ORIGINAL RESEARCH

Sentiment analysis tracking of COVID-19 vaccine through tweets

Akila Sarirete[1,2]

Abstract

Recent studies on the COVID-19 pandemic indicated an increase in the level of anxiety, stress, and depression among people of all ages. The World Health Organization (WHO) recently warned that even with the approval of vaccines by the Food and Drug Administration (FDA), population immunity is highly unlikely to be achieved this year. This paper aims to analyze people's sentiments during the pandemic by combining sentiment analysis and natural language processing algorithms to classify texts and extract the polarity, emotion, or consensus on COVID-19 vaccines based on tweets. The method used is based on the collection of tweets under the hashtag #COVIDVaccine while the nltk toolkit parses the texts, and the tf-idf algorithm generates the keywords. Both n-gram keywords and hashtags mentioned in the tweets are collected and counted. The results indicate that the sentiments are divided into positive and negative emotions, with the negative ones dominating.

1 Introduction

On February 11, 2020, the World Health Organization (WHO) officially announced the name of the 2019 new coronaviruses outbreak "COVID-19", an abbreviation of "COrona VIrus Disease" (Lovelace 2020). Following dreadful disease spread levels, on March 11, 2020, WHO officially declared the COVID-19 a Pandemic that affected more than 110 countries (Beaubien 2020). Today, after a year of combating this pandemic, WHO reported that more than two million people have died and more than one hundred million infection cases, as shown in Fig. 1 (WHO 2021). This situation created a severe challenge for all countries to monitor and slow down the pandemic using various measures, including social distancing, lockdowns, limiting movement of persons, and avoiding the 3Cs (closed spaces, crowded, or involve close contact). After some relief during the summer period and returning to school and universities, a new vague appeared in fall 2020. The highly contagious mutation of the novel coronavirus variants was reported in the UK, South

Africa, and Brazil (Bollinger and Ray 2020). It added new fears to the international health organization. Figure 1 shows the distribution of the COVID-19 total cases worldwide and daily deaths from January 21, 2020, to December 31, 2020.

As reported by many health organizations, the top priority of 2021 is to continue fighting COVID-19, repairing and strengthening the existing health systems, speeding up access to COVID-19 treatments, and achieving equitable and safe vaccines for all (WHO 2021). Several developed vaccinations are currently being approved in several countries (Hindustan Times 2020). The majority of these vaccines were produced in North America (46%), China (18%), Asia (excluding China) and Australia (18%), and Europe (18%) (Le et al. 2020). However, traditionally the development of a vaccination pipeline could take on average ten years. It is not possible with the speed of the COVID-19 pandemic.

Consequently, new strategies have been adopted to accelerate the development of vaccines (Lurie et al. 2020). The WHO reported more than sixty COVID-19 vaccine candidates are currently in clinical development and more than 170 in preclinical development, such as Pfizer–BioNTech, Moderna, Sputnik V, Sinovac, and Oxford–AstraZeneca (Statista 2021). However, the WHO recently warned that despite vaccines approval by the Food and Drug Administration (FDA), such as Pfizer-BioNTech and Moderna's COVID-19 vaccines, population immunity is highly unlikely achieved this year (Cheng and Keaten 2021). The WHO

✉ Akila Sarirete
asarirete@effatuniversity.edu.sa

[1] Computer Science Department, Effat College of Engineering, Effat University, Jeddah, Saudi Arabia

