences that answer these questions. Subsequently, the related data collected is studied further and the findings are evaluated. The chief methodology of conducting a comprehensive literature review is selection of credible sources from which the research papers are collected. Hence, the research papers for this review are collected from the various trustworthy electronic sources like IEEE Explore, Springer, ScienceDirect, ACM Digital Library, Google Scholar, etc., using appropriate keywords like COVID-19, Sentiment Analysis, Opinion Mining, etc. In total, 48 research papers were downloaded using a combination of these keywords. The following figure 3 demonstrates the year-wise publication of these research articles.
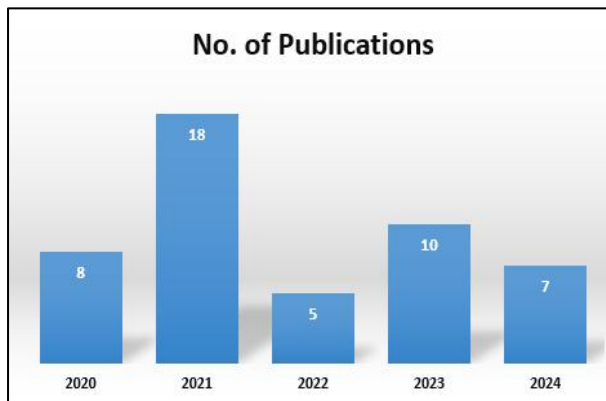

Fig 3. The Year-wise Paper Publications Included in Survey

After defining the review protocol and downloading of more than 75 papers, full text reading was done to remove the unrelated papers. This full text reading was conducted to extract meaningful and valuable information; and based on this the 48 research papers were selected.

Table 1. Data Extraction Process

Data Item	Description
Title	Title of Research Paper
Year	Year of Publication
Data Source	Source from which data is extracted (Twitter, Facebook, Reddit, etc.)
Goal	Primary Goal of Research
Language	Primary natural language in which data is written
Scope	Whether study is specific to a geographical location (a country) or global
Volume	Size of data considered for analysis
Aspect Discussed	Either a specific topic like Covid vaccine, fake news/ misinformation spread is discussed or broadly COVID-19
Technology	If the underlying technology is Machine Learning based, Deep learning based, Large Language Model based or Lexicon-based. Also myriad types of sentiment analyzers (VADER, TextBlob, SentiWordNet, etc.) or kind of word embedding used in the study

The following table 1 displays the chief data items which were considered to finally shape the review.

3. LITERATURE REVIEW

The term "sentiment analysis" was first used by Nasukawa and Yi in their research article published in 2003 and the term "opinion mining" was also introduced in the same year by Dave et. al. [14]. Both terminologies are used to determine emotions of people towards any peculiar entity like person, place or event. This incredible idea caused turbulence in the field of research and innumerable academicians and scholars turned to utilize in myriad fields. Just like, Păvăloaia et. al. (2019) demonstrated that sentiment analysis can be applied in the sphere of user preferences by the businesses. The study suggests that social media posts can be combined with customer-centric management systems to form a novel Social Customer Relationship Management (sCRM) that will be extremely beneficial for the existing and potential customers. A total of six social media sites were included in this research: Facebook, Twitter, Instagram, YouTube, Google+, Pinterest and two types of media were included: photos and videos [15].

A brilliant work by Núñez et. al. (2020) conducted a similar survey on 919 research works from 2010 to 2019 were extracted from Web of Science. This bibliometric survey was done to examine the role of opinion mining in the advertising domain. The article clearly defines that the in understanding the emotion in the marketing and advertising field through sentiment analysis takes a huge leap and will reap higher profits than ever before [16].

Shaik et. al. (2022) have highlighted the importance of using sentiment analysis in the field of Education. Here, the authors have highlighted that it can be vastly applied to understand the student opinions and boost their teaching-learning practices. It is done at various levels like document, sentence, entity and aspect by means of several lexicon and corpus-based approaches. The paper also defines how online education during the pandemic revolutionized the education sector and why opinion mining is more significant during the covid era. The researchers also drew attention to the practice of involving students in the teaching-learning process and emphasized that the feedback of students can exponentially improve the overall

pedagogical structure [17]. Similar approach was taken Soong et. al. (2019) to identify the recent approaches of sentiment analysis that could be applied on the data mined from social networks. The authors also advocate that the evolution of social networks and increased web activity of mundane people, opinions and emotions play a key role in purchasing of items/ products and the opinion mining is different from emotion mining. The scholars also state that lexicon-based approach outperforms supervised machine learning approach not only in terms of standard performance metrices but also in terms of economy of time and efforts [18].

Stefanis et. al. (2023) studied the comments and posts on the Greek National Public Health Organization (EODY) Facebook page from November 2021 to January 2022. This daily surveillance report of COVID-19 disease was scrutinized on regular basis and with the help of machine learning algorithms, the underlying sentiments were classified into positive, negative and neutral. The overall opinion mining shows that nearly 57% of the comments were negative and merely 9% were positive. During this particular pandemic era, two prime factors were gender and time, which determine the public belief on vital medical topics. Out of all the machine learning classifiers and neural networks, Bayes point machine yielded more productivity. The following figure 4 depicts the most frequent words shown in the evaluation of audience perspective. A major dearth is the need of comparison with other research papers on associated topics [19].
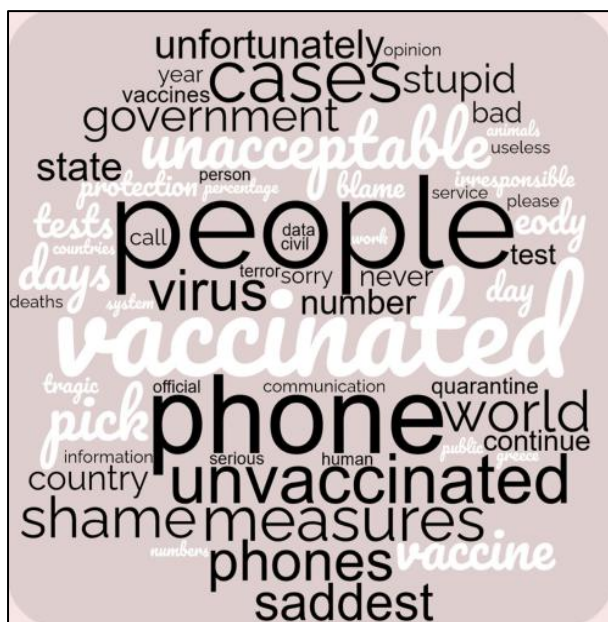

Fig 4. Word cloud showing frequent words in public perception [19]

Another masterpiece work by Iwendi et. al. (2022) tries to differentiate between real and fake news. The real news were collected by Government and authorized health sites, while the fake news were collected from social media platforms. After collecting the data in multimedia form, 39 peculiar features were extracted and then deep learning models were employed on these features which enhanced the accuracy by approximately 17%. The other performance parameters also depict improved values and the proposed system outperforms the standard machine learning algorithm counterparts on the same dataset. Overall, this work establishes a strong defense

against all major misinformation like conspiracy theories, fake cures, false vaccination side-effects, etc. [20].