[2] Energy and Technology Research Center, Effat University, Jeddah, Saudi Arabia
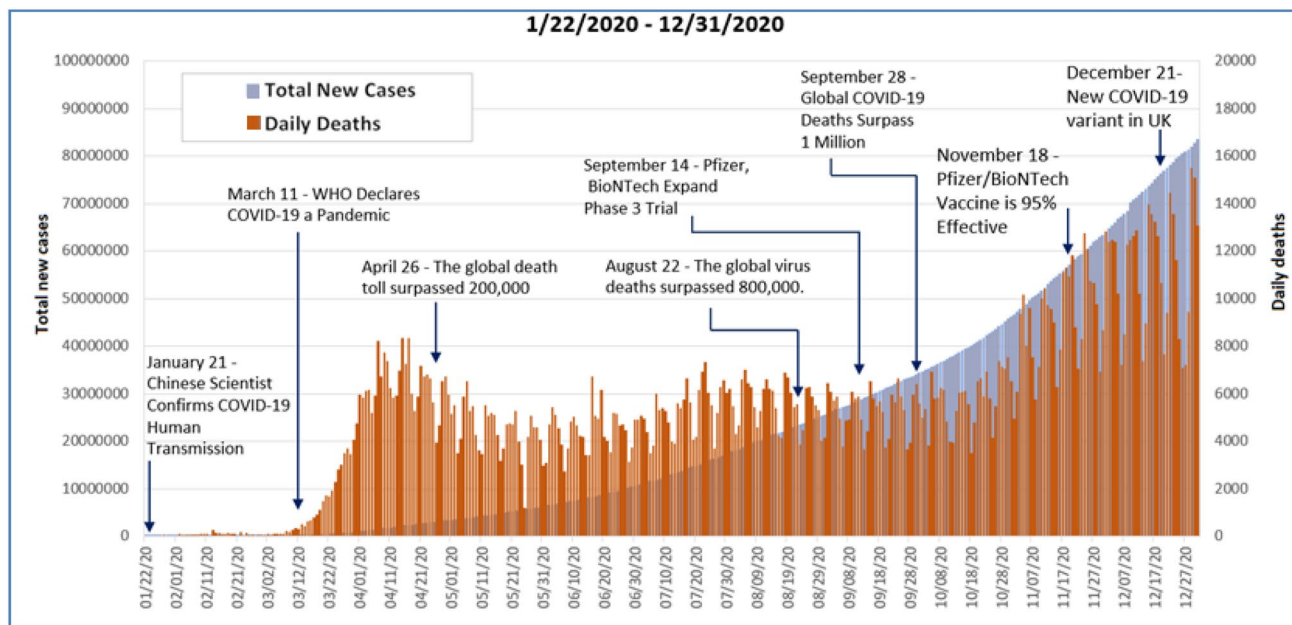
Fig. 1 COVID-19 worldwide total cases and daily deaths from 1/22/2020 to 12/31/2020

Table 1 List of available COVID-19 vaccines (WHO 2020)

COVID-19 vaccine	Country of Origin	Method	Estimated efficacy (%)
Pfizer/BioNTech	USA & Germany	Uses Message RNA (mRNA) in accordance with SARSCov2 virus responsible for COVID-19	95.03
Moderna/NIH	USA	Uses Message RNA (mRNA)	94.08
AstraZeneca/Oxford	UK and Sweden	Uses genetically altered virus	66.84
Sputnik V	Russia	Uses adenoviral vectors, viruses responsible for human cold	90.97
Sinopharm	China	Uses inoculation technique	79
Sinovac	China	Uses inoculation technique	78
Johnson & Johnson	USA	Uses genetically altered virus	66.62

concern is the unequal COVID-19 vaccine distribution, the rich and poor vaccine divide, and vaccine nationalism. During the recent WorldEconomic Forum, China President (World Economic Forum 2021) stressed the importance of cooperation to develop and produce vaccines for all people worldwide. Table 1 shows the available COVID-19 vaccines within the WHO emergency use (Olliaro et al. 2021), including the Chinese vaccines Sinopharm and Sinovac recently approved by the WHO (2020). While these vaccines use different methods, the efficacy (Relative Risk Reduction) shows Pfizer-BioNTech and Moderna among the highest).

The present work aims to analyze people's sentiments during the pandemic by combining sentiment analysis and natural language processing algorithms. The remainder of the paper is organized as follows: first, a review of the relevant background and literature of sentiment analysis is

presented, followed by the research methodology applied in this study. Then, results are described with a discussion, and finally, the study's main conclusion.

2 Background

Sentiment analysis has been applied in different domains such as social media monitoring, business, product analysis, stock market, tourism, health, and education to understand events and trends (Vijaykumar et al. 2017), (Ainin et al. 2020), (Hassan et al. 2020), (Abualigah et al. 2020), (Yadav et al. 2018). With the increase of social media, sentiment analysis presents the best way in understanding people opinions, whether it is positive, negative, or neutral. In healthcare, for example, sentiment analysis classifies patient

comments and quantifies the performance of the employees. This sentiment analysis extends to medical sentiment analysis to provide professionals with an automated support system, determine patients 'concerns, or analyze patients' emotions (Kushwah et al. 2021). Birjali et al. (2021) performed a comprehensive survey on sentiment analysis, which discussed traditional and recent approaches, challenges, and future trends. Srivastava et al. (2022) explored the challenges of sarcasm detection, negation detection, and word multipolarity and ambiguity. Most literature divides sentiment analysis into three categories: machine learning approach, lexicon-based approaches, and hybrid approaches (Duong and Nguyen-Thi 2021; Soong et al. 2021). One of the main problems in sentiment analysis is the classification of the sentiment polarity as positive, negative, or neutral. Several research articles indicated that many outbreaks and pandemics could have been quickly monitored if experts considered social media data (Alamoodi et al. 2021). According to Statista (Statista 2021), as of February 12, 2021, nearly three billion AstraZeneca/Oxford's vaccines, one billion Pfizer-BioNTech vaccines, and more than six hundred million Moderna vaccine were pre-purchased. However, in a survey conducted on December 2020 by KFF (Kaiser Family Foundation) under the research project "COVID-19 Vaccine Monitor" (Muñana 2020), 71% of respondents said they would agree to take the COVID-19 vaccine if it is free of charge and safe. While about 27% remained hesitant fearing side effects. In another survey, based Health Belief Model (HBM), conducted in Hong Kong (Wong et al. 2021) during the peak of the third wave of the pandemic between July 27 and August 27, 2020, the acceptance rate was 37.2%, while the confidence of newer vaccines was 43.4% and manufacturer received 52.2% of the respondents. In a study by Raamkumar, Tan, and Wee (Raamkumar et al. 2020), the authors used Facebook to investigate and understand the public sentiment and responses to the COVID-19 pandemic. The study aimed to improve the communication strategies conducted by various public health authorities (PHAs) in the United States, Singapore, and England. This study showed that social media analyses could provide insights into PHAs' communication strategies during a pandemic. In another study, Raamkumar, Tan, and Wee (Sesagiri Raamkumar et al. 2020) developed a deep-learning text classifier to investigate public perception and reaction to physical distancing. The authors conclude that public health authorities (PHAs) can characterize the public's health behaviors using classification models. Vijaykumar, Meurzec, Jayasundar (Vijaykumar et al. 2017) studied Zika outbreaks in Singapore using Facebook to bridge the gaps between people and public officials. Results show the importance of social media in sharing information on pandemics. Manguri, Rebaz, and Pshko (Manguri et al. 2020) analyzed Twitter sentiment on COVID-19 outbreaks by pulling Twitter data from Twitter

social media from April 9, 2020, to April 15, 2020, using tweepy library. The study found that people's sentiments and reactions vary daily, while the neutral toll was substantially significant for coronavirus and COVID-19 keywords. Prabhakar and Krishna (2020) used the Latent Dirichlet Allocation (LDA) to study the flow on Twitter during the COVID-19 pandemic. The sentiment analysis results confirmed people's expected reaction with negative sentiment towards the COVID-19 pandemic. The authors conclude that the knowledge flow on Twitter about the Coronavirus outbreak was necessary and correct. Some minor misinformation was spread compared to previous Ebola and Zika virus outbreaks, where Twitter users widely disseminated misinformation. Jia Xue et al. (2020) also used LDA machine learning algorithm to conduct COVID-19 sentiment analysis on 4 million Twitter messages. Results showed that public tweets have significant fear in their discussion. Shamrat et al. (2021) used Twitter to analyze people's sentiment towards three vaccines: Pfizer, Moderna, and AstraZeneca. The authors reprocessed the raw tweets using Natural Language Processing (NLP), while the algorithm for KNN classification was used to classify the processed data. The authors found a higher positive sentiment towards all three vaccines, with 47.29% for Pfizer, 46.16% for Moderna, and 40.08% for AstraZeneca Shamrat et al. (Shamrat et al. 2021). Another sentiment analysis study involving 3242 tweets on Pfizer and Sinovac vaccines was conducted during October–November 2020 in Indonesia (Nurdeni et al. 2021). Results revealed that Pfizer had 81% positive perceptions while Sinovac showed 77% positive perceptions.