Dubey (2020) has done a country-wise sentiment analysis of the tweets by extracting tweets for a certain lockdown period of 12 countries. The research also defines how netizens residing in various countries are affected by dire covid situations but are still reflecting positive and hopeful emotions. In this work, tweets in English language are only considered to maintain a standard and NRC Lexicon methodology is implemented for eight different emotions. An interesting finding in this paper was that Donald Trump was the most active user across all 12 countries [21].

The scholarly article by Manguri, Ramadhan & Amin (2020) has done a twitter sentiment analysis on the COVID-19 data downloaded from Twitter using Tweepy library. The TextBlob library has been extensively used in this study for sentiment analysis and also visualized for the period of April 2024. A dataset of 5,30,232 tweets is formed by using vital keywords like coronavirus and COVID-19 [22].

In the research paper by Nemes & Kiss (2020) performed opinion mining on the comments, hashtags, tweets and posts obtained from the social networking sites regarding COVID-19. The major finding of this research is that the authors have tried to trace not only the underlying sentiments, but also the manifestations of the users on Twitter platform. The classification of the tweets is done using deep learning techniques i.e. Recurrent Neural Network (RNN) which was also compared to the TextBlob model. Surprisingly, both models gave good efficiency, but the TextBlob model resulted in more than 30% neutral results than the RNN model [23].

An original research work by Gupta et. al. (2021) analyzed tweets by Indian netizens from April 2020 spell using Tweepy API. These tweets were annotated using the VADER and TextBlob lexicons and then preprocessed before submitting them to eight supervised machine learning classifiers, out of which LinearSVC and unigram showed highest accuracy (84.4%). The research suggested that the Indian commoners were more positive during this time and supported the strong steps taken by the Indian government during the second lockdown phase. The major drawback of this research was the non-consideration of emoticons and hashtags used in the tweets [24].

Pedrosa, Núñez, & Peláez (2020) conducted a semantic investigation in Spanish society during the pandemic time. The researchers scrutinized the digital ecosystems during the months of March and April of the year 2020 and formed a dataset of 1,06, 261 posts with the help of APIs and web-scraping techniques. The main idea in this paper is effective risk communication whose objective is to reduce the communication gap between information issuers and common people so that no fear or panic is created in the public. The paper has considered all major players related to public health like international and national health organizations, non-governmental organizations, broadcast and social media and mundane people. A thought-provoking discovery is that seven out of ten Spaniards form their opinions based on news broadcasted on the television [25].

This alternative research by Boon & Skunkan, (2020) has identified three focal aspects of public consciousness and concern about the SARS-COV-2 virus spread. The paper also discusses the three main stages of the pandemic and the main concerns in these periods. The results show that most people express negative emotions on social media and there is dire

need for more proactive public health response. This study included approximately 1,07,990 tweets and used topic modelling technique i.e. Latent Dirichlet Allocation (LDA) to conduct the research [26].

Al-Shabi (2020) recommended that lexicon-based methodology is far superior to the machine learning techniques for the task of opinion mining. The study undertakes the standard twitter datasets (Stanford Twitter Sentiment Set and Sanders Twitter Dataset Test Set) and employs five renowned lexicons namely SentiWordNet, VADER, SentiStrength, Liu and Hu opinion lexicon and AFINN-111. All these lexicon-based techniques were employed on both datasets and their performance based on Accuracy, Precision, Recall and F1-score were measured; out of which VADER indicated best accuracy rates [27].

The research paper by Pastor (2020) examined the pros and cons of the Community Quarantine that was imposed by the Philippines government in the March 2020. This analysis unveils the emotions of Filipino citizens especially residing in the Luzon area. A noteworthy finding is that most of the Filipinos had negative feelings during this time and mostly concerned about food shortage in the grim stretch. The study also proposes that the negative emotions like fear, anger and indignation will increase with time and government must take timely measures for their citizens. AYLIEN Sentiment Analysis API is used here for conducting sentiment analysis on the data extracted using RapidMiner [28].

Th exceptional research was done by Imran et. al. (2020) advocates that a deep learning-based sentiment analysis could prove fruitful for calculating a cross-cultural polarity and emotion detection. The paper discusses how people from different background and living in diverse countries react to coronavirus, even though their national boundaries are connected to each other. For instance, Danes and Swedes reacted contrarily to their government measures against COVID-19. Also, citizens of South Asian countries displayed more fear and anxiety than other people residing at other parts of the world. Deep long short-term memory (LSTM) models were employed on some established datasets (and also COVID-19 tweets) for tracing six primary emotions: joy, surprise, sad, disgust, fear and anger. This study is magnificent in terms of identifying cross-cultural trends despite the variety in socio, economic and cultural differences between people [29].

Aljameel et. al. (2021) on the other hand conducted extensive research on a specific nation – Saudi Arabia regarding coronavirus. The research aims to develop an efficient model that measures the awareness of commoners living in five distinctive regions of Saudi Arabia. This way the government can analyze the sentiments of public in a better way and can take better decisions for people living in separate regions of the country. Various machine learning models were applied along with n-gram feature extraction techniques on the Twitter data, and it was discovered that SVM in conjunction with TF-IDF offers nearly 85% accuracy. It was found out that the Southern part of the country has the most aware citizens while people living in the middle exhibited lowest alertness regarding the pandemic [30].

Situla et. al. (2021) also did country specific research on coronavirus related tweets i.e. Nepal. A novel dataset called NepCOV19Tweets, which collects the posts on Twitter platform in Nepali language, on which feature extraction multiple methods —fastText-based (ft), domain-specific (ds), and domain-agnostic (da) are applied. Then this dataset is submitted to three CNN models including one ensemble CNN

approach. Research shows that the newly formed dataset could act as a benchmark dataset for the future and the ensembled CNN approach yields better efficiency than the traditional machine learning models. The limitations of the study are also mentioned in the paper like ignoring the token sequence and established word embeddings like Word2vec and GloVe [31].

Another excellent work by Yin et. al. (2022) described that over 20 vaccines were developed for COVID-19 which were approved by the World Health Organization (WHO). This study covers the public sentiments from nearly 75,000 tweets towards the coronavirus vaccines especially in four countries: India, England, USA and Canada. The overall trend in these countries exhibit a positive response towards vaccination as people show their gratitude towards their governments for getting them vaccinated. However, there are some vaccines which are much disliked on a much larger scale like Sputnik V in USA and Sinovac in England and Canada. A higher percentage of positive posts show that people are more satisfied with government vaccination schemes and hopeful that it will definitely help to contain COVID-19 disease [32].