3 Research methodology

Sentiment analysis (SA) is a text classification technique used to analyze natural language text. It uses Natural Language Processing (NLP) to determine whether the sentiments expressed towards a subject are positive, negative, or neutral (Solangi et al. 2018; Shamrat et al. 2021). It also helps decide people's emotions such as happiness, depression, anxiety, fear, and sadness. As shown in Fig. 2, there are three different approaches used in sentiment analysis (Kausar et al. 2019), (Rokade and Aruna 2019), (Hauthal et al. 2020): (1) a machine learning approach that includes linear classifiers (support vector machines and neural networks), probabilistic classifiers (Naïve Bayes and Bayesian network), and decision tree classifiers; (2) a lexicon-based approach which includes a corpus-based approach (statistical and semantic) and a dictionary-based approach; finally, (3) a hybrid approach which is a combination of the previous two approaches. The machine learning approach has better accuracy than the lexicon-based approach, primarily when it uses an extensive database. In contrast to machine
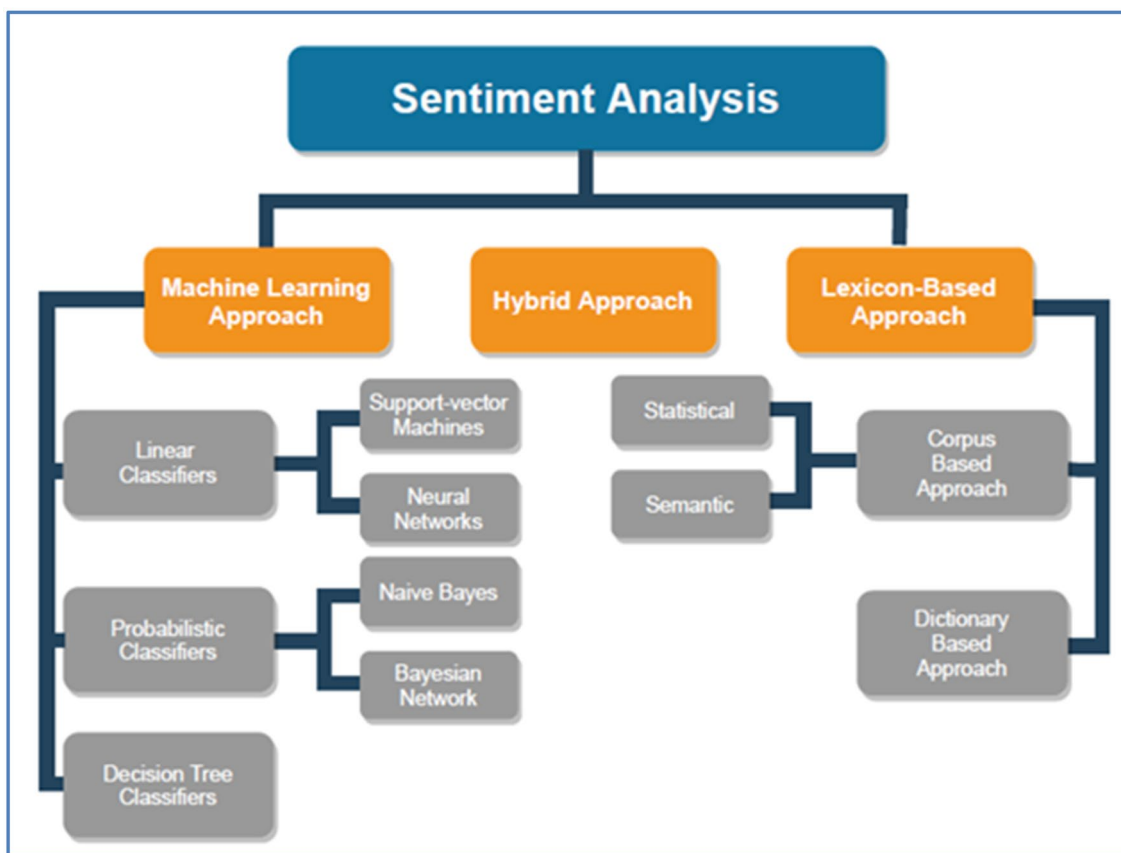
Fig. 2 Sentiment analysis approaches

learning, the lexicon-based method requires the involvement of humans to process text analysis. Four main steps are needed in sentiment analysis (i) a social media platform, (ii) a data collection, (iii) a pre-processing, and (iv) a data analysis. In this study, a hybrid approach is used, including a corpus-based method with semantic analysis. The sentiment analysis process begins with data collection and identification, then an extraction and classification of the features. Finally, a decision process is conducted during the stage of sentiment polarity and subjectivity. Data are collected in the form of tweets through social media during the first period between December 16, 2020 to January 27, 2021, and the second period from January 20, 2021 to April 8, 2021. The collected tweets are then pre-processed, and the cleaned data classified, analyzed and evaluated.

3.1 Data collection and pre-processing

The pre-processing of the data collected is mainly used to clean the raw data and minimize the vocabulary of words detected in the text message using Natural Language Processing with Python's NLTK Package (Solangi et al. 2018; Rajput 2019). Since the collected data is text, it needs to

be transformed into numerical representation using a fixed-length feature vector. This representation is commonly based on the bag-of-words approach (BOW) (Minaee et al. 2021). This approach is a simple and flexible way to extract features from documents and track the number of used words. In this work, the #COVIDvaccine hashtag is used through two primary sources:

- Online tweets' monitoring tool to monitor daily tweets on the hashtag (TAGS Google sheet with Macros (https://tags.hawksey.info/).
- Tweets collection and analysis using Python code monitoring the #COVIDvaccine hashtag over one month (30 days) using Twitter API and python Natural Language Processing (NLP) libraries.

This hashtag's total number of tweets is 230,672, with 230,623 unique tweets (till April 8, 2021). The author selected the subset of the collected tweets (16/12/2020 to 26/1/2021) and parsed the texts using the Natural Language Toolkit in Python. Then, the stopwords were eliminated from the collected corpus (such as articles and common words). Afterward, the author used the tf-idf algorithm to generate

keywords (Schvaneveldt et al. 1976). The n-gram model was used afterwards to identify the top words 1-gram (1-g) ($n = 1$; single words); bi-gram ($n = 2$; two adjacent words); and three-gram ($n = 3$; three adjacent words)).