Singh et. al. (2022) have also developed an LSTM-RNN-based network and augmented featured weighting by an attention layer for opinion mining of the tweets based on the COVID-19 data. This data is taken from a public dataset available on the Kaggle site and it contains 1,79,108 posts that are analyzed and based on the core emotions divided into four classes - sad, joy, fear, and anger. The study displays an increase in approximately 12% precision and 20% accuracy that makes this proposed alternative more practical and effective than the customary standards. As this study is based on feature mapping and weight adjustment, there are some limitations also which will definitely affect the expected outcomes. It is also vital to understand that the recall parameter is not that much altered by the proposed algorithm [33].

A similar study was conducted by Shofiya & Abidi (2021) on coronavirus tweets, but this was done with the help of machine learning instead of deep learning. Here, a particular part of the COVID problem was explored i.e. social distancing. Hence, peculiar tweets regarding social distancing in the geographical domain of Canada was extrapolated using specific keywords and then examined using SentiStrength tool, followed by the application of Support Vector Machine (SVM) machine learning algorithm. The study showed that approximately 40% of tweets were neutral in nature and out of the remaining posts, maximum posts contained negative sentiments and very few posts had positive emotions in them. Although the given algorithm is quite simple in nature, still it enhances the accuracy up to 81% [34].

Elaborative research by Sanders et. al. (2021) to demonstrate public thoughts towards mask usage during the COVID era. For this task, a herculean dataset of 1 million tweets was created by downloading posts related to wearing masks and a topic-clustering approach was taken to analyze them. This research allowed speedy summarization of user opinions about prevention measures taken in this disastrous age and societal awareness of COVID-19 were extracted as a result. This research was designed to help the health care workers and policy makers to better understand the public perceptions by identifying them into sub-groups which need exceptional care during this time [35].

Kausar, Soosaimanickam & Nasar (2021) did extensive research by doing sentiment analysis across the globe by selecting some prime nations and analyzing posts and comments of their citizens. These countries are USA, Brazil,

India, Russia, South Africa, Peru, Mexico, Chile, Spain, UK and Oman. Specifying the above geographical locations, more than 13,05,000 tweets were extracted from the Twitter social media site using RTweet package and more than 16 keywords and then a substantial size and versatile dataset was formed. This study is bit different from others discussed so far as it took diverse tools for opinion mining like R programming language and packages like TM, tidytext, word cloud, dplyr, syuzhet, etc. Experimental results show more than 50% of the population in the selected countries displayed positive and hopeful emotions during the COVID era [36].

Chandra & Krishna (2021) undertook tweets related to only one country – India and studied the online behavior of Indians in this period. For this, a new dataset of 10,000 tweets from peak seasons of COVID was formed and then manually annotated for better training of the algorithm. The proposed system here is based on the concept of Long Short-Term Memory (LSTM) Recurrent Neural Networks, which is a special type of deep learning architecture. This LSTM model was combined with GloVe and state-of-art BERT model for deep understanding of human psychology during the catastrophic era. The experimental results reveal that joking, optimistic and infuriated posts dominated the Indian Twitter media during lockdown peaks, but still an exceedingly high number of netizens expressed optimism and extremely annoyed by the handling of pandemic by the government [37].

The research paper by Mujahid (2021) focused on the aspect of online education and e-learning during the covid time. The researchers underlined that this gloomy time of coronavirus was an unprecedented one as mundane man had not experienced such things in the modern times. As most of the educational institutions including schools, colleges and universities were closed during this period, providing education online was the online option left to the academicians and knowledge-seekers. Hence, this study took a dataset of tweets regarding the same and analyzed it using myriad machine learning and deep learning mechanisms. For checking the sentiment polarity and subjectivity, SentiWordNet, TextBlob and VADER were employed here. For feature extraction also, Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) were utilized, followed by the usage of various learning models. An interesting finding was that data balancing with SMOTE boosts the accuracy of classification. Also, the superiority of TextBlob was displayed by the experimental results as it performs better than VADER and SentiWordNet when it comes to data annotations. The researchers also solidify the fact that the deep learning models have more efficacy than the machine learning models [38].

During the year 2021, multiple researchers turned to the topic of covid vaccination as it was the hot potato of that period. Lyu, Han & Luli (2021) explain how the vaccination became the most debated subject on social media first and then on mainstream news when Russia announced the approval of first COVID-19 vaccine on 11th August 2020. The discussions rose and fall frequently; but the emotions moved its peak when Pfizer announced that its vaccine is 90% effective on 9th November 2020. The study showed that the trust factor was the most dominant one in the discussions of covid vaccines and the netizens displayed incredibly positive behavior towards the development and acceptance of coronavirus vaccines [39]. In the same time frame, Liu and Liu (2021) also studied the same subject by collecting English-language tweets related to covid vaccines. A humungous total of 26,78,372 vaccine-related tweets were stored in a dataset and with the help of VADER their sentiment polarity was discovered. One important aspect

here is that the researchers identified five themes each for positive (trial results, administration, life, information, and efficacy) and negative sentiments (trial results, conspiracy, trust, effectiveness, and administration). The experiments were conducted state-wise in the USA and also nation-wise. Th experimental results display that the sentiment score was highest in the United Arab Emirates (UAE) and lowest in Brazil [40].

Marcec & Likic (2021) have focused their research on some specific SARS-CoV-2 vaccines namely AstraZeneca/Oxford, Pfizer/BioNTech and Moderna for four months of peak discussion. Another notable point here is that AFINN lexicon was employed here to calculate the average sentiment longitude wise on a daily basis and compared. The study uses a total of 7,01,891 tweets for carrying out the research and results reveal that most of the vaccines were officially supported by the public online, however AstraZeneca/Oxford is the only vaccine about which the sentiment kept on decreasing and furthered the hesitancy towards this particular one while accepting other vaccines [41].

An analogous investigation was carried out by Nezhad & Deihimi (2021) which concentrated on social media posts in Iran region only. A total of 8,03,278 Persian tweets were extracted from Twitter which mentioned the following vaccines - Pfizer/BioNTech, AstraZeneca/Oxford, Moderna, Sinopharm and COVIran Barekat. A perceptible point is that the later vaccine was developed in Iran. A CNN-LSTM based deep learning architecture was used here for carrying out opinion mining of the Iranian people. Overall, it was discovered that the Iranian public displayed more positive attitude towards the imported vaccines than the home-grown vaccine. Hence the government and public health care services should turn to social networking sites to counter the negative opinions about COVIran Barekat and promote its wide usage across the nation [42].