3.2 Data analysis and evaluation

After cleaning the data and removing the stop words, the author vectorized the tweets using the sklearn library and built the corpus of words from the tweets. A corpus analysis was established and common words and emoji related to sentiments were identified such as {'happi'; 'nice'; 'good'; 'bad'; 'sad'; 'mad'; 'best'; 'pretti'; '♡'; ':)'; ':('; '😣'; '😩'; '😊'; '😎'; '😊';'😊';'😠';'😣';'😊';'😋'; critical; decline; die; disaster; collapse; lockdown; angry; risk; sad; serious; homeless; scam; reject; efficient; vaccine}. Tweets are divided into two groups: positive and negative tweets. After a comparison of the tweets with the keywords, positive and negative tweets are classified.

4 Results and discussion

In this section, the author investigates, in detail, the natural language processing of the tweets related to the main hashtag #COVIDvaccine.

4.1 Use of the TAGS tool

The author used the TAGS monitoring tool, freely offered by Martin Hawksey, to collect and monitor the tweets on the #COVIDvaccine hashtag. This tool helped in the social media analytics of the tweets and identified the trends in using a specific hashtag.

As shown in Fig. 3, more than 230,000 tweets were collected. The tool monitors the tweets daily. As shown, the top tweeter is *VaccineCa*. On the right side of the figure, the Twitter activity for the last three days is shown.

4.2 Data pre-processing and n-gram analysis

Using the TAGS tool helped figure out the tweets' activity over a period and show the tweets' trends analytically. Later, using Python code, the author analyzed the selected tweets deeper (based on collected data from December 16, 2020, till January 26, 2021) and applied tf-idf analysis on the tweets. The results include the 1-g, bi-gram, tri-gram keywords, and the collected hashtags.

Table 2 lists the top twenty 1-g using the tf-idf analysis. The words are organized by type, availability, and feelings. As shown, most of the top 1-g concerning the type is related to allergy, care, effectiveness, safety, reactions, isolation.

Table 2 tf-idf analysis—1-g model

Type	Availability	Feeling
Allergic	Appointment	Bragging
Care	Available	Convinced
Cause	Beginning	Grateful
Effective	Offered	Happy
Isolated	Remote	Hesitant
Reactions	Trial	Honest
Safe		Glad

Fig. 3 TAGS Archive for #COVIDvaccine hashtag

Table 3 Selected tweets related to the 1-gram analysis

Tweets
Toronto-based Providence Therapeutics begins a Phase I clinical trial for a Canadian-made #COVID19 vaccine.… https://t.co/3NlaPVoT2z
When $NVAX reports South Africa #covidvaccine data, they should also show first preclinical data w/ their SA-strain… https://t.co/kEt5Hqetht
"RT @bunsenbernerbmd: Hey while the allergic reactions to the #CovidVaccine are real, they are very rare With millions of jabs so far, per…."
While I'm glad we will have more doses of #covidvaccine available for Ohioans, I'm sad that the reason is because m… https://t.co/dyJG4LncxR
RT @IslamRizza: How many will it take to convince you? You are involved in the #Covidexperiment 🙂 —#Covid_19 #CovidVaccine https://t.co/TX…

Also, other words that come in the list are associated with the availability of the vaccine, as shown in the table. Other words in the 1-g analysis are related to bragging, convinced, grateful, happy, hesitant, honest. This shows the research is already finding some sentiments related to the vaccine.

Table 3 lists some top tweets that appear in the data analysis. As shown, some of the feelings are carried in the tweets.

Moreover, besides the 1-g analysis, the 2-g and 3-g models are also considered. Table 4 and Table 5 list the most common 2-g and 3-g, respectively, based on the availability and the feelings. As shown, several 2-g appear to tackle the safety of the vaccine and awareness of the people and the hesitancy towards the vaccine. Tweets show that people have mixed feelings about the vaccine. The bi-gram 'take covidvaccine' often comes in the tweets as the bi-gram 'covidvaccine bragging', considering that people are scamming about the vaccine.

The 3-g analysis shows a close relationship with the 1-g and 2-g analysis. As shown in Table 5, similar results appear related to the feelings. Some of the keywords are introduced, such as '2nd nation number' and 'community take covidvaccine', which are significantly related to the importance of the vaccine and the people taking the vaccine. For the feelings, new phrases appear, such as 'excited dose wrap' and 'reactions covidvaccine real' and 'honest discussion covidvaccine'.