Pristiyono et. al. (2021) also conducted a similar study on COVID-19 vaccines using the tweets extracted in the Bahasa Indonesia language using Twitter API and RapidMiner tool. Another tool that was used for data crawling was Drone Emprit Academic Streaming. After the data collection and pre-processing, data was subjected to Naïve Bayes machine learning algorithm. The results showed that around 56% people had a negative sentiment towards the covid vaccines and were reluctant to get vaccinated [43].

The scholarly article by Rahman & Islam (2021) highlights the role of social media in fighting against the coronavirus dilemma and fear. In this study, twelve thousand tweets from United Kingdom were examined and annotated by three different reviewers. After the rigorous data labeling, the data is exposed to three ensemble machine learning models to classify them into positive, neutral and negative posts. It was found out that the Stacking classifier (SC) achieved the highest f1 score i.e. 83.5%, followed by 83.3%o of the Voting classifier (VC) and 83.2% of the Bagging classifier (BC). The results were far superior to any other machine learning approach [44].

Basiri et. al (2021) did an exhaustive study by employing 16,00,000 tweets for understanding the human sentiments regarding SARS-COV-2 virus. This colossal data was collected from Twitter platform from people in eight different countries in duration of four months. This data was preprocessed and submitted to five basic deep learning models to carry out the opinion mining task. An ensemble approach by combining the five deep learning approaches is also operated on this data for proper analysis and it yielded the maximum accuracy. This

paper also consider "Google Trends" for better comprehension of the social sentiments of individuals living in different spheres of the world. The major findings of this study are as follows: first, every nation has its own unique sentiment pattern, second the negative sentiments rise exponentially after report of new death cases due to coronavirus [45].

The brainy researchers Chintalapudi, Battineni & Amenta (2021) turned towards Twitter social networking platform to collect posts across 5 months to study the opinion pattern of Indian netizens during the covid era. The data was classified into four categories: fear, anger, sad and joy using the standard machine learning classifiers - Logistic Regression (LR) and Support Vector Machines (SVM). The results were then compared with Long-Short Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT) model. It was found that the BERT model attained highest accuracy of 89% while all other remaining models produced much lesser accuracy [46].

Kaur et. al. (2021) explained how the bizarre situation arose for every country when coronavirus surfaced in 2019 and to tackle it in a constructive way, governments explored the social media platforms for better health care advice. The authors have designed a new algorithm called Heterogenous Support Vector Machines (H-SVM) that produces higher values of accuracy, precision and recall for the data collected from Twitter. The programming language used here is R and the results obtained after applied model is also compared to RNN and SVM [47].

Naseem et. al. (2021) developed a new large-scale sentiment dataset COVIDSENTI which encompasses 90,000 tweets collected during covid time. A diverse set of features and classifiers were operated on this dataset to condition the public sentiments. The experimental results showcase that commoners were quite positive during first stage of lockdown but the opinion momentum shifted towards negative value as the time passed. This shift in sentiment is uncertain but it may be due to distortion of information spread on social networking sites, hence the government authorities should also employ social media to create well-informed awareness among the netizens [48].

This alternative research by Dangi, Dixit & Bhagat (2022) developed a novel method called Sentimental Analysis of Twitter social media Data (SATD) to understand the feelings of common people towards COVID-19 disease. This innovative approach is based on five established machine learning standards- Logistic Regression, Random Forest Classifier, Multinomial NB Classifier, Support Vector Machine, and Decision Tree Classifier. The performance metrices obtained after applying new model predict better efficacy than the state-of-the-art approaches [49].

The research paper by Qorib (2023) evaluates the public sentiments towards the vaccines developed for the deadly SARS-COV-2 virus. The researchers have deployed three opinion computation techniques- VADER, TextBlob & Azure Machine Learning along with myriad vectorization methods- Doc2Vec, CountVectorizer, and TF-IDF. The entire process is subjected to multiple classifiers like LinearSVC, Naïve Bayes, Logistic Regression, Decision Trees and Random Forest for finding out the most efficient option available out of 42 algorithms applied on the twitter data. The experiments reveal that LinearSVC combined with TextBlob and TF-IDF produces maximum values for f1-score, precision, recall and accuracy performance metrices [50].

Braig et. al. (2023) did an extensive research by searching the vital five databases- IEEE Xplore, ScienceDirect, SpringerLink, ACM and AIS; and studying over 40 research papers on opinion mining on COVID-19 data. The chief conclusion that arrived after this exhaustive study is that the ensemble approaches are far superior in efficiency than the established machine learning classifiers. Another finding is that the pre-trained models (BERT and RoBERTa) fine-tuned on Twitter data provide even better results. This research contributed not only to the natural language processing field, but also to the social and behavioral sciences [51].
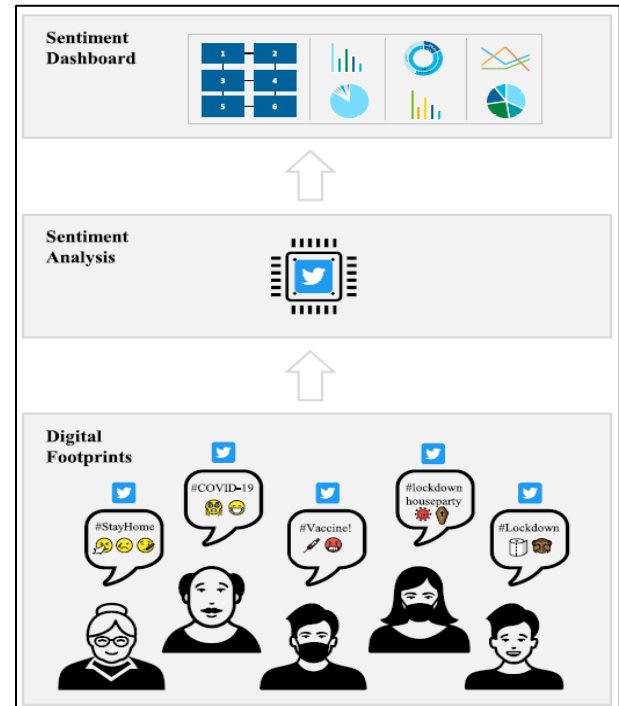

Fig 5. Role of Sentiment Analysis in gaining insights for effective management of COVID-19 [51]

An article by Joloudari et. al. (2023) highlights that posts written on social networking platforms are primary source for opinion mining in today's world. However, the social media posts are informal in nature which brings oodles of challenges like careless grammar and sentence formation, use of slangs, misspelled words, etc., which make the tasks of determining sarcasm, subjectivity detection and aspect extraction quite a daunting task. Hence, the authors proposed a contemporary method that utilizes BERT, CNN and Bi-LSTM for detecting human emotions present in the social media posts and comments [52].