Table 4 tf-idf analysis—2-g model

Availability	Feeling	
Currently offered	Covidvaccine bragging	Tell truth
Number people	Take covidvaccine	Aware scams
Offered people	Covidvaccine safe	Please aware
People cigarettes	Safe remember	Vaccine need
Single dose	Companies lied	Need apply
Trial	Apply vaccine	
First dose	Even hesitant	
Second dose		

Table 5 tf-idf analysis—3-g model

Availability	Feeling
Community take covidvaccine	Reactions covidvaccine real
Covidvaccine currently offered	Scams circulating around
Currently offered people	Covid_19 vaccine need
Dose covid vaccine	Covidvaccine bragging abou
2nd nation number	Covidvaccine safe remember
Around covid_19 vaccine	Even hesitant beginning
	Excited dose wrap
	Government convinced people
	Honest discussion covidvaccine

Table 6 lists extracts of tweets related to the 3-g analysis above. As shown, some of the sentiments are not related directly to the vaccine itself; for example, the 3-g 'government convinced people' is related to some comparison of the vaccine and the cigarettes. Another tweet is associated with the COVID-19 vaccine experiment and the excitement of taking it. Therefore, some mixed feelings are being retrieved when analyzing the tweets.

4.3 Data classification and sentiment analysis

The collected tweets are vectorized using the sklearn library, and the corpus of words is built. As stated in Sect. 3, the corpus analysis identified specific keywords and emojis related to sentiments. Therefore, the tweets are divided into two groups: positive and negative tweets. The tweets analysis went through several steps, as shown below:

1. Identification of most common keywords
2. Division of the tweets into positive and negative tweets
3. Stemming and cleaning the tweets
4. Calculating the frequency of the keywords in the tweets

Table 7 summarizes three tests done on the tweets after stemming and removing the stop words.

**Table 6** Selected tweets related to the 3-g analysis

Tweets
RT @CityWestminster: Let's have an honest discussion about the #CovidVaccine
Join Lord Simon Woolley, @ProfKevinFenton and @leader_wcc f…
RT @IslamRizza: People say: "Take the #CovidVaccine it's safe." Remember when the government convinced people that cigarettes DIDN'T cause…
RT @NRochesterMD: I am so excited! Dose #1 is a wrap! Guess what? Even I was hesitant in the very beginning. I had to read the data and get…
RT @IslamRizza: How many will it take to convince you? You are involved in the #Covidexperiment 😐 —#Covid_19 #CovidVaccine https://t.co/TX…
"Today, #ygk is on fire with #healthcare #innovation
Excited for this @MESHScheduling announcement to help with… https://t.co/g9Cqgka9vi"
"RT @WYRForum: ⚠️ Please be aware of scams circulating around the #Covid_19 #vaccine ⚠️
➡️ You don't need to apply for the vaccine—you'll be…"