A latest research by Storey & O'Leary (2024) revolves around the emotion evolution of communication on Twitter platform during pandemic related to coronavirus. Secondly, linguistic sentiment concepts—such as analytic thinking, authenticity, clout, and tone—were examined across various Twitter contexts, including variances in sentiment between tweets with and without pictures, as well as between original tweets and retweets. The researchers have also analyzed other forms of online activities like using Google Search and Wikipedia related to content present in the tweets. A stimulating finding here is that the initial tweets tended to be informational while the later tweets were more on the anger side showing frustration and rage. By examining all such detailed information, a dictionary was formed that is significant in investigating

influence of Twitter on the netizens in regard to COVID-19 [53].

Abiola (2023) published a brilliant article that did emotion mining of SARS-COV-2 related tweets in Nigeria using TextBlob and VADER sentiment analyzers. Over 10,48,575 tweets were extracted from Twitter for this purpose and submitted to Latent Dirichlet Allocation (LDA) to identify key topics discussed in the tweets, which were visualized using multidimensional scaling. The classification percentage of tweets in positive, neutral and negative were different for both VADER and TextBlob, although the chunks remained in same proportion. An imperative realizing is that people are divided into two halves where one section of people think COVID-19 will eventually be outdone and other section think that it is here to stay and will stay permanently in the world [54].

The scholarly article by Ainapure et. al. (2023) decodes the hidden sentiments of Indian people in the tweets related to the deadly virus and its vaccines. For tweet classification, this study has utilized two approaches: first, lexicon-based techniques- VADER and NRCLex and second deep learning methods- Bi-LSTM and GRU. Needless to say, the deep learning based approaches gave better accuracy than the lexicon-based methods. The study highlights the enhanced usage of social media during pandemic and study of netizens' posts for understanding situations in an effective manner and a better response [55].

In the research paper by Sandu et. al. (2023), the comprehensive bibliometric analysis of opinion mining research during the COVID-19 pandemic is carried out. The study deeply inspected 646 papers from the ISI Web of Science database that exhibited an exponential annual growth in this particular sphere by nearly 153%. King Saud University and Harvard Medical School are among the top contributing institutions; and the journals related to health sector published more articles on this topic than other scientific journals. Although, it is a global topic and scholars from around the world have contributed, but bulk of research is done by data scientists from USA and China. In short, this paper underlines impact of social networking sites in valuable understanding of the pandemic that is extremely helpful for medical and political persons [56].

Mir & Sevukan (2024) had also done an interesting study of sentiment analysis on COVID-19 vaccines in reference to social media posts of Indians. A special API known as Tweet Archiver was used for downloading of tweets having special hashtags- # Covid19vaccine and # Coronavirusvaccine for a period of two months. Post data collection and preprocessing, two special tools- Orange software and VOSviewer were employed for emotion analyzing. The study claims that most of the twitter users in India feel positively opinionated towards the vaccination programs of the Indian government. A major limitation of this research were: diminutive number of tweets undertaken for examination, restricting to posts and comments in English language only and that too for one nation only [57].

Another inventive article by Alqarni & Rahman (2023) underscores the importance of opinion mining in the Kingdom of Saudi Arabia before and after the ghastly pandemic. For this, two datasets were formed using Twitter data (from three main regions- Riyadh, Jeddah and Dammam): one that contains sentiments of people before covid and the other contains their emotions after facing pandemic. The major findings of this study are as follow: AraVec-SG-300-3 word2vec model gives quite good efficiency as compared to other available Arabic word embeddings. Also, Convolution Neural Network based

classifiers produce higher accuracy than the Bi-LSTM based algorithm. Obviously, commoners exhibited more negative sentiments during the covid era than pre-covid period [58].

Yet additional research by Thakur (2023) concentrated towards two specific diseases: Mpox and Coronavirus. The probe explored 61,862 tweets posted between May 2022 and March 2023, focusing on both viruses simultaneously that included top 50 hashtags. The paper also emphasizes that Twitter data is an indispensable tool for public discourse on health crises and must be taken seriously to identify and understand emotional status of mundane man. A major feature of this paper is comparison with 49 other vital research papers that upholds the novelty of the work carried out by author [59].

Catelli et. al. (2023) exploited the social networking comments and posts on Twitter for Italian netizens only for vaccination of SARS-COV-2 virus. For this purpose, 16,02,940 tweets were extracted and filtered. Then Lexicon-based opinion mining was carried out on this data and also the categorization of online users was done into Common users, Media, Medicine and Politics based on myriad lexicons. The sentiment lexicons were utilized to deduce the underlying emotions, extracting semantic orientation, tracing polarized and intensive Italian words of high importance. The primary discovery of this study is that most of the commoners had a strong negative sentiment which they expressed on Twitter platform [60].

An analogous research was carried out by Ariff et. al. (2024) which did sentiment analysis on the COVID-19 vaccines for netizens of Malaysia. The Machine Learning Life Cycle (MLLC) method was applied on the twitter data collected by StartBot tool. Subsequently, the data was analyzed by a web dashboard designed to visualize the results for various time periods, tweet traits and user demographics. The performance metrices of applied model showed nearly 94% f1-score and recall, 93% precision and 97.3% accuracy, which is far better than established standards. The study revealed that the Malaysian people discussed the Pfizer vaccine more than AstraZeneca and Sinovac [61].

Xiong et. al. (2024) presented the intricate feelings expressed in the dynamic landscape of coronavirus associated social networking dialogues using multiple deep learning models namely RNN, CNN, LSTM, GRU and BERT-RCNN. It was found out the BERT-RCNN model provides superior efficacy due to the features like bidirectional contextual insights obtained from BERT and enhanced feature mining delivered by RCNN. The paper also suggests that incorporating the Large Language Models (LLMs) can further increase the effectiveness of opinion mining [62].

The research by Park, Jang & Kwak (2024) undertakes the sentiment analysis of data related to later variants of coronavirus- Omicron substrain, BA4 and BA5. Interestingly, this data is collected from Reddit instead of Twitter, which is a more renowned alternative. Approximately, 3,38,465 COVID-19 vaccine-related comments were extracted from Reddit for a period of one year and five months. The analysis of this data is also compared with data related to other societal events and it was discovered that vaccine related discussions tend to spike during the numerous social and political programmes [63].

In a research article by Ahammad (2024), the data collected from various sources to comprehend how the fake news and misinformation related to COVID-19 is spread to strike fear in the common people. Also, a rule-based sentiment analysis was employed to label the dataset with three sentiment tags-positive, neutral and negative. Here, the idea of opinion mining

is combined with topic modeling where three vital models- Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA) and Matrix Factorization (NMF) are employed on the data composed from diverse news sources. The experimental results demonstrate that LDA is more efficient and versatile than other two by identifying more than 18 positive fake news topics and 20 negative topics and that too with highest coherence score [64].