Table 7 Tweets analysis and keywords classification

Tests	Total frequent words in tweets	Most common keywords
Fist test December 21, 2020 Tweets	746 words	Good; safe; die; vacc lockdown; risk; scam
Second test Adding January 25, 2021 Tweets	1197 words (more tweets added to the set)	Fight; good; safe; help; spread; scam; die; close; lockdown; risk
Third test Adding January 26, 2021 Tweets	1532 words (more tweets added to the set)	Happi; good; close; die; lockdown; risk; spread; scam; safe; fight; help; vacc
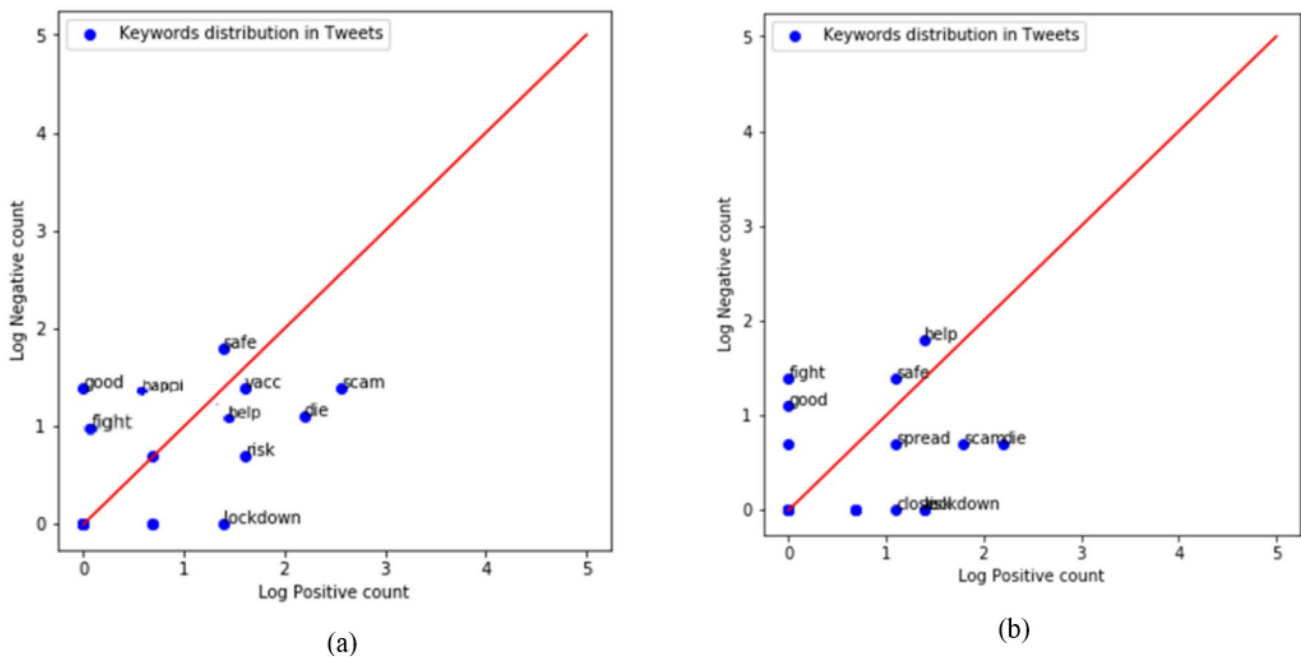
Fig. 4 Tweets keywords classification—**a** First test; **b** Second test
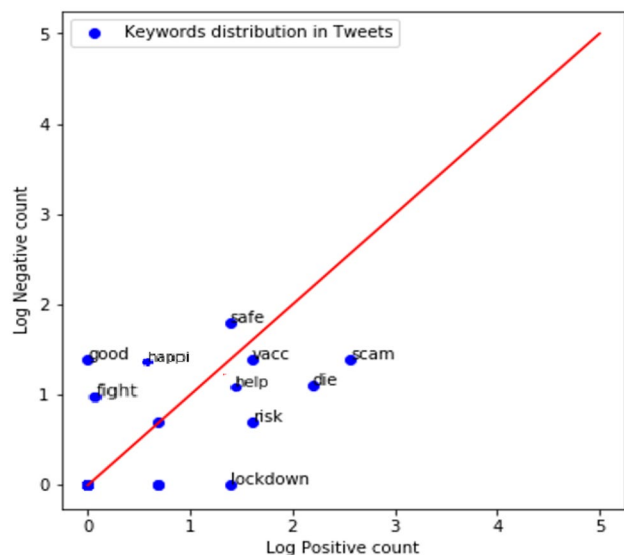
Fig. 5 Tweets keywords classification—third test

Several keywords representing sentiments appeared in the tweets' analysis in the different trials based on the tests as shown in Figs. 4 and 5.

Both figures above show the distribution of the keywords in the selected Tweets for each test. As shown in the figures, the more tweets are selected, the more keywords are identified. Note that the Log function is used to scale the data and avoid showing large values in the graph. The tweets are classified based on their frequencies separating the positive and negative sentiments. This classification of the keywords brings back some common words retrieved in the n-gram analysis, such as 'lockdown', 'happy', 'scam', 'risk'. Based on the results, people's sentiments are mixed, and some positive and negative emotions are coming back with the tweet analysis.

5 Conclusion

This study attempts to explore and analyze people's sentiments on the vaccine during the COVID-19 pandemic by using natural language processing algorithms for sentiment analysis to classify texts and extract the corresponding polarity. As shown, the sentiments are mixed with some positive and others with negative feelings. The main keyword used for retrieving the tweets is #COVIDVaccine. From this work, the author concludes that the analysis of microblogging such as Twitter can help determine people's opinions and feelings. This work focused mainly on the COVID-19 vaccine and what tweeters feel about it. Several tweets were collected, analyzed, classified based on keywords frequency.

For future work, more testing shall be done on tweets using machine learning techniques, comparing the results with the NLP techniques, and generalizing the algorithm to different hashtags and other applications.

References

Abualigah L, Alfar HE, Shehab M, Hussein AMA (2020) Sentiment analysis in healthcare: a brief review. In: Abd Elaziz M, Al-qaness M, Ewees A, Dahou A (eds) Recent advances in NLP: the case of Arabic Language. Studies in computational intelligence, vol 874. Springer, Cham. https://doi.org/10.1007/978-3-030-34614-0_7

Ainin S, Feizollah A, Anuar NB, Abdullah NA (2020) Sentiment analyses of multilingual tweets on halal tourism. Tour Manag Perspect 34:100658. https://doi.org/10.1016/j.tmp.2020.100658