The study done by Chen & Xu (2024) reconnoiters the emotions arising from Sina Weibo, a popular Chinese social media platform, during the duration of coronavirus. A total of 72,084 Weibo posts correlated to the 2022 Shanghai public health programme. An interesting finding is that people in China were more concerned about the impact of COVID-19 disease on their professional lives rather than life-threatening danger. Hence, the initial goals (death and illness) related to this deadly virus appear to have altered and the suggests that authorities should act accordingly. The results also disclose that Weibo platform reveals imperative user characteristics and these individual features have an enormous impact. For instance, the posts of middle-aged man were significantly different than those of elderly persons [65].

4. RESULTS AND DISCUSSION

SARS-COV-2 virus that originated from Wuhan, China in December 2019 caused a havoc among the people across the globe and it caused unprecedented situations like millions of deaths, community lockdown, wearing masks, social distancing, etc. The unusual circumstances forced the mundane people to stay inside their homes and spending time on social media for gaining information, learning new things and expressing their concern or fear became the new normal. This humongous amount of data generated by people of entire world made the data scientists curious and oodles of researches were conducted to examine the sentimental status of commoners which could hugely benefit the health care workers and policy makes in the government sector to understand needs of common people and help them in an effective manner.

The following table 2 summarizes the exceptional research work carried out by the leading academicians that contain opinion mining on specific geographical locations (chronological order):

Table 2. Sentiment Analysis on Specific Geographical Locations

Authors	Year	Geographical Location
Akash D Dubey	2020	Belgium, India, Australia, Italy Germany, France, Switzerland, Netherlands, USA, China, Spain & UK
Carlos de las Heras-Pedrosa, Pablo Sánchez-Núñez & José Ignacio Peláez	2020	Spain
Cherish Kay L. Pastor	2020	Philippines
Prasoon Gupta , Sanjay Kumar, R. R. Suman & Vinay Kumar	2021	India
Sumayh S. Aljameel et. al.	2021	Saudi Arabia
C. Sitaula, A. Basnet, A. Mainali & T. B. Shahi	2021	Nepal

Authors	Year	Geographical Location
Carol Shofiya & Samina Abidi	2021	Canada
Mohammad Abu Kausar, Arockiasamy Soosaimanickam & Mohammad Nasar	2021	USA, Brazil, India, Russia, South Africa, Peru, Mexico, Chile, Spain, UK & Oman.
Rohitash Chandra & Aswin Krishna	2021	India
Siru Liu & Jialin Liu	2021	USA, UAE & Brazil
Zahra Bokaee Nezhad & Mohammad Ali Deihimi	2021	Iran
Pristiyono et. al.	2021	Indonesia
Md Mahbubar Rahman & Muhammad Nazrul Islam	2021	UK
Nalini Chintalapudi, Gopi Battineni & Francesco Amenta	2021	India
Hui Yin, Xiangyu Song, Shuiqiao Yang & Jianxin Li	2022	UK, Usa, India & Canada
Odeyinka Abiola et. al.	2023	Nigeria
Andra Sandu et. al.	2023	Saudi Arabia
Stefanis Christos et. al.	2023	Greece
Arwa Alqarni & Atta Rahman	2023	Saudi Arabia
Rosario Catelli et. al.	2023	Italy
Mohamed Imran Mohamed Ariff et. al.	2024	Malaysia
Lixiong Chen & Nairui Xu	2024	China

The research papers discussed in this survey can be categorized into five different classes: Machine Learning (ML)-based, Deep Learning (DL)-based, Lexicon-based, Hybrid & Miscellaneous depending on the type of classifiers that are used by the researchers in their study. As evident from the nomenclature, ML-based category of papers consist of research papers solely focused on the different algorithms of machine learning approach. The DL-based grouping consists of all the research articles that undertook deep learning as well as Large Language Models (LLMs) like BERT, RoBERTa, etc. The Hybrid architecture is the most popular one as more researchers deploy more than one type of classifier to obtain higher values of performance metrices. All other research papers are categorized into Miscellaneous. The following image (Figure 6) shows the number of research papers belonging to above categories.
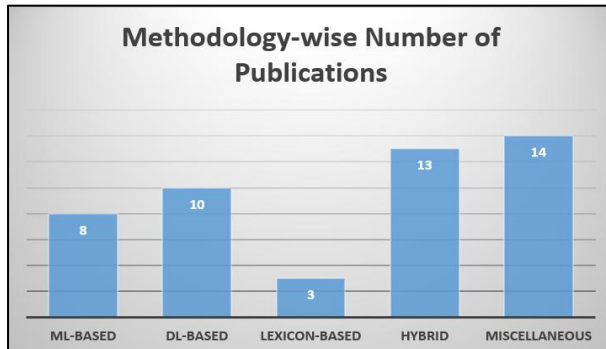
Fig 6. Number of Publications according to Chief Methodology used in Research

As evident from the figure above, more researchers tend to employ the deep learning methods rather than machine learning and lexicon-based techniques; however the hybrid approaches are gaining popularity as the ensemble approaches definitely provide better output and thus getting more utilized in the contemporary investigation of opinion mining in case of SARS-COV-2 virus.

Although depending upon geographical location and number of posts under review, the technique for sentiment analysis may vary, but generally it can be concluded that best performance results are obtained when Hybrid or Ensemble methodologies are applied with either TextBlob or VADER.

5. CONCLUSION

Data analysis plays a pivotal role in various research areas and that includes the deadly coronavirus also. COVID-19 spread like wild fire and soon took over the universe and was aptly declared a pandemic by WHO due to its colossal rate of spread infecting billions and killing millions. A lot of researchers have put brilliant work in doing research on the vital and vicious topic of COVID-19 sentiment analysis. This significant studies had made it easy for the people involved in health sector and governance to improved comprehension of needs and anxieties of mundane man. This understanding will lead to better governance steps and will benefit the entire population. In this study, a comprehensive survey of research papers for last five years has been conducted. This systematic review addressed the main highlights of each of the forty eight research articles considered like chief methodology employed (ML, DL, LLM, Lexicon or Hybrid), nation specific opinion mining, data volume, topic modeling and the decisive discussions by common people on social media regarding coronavirus. The key submission of this research that maximum people exhibited positive emotions in regards to COVID-19 prevention steps, vaccines, imposed lockdowns and the efficacy of national governments in fighting against the lethal coronavirus. Another crucial verdict is that hybrid approaches were more employed by data scientists as they offered superior productivity than either machine learning, deep learning or lexicon-based classifiers.

6. REFERENCES

[1] Zanke, A. A., Thenge, R. R., & Adhao, V. S. (2020). COVID-19: A pandemic declare by world health organization. IP International Journal of Comprehensive and Advanced Pharmacology, 5(2), 49-57.

[2] Chen, T. (2021). Fomites and the COVID-19 pandemic: An evidence review on its role in viral transmission. National Collaborating Centre for Environmental Health: Vancouver, BC, Canada, 1-24.

[3] Madabhavi, I., Sarkar, M., & Kadakol, N. (2020). COVID-19: a review. Monaldi Archives for Chest Disease, 90(2).

[4] Web Desk, "X-rays size up coronavirus protein structure at room temperature", Article Date: 30 Jun., 2020, Access Date: 31 Oct., 2024, An online article available on https://www.theweek.in/news/health/2020/06/30/xrays-size-up-coronavirus-structure-at-room-temperature.html.

[5] Shapiro, L. C., Thakkar, A., Campbell, S. T., Forest, S. K., Pradhan, K., Gonzalez-Lugo, J. D., ... & Halmos, B. (2022). Efficacy of booster doses in augmenting waning immune responses to COVID-19 vaccine in patients with cancer. Cancer Cell, 40(1), 3-5.

[6] Auxier, B., & Anderson, M. (2021). Social media use in 2021. Pew Research Center, 1(1), 1-4.

[7] Kang, H., Wang, Y., Wang, M., Al Imran Yasin, M., Osman, M. N., & Ang, L. H. (2023). Navigating digital network: Mindfulness as a shield against cyberbullying in the knowledge economy era. Journal of the Knowledge Economy, 1-39.

[8] Alam, S., & Yao, N. (2019). The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis. Computational and Mathematical Organization Theory, 25, 319-335.

[9] SA image - Devendra, "Sentiment Analysis with NLP & Deep Learning", Article Date: 24 Feb., 2022, Access Date: 13 Nov., 2024, An online article available on https://www.analyticsvidhya.com/blog/2022/02/sentiment-analysis-with-nlp-deep-learning/.

[10] Sadia, A., Khan, F., & Bashir, F. (2018, February). An overview of lexicon-based approach for sentiment analysis. In 2018 3rd International Electrical Engineering Conference (IEEC 2018) (pp. 1-6).

[11] Boppiniti, S. T. (2022). Exploring the Synergy of AI, ML, and Data Analytics in Enhancing Customer Experience and Personalization. International Machine learning journal and Computer Engineering, 5(5).

[12] Tsoy, D., Tirasawasdichai, T., & Kurpayanidi, K. I. (2021). Role of social media in shaping public risk perception during COVID-19 pandemic: A theoretical review. International Journal of Management Science and Business Administration, 7(2), 35-41.

[13] Ghosh, S. K. (2022). MASK AS A FASHION PRODUCT: A CULTURAL STUDY OF COVID-19 IN INDIA. Society, Pedagogy, Politics: A Multidimensional Approach to COVID-19, 130.

[14] Bing Liu, "Sentiment Analysis: mining sentiments, opinions, and emotions", 2nd edition, Cambridge University Press, 2020.

[15] Păvăloaia, V. D., Teodor, E. M., Fotache, D., & Danileţ, M. (2019). Opinion mining on social media data: sentiment analysis of user preferences. Sustainability, 11(16), 4459.

[16] Sánchez-Núñez, P., Cobo, M. J., De Las Heras-Pedrosa, C., Pelaez, J. I., & Herrera-Viedma, E. (2020). Opinion mining, sentiment analysis and emotion understanding in

advertising: a bibliometric analysis. IEEE Access, 8, 134563-134576.

[17] Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). Sentiment analysis and opinion mining on educational data: A survey. Natural Language Processing Journal, 2, 100003.

[18] Soong, H. C., Jalil, N. B. A., Ayyasamy, R. K., & Akbar, R. (2019, April). The essential of sentiment analysis and opinion mining in social media: Introduction and survey of the recent approaches and techniques. In 2019 IEEE 9th symposium on computer applications & industrial electronics (ISCAIE) (pp. 272-277). IEEE.

[19] Stefanis, C., Giorgi, E., Kalentzis, K., Tselemponis, A., Nena, E., Tsigalou, C., ... & Bezirtzoglou, E. (2023). Sentiment analysis of epidemiological surveillance reports on COVID-19 in Greece using machine learning models. Frontiers in Public Health, 11, 1191730.

[20] Iwendi, C., Mohan, S., Ibeke, E., Ahmadian, A., & Ciano, T. (2022). Covid-19 fake news sentiment analysis. Computers and electrical engineering, 101, 107967.

[21] Dubey, A. D. (2020). Twitter sentiment analysis during COVID-19 outbreak. Available at SSRN 3572023.

[22] Manguri, K. H., Ramadhan, R. N., & Amin, P. R. M. (2020). Twitter sentiment analysis on worldwide COVID-19 outbreaks. Kurdistan Journal of Applied Research, 54-65.

[23] Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. Journal of Information and Telecommunication, 5(1), 1-15.

[24] Gupta, P., Kumar, S., Suman, R. R., & Kumar, V. (2020). Sentiment analysis of lockdown in india during covid-19: A case study on twitter. IEEE Transactions on Computational Social Systems, 8(4), 992-1002.

[25] de Las Heras-Pedrosa, C., Sánchez-Núñez, P., & Peláez, J. I. (2020). Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. International journal of environmental research and public health, 17(15), 5542.

[26] Boon-Itt, S., & Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR public health and surveillance, 6(4), e21978.

[27] Al-Shabi, M. (2020). Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining. IJCSNS, 20(1), 1.

[28] Pastor, C. K. Sentiment analysis of Filipinos and effects of extreme community quarantine due to coronavirus (COVID-19) pandemic. 2020. Available at SSRN.

[29] Imran, A. S., Daudpota, S. M., Kastrati, Z., & Batra, R. (2020). Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. Ieee Access, 8, 181074-181090.

[30] Aljameel, S. S., Alabbad, D. A., Alzahrani, N. A., Alqarni, S. M., Alamoudi, F. A., Babili, L. M., ... & Alshamrani, F. M. (2021). A sentiment analysis approach to predict an individual's awareness of the precautionary procedures to prevent COVID-19 outbreaks in Saudi Arabia. International journal of environmental research and public health, 18(1), 218.

[31] Sitaula, C., Basnet, A., Mainali, A., & Shahi, T. B. (2021). Deep Learning-Based Methods for Sentiment Analysis on Nepali COVID-19-Related Tweets. Computational Intelligence and Neuroscience, 2021(1), 2158184.

[32] Yin, H., Song, X., Yang, S., & Li, J. (2022). Sentiment analysis and topic modeling for COVID-19 vaccine discussions. World Wide Web, 25(3), 1067-1083.

[33] Singh, C., Imam, T., Wibowo, S., & Grandhi, S. (2022). A deep learning approach for sentiment analysis of COVID-19 reviews. Applied Sciences, 12(8), 3709.

[34] Shofiya, C., & Abidi, S. (2021). Sentiment analysis on COVID-19-related social distancing in Canada using Twitter data. International Journal of Environmental Research and Public Health, 18(11), 5993.

[35] Sanders, A. C., White, R. C., Severson, L. S., Ma, R., McQueen, R., Paulo, H. C. A., ... & Bennett, K. P. (2021). Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse. AMIA Summits on Translational Science Proceedings, 2021, 555.

[36] Kausar, M. A., Soosaimanickam, A., & Nasar, M. (2021). Public sentiment analysis on Twitter data during COVID-19 outbreak. International Journal of Advanced Computer Science and Applications, 12(2).

[37] Chandra, R., & Krishna, A. (2021). COVID-19 sentiment analysis via deep learning during the rise of novel cases. PloS one, 16(8), e0255615.

[38] Mujahid, M., Lee, E., Rustam, F., Washington, P. B., Ullah, S., Reshi, A. A., & Ashraf, I. (2021). Sentiment analysis and topic modeling on tweets about online education during COVID-19. Applied Sciences, 11(18), 8438.

[39] Lyu, J. C., Han, E. L., & Luli, G. K. (2021). COVID-19 vaccine–related discussion on Twitter: topic modeling and sentiment analysis. Journal of medical Internet research, 23(6), e24435.

[40] Liu, S., & Liu, J. (2021). Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis. Vaccine, 39(39), 5499-5505.

[41] Marcec, R., & Likic, R. (2022). Using twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgraduate medical journal, 98(1161), 544-550.

[42] Nezhad, Z. B., & Deihimi, M. A. (2022). Twitter sentiment analysis from Iran about COVID 19 vaccine. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 16(1), 102367.

[43] Pristiyono, Ritonga, M., Ihsan, M. A. A., Anjar, A., & Rambe, F. H. (2021, February). Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm. In IOP Conference Series: Materials Science and Engineering (Vol. 1088, No. 1, p. 012045). IOP Publishing.

[44] Rahman, M. M., & Islam, M. N. (2022). Exploring the performance of ensemble machine learning classifiers for sentiment analysis of COVID-19 tweets. In Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021 (pp. 383-396). Springer Singapore.

[45] Basiri, M. E., Nemati, S., Abdar, M., Asadi, S., & Acharrya, U. R. (2021). A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. Knowledge-Based Systems, 228, 107242.

[46] Chintalapudi, N., Battineni, G., & Amenta, F. (2021). Sentimental analysis of COVID-19 tweets using deep learning models. Infectious disease reports, 13(2), 329-339.

[47] Kaur, H., Ahsaan, S. U., Alankar, B., & Chang, V. (2021). A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. Information Systems Frontiers, 23(6), 1417-1429.

[48] Naseem, U., Razzak, I., Khushi, M., Eklund, P. W., & Kim, J. (2021). COVIDSenti: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis. IEEE transactions on computational social systems, 8(4), 1003-1015.

[49] Dangi, D., Dixit, D. K., & Bhagat, A. (2022). Sentiment analysis of COVID-19 social media data through machine learning. Multimedia tools and applications, 81(29), 42261-42283.

[50] Qorib, M., Oladunni, T., Denis, M., Ososanya, E., & Cotae, P. (2023). Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset. Expert Systems with Applications, 212, 118715.

[51] Braig, N., Benz, A., Voth, S., Breitenbach, J., & Buettner, R. (2023). Machine learning techniques for sentiment analysis of COVID-19-related twitter data. IEEE Access, 11, 14778-14803.

[52] Joloudari, J. H., Hussain, S., Nematollahi, M. A., Bagheri, R., Fazl, F., Alizadehsani, R., ... & Talukder, A. (2023). BERT-deep CNN: State of the art for sentiment analysis of COVID-19 tweets. Social Network Analysis and Mining, 13(1), 99.

[53] Storey, V. C., & O'Leary, D. E. (2024). Text analysis of evolving emotions and sentiments in COVID-19 Twitter communication. Cognitive Computation, 16(4), 1834-1857.

[54] Abiola, O., Abayomi-Alli, A., Tale, O. A., Misra, S., & Abayomi-Alli, O. (2023). Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. Journal of Electrical Systems and Information Technology, 10(1), 5.

[55] Ainapure, B. S., Pise, R. N., Reddy, P., Appasani, B., Srinivasulu, A., Khan, M. S., & Bizon, N. (2023). Sentiment analysis of COVID-19 tweets using deep learning and lexicon-based approaches. Sustainability, 15(3), 2573.

[56] Sandu, A., Cotfas, L. A., Delcea, C., Crăciun, L., & Molănescu, A. G. (2023). Sentiment Analysis in the Age of COVID-19: A Bibliometric Perspective. Information, 14(12), 659.

[57] J Mir, A. A., & Sevukan, R. (2024). Sentiment analysis of Indian Tweets about Covid-19 vaccines. Journal of Information Science, 50(5), 1308-1320.

[58] Alqarni, A., & Rahman, A. (2023). Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: a deep learning approach. Big Data and Cognitive Computing, 7(1), 16.

[59] Thakur, N. (2023). Sentiment analysis and text analysis of the public discourse on Twitter about COVID-19 and MPox. Big Data and Cognitive Computing, 7(2), 116.

[60] Catelli, R., Pelosi, S., Comito, C., Pizzuti, C., & Esposito, M. (2023). Lexicon-based sentiment analysis to detect opinions and attitude towards COVID-19 vaccines on Twitter in Italy. Computers in Biology and Medicine, 158, 106876.

[61] Ariff, M. I. M., Zubir, N. E. S., Azizan, A., Ahmad, S., & Arshad, N. I. (2024). Malaysian views on COVID-19 vaccination program: a sentiment analysis study using Twitter. Bulletin of Electrical Engineering and Informatics, 13(1), 436-443.

[62] Xiong, J., Feng, M., Wang, X., Jiang, C., Zhang, N., & Zhao, Z. (2024). Decoding sentiments: Enhancing covid-19 tweet analysis through bert-rcnn fusion. Journal of Theory and Practice of Engineering Science, 4(01), 86-93.

[63] Park, B., Jang, I. S., & Kwak, D. (2024). Sentiment analysis of the COVID-19 vaccine perception. Health Informatics Journal, 30(1), 14604582241236131.

[64] Ahammad, T. (2024). Identifying hidden patterns of fake COVID-19 news: An in-depth sentiment analysis and topic modeling approach. Natural Language Processing Journal, 6, 100053.

[65] Chen, L., & Xu, N. (2024). To live or to stay alive? A thematic and sentiment analysis of public posts on social media during the 2022 Shanghai COVID-19 outbreak. Digital Health, 20552076241288731.