Alamoodi AH, Zaidan BB, Zaidan AA et al (2021) Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: a systematic review. Expert Syst Appl 167:114155. https://doi.org/10.1016/j.eswa.2020.114155

Beaubien J (2020) COVID-19 Is Officially A Pandemic, Declares World Health Organization : NPR. https://www.npr.org/2020/03/11/814603236/covid-19-is-officially-a-pandemic-declares-world-health-organization. Accessed 11 Feb 2021

Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends. Knowl-Based Syst 226:107134. https://doi.org/10.1016/j.knosys.2021.107134

Bollinger R, Ray S (2020) New variants of coronavirus: what you should know | Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/a-new-strain-of-coronavirus-what-you-should-know. Accessed 20 Feb 2021

Cheng M, Keaten J (2021) UN: COVID-19 herd immunity unlikely in 2021 despite vaccines. https://apnews.com/article/canada-coronavirus-pandemic-united-nations-acffe5031e65b472e5a4848ab516163f. Accessed 11 Feb 2021

Duong H-T, Nguyen-Thi T-A (2021) A review: preprocessing techniques and data augmentation for sentiment analysis. Comput Soc Netw 8:1. https://doi.org/10.1186/s40649-020-00080-x

Hassan S-U, Aljohani NR, Idrees N et al (2020) Predicting literature's early impact with sentiment analysis in Twitter. Knowl-Based Syst 192:105383. https://doi.org/10.1016/j.knosys.2019.105383

Hauthal E, Burghardt D, Fish C, Griffin AL (2020) Sentiment analysis. In: Kobayashi A (ed) International encyclopedia of human geography, 2nd edn. Elsevier, Oxford, pp 169–177

Hindustan Times (2020) Covid-19: list of countries which have already approved emergency use of vaccine

Kausar S, Xu H, Shabir MY, Ahmad W (2019) A sentiment polarity categorization technique for online product reviews. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2963020

Kushwah S, Kalra B, Das S (2021) Sentiment analysis of healthcare big data: a fundamental study. In: Bansal JC, Paprzycki M, Bianchini M, Das S (eds) Computationally intelligent systems and their applications. Springer, Singapore, pp 53–70

Le T, Andreadakis Z, Kumar A et al (2020) The COVID-19 vaccine development landscape. Nat Rev Drug Discov. https://doi.org/10.1038/d41573-020-00073-5

Lovelace B (2020) World Health Organization names the new coronavirus: COVID-19. In: CNBC. https://www.cnbc.com/2020/02/11/world-health-organization-names-the-new-coronavirus-covid-19.html. Accessed 20 Feb 2021

Lurie N, Saville M, Hatchett R, Halton J (2020) Developing Covid-19 vaccines at pandemic speed. N Engl J Med 382:1969–1973. https://doi.org/10.1056/NEJMp2005630

Manguri HK, Rebaz NR, Pshko RMA (2020) Twitter sentiment analysis on worldwide COVID-19 outbreaks. Kurd J Appl Res https://doi.org/10.24017/covid.8

Minaee S, Kalchbrenner N, Cambria E et al (2021) Deep learning–based text classification: a comprehensive review. ACM Comput Surv 54:1–40. https://doi.org/10.1145/3439726

Muñana C, 2020 (2020) KFF COVID-19 Vaccine Monitor: Dece 2020

Nurdeni DA, Budi I, Santoso AB (2021) Sentiment analysis on Covid19 vaccines in Indonesia: from the perspective of Sinovac and Pfizer. In: 2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT). pp 122–127

Olliaro P, Torreele E, Vaillant M (2021) COVID-19 vaccine efficacy and effectiveness—the elephant (not) in the room. Lancet Microbe 2:e279–e280. https://doi.org/10.1016/S2666-5247(21)00069-0

Prabhakar Kaila DR, Prasad DAVK (2020) Informational flow on Twitter – Corona virus outbreak – topic modelling approach. Social Science Research Network, Rochester, NY

Raamkumar AS, Tan SG, Wee HL (2020) Measuring the Outreach efforts of public health authorities and the public response on Facebook during the COVID-19 pandemic in early 2020: cross-country comparison. J Med Internet Res 22:e19334. https://doi.org/10.2196/19334

Rajput A (2019) Natural language processing, sentiment analysis and clinical analytics. CoRR abs/1902.00679:

Rokade P, Aruna D (2019) Business intelligence analytics using sentiment analysis-a survey. Int J Electr Comput Eng 9:613. https://doi.org/10.11591/ijece.v9i1.pp613-620

Schvaneveldt RW, Meyer DE, Becker CA (1976) Lexical ambiguity, semantic context, and visual word recognition. J Exp Psychol Hum Percept Perform 2:243–256. https://doi.org/10.1037/0096-1523.2.2.243

Sesagiri Raamkumar A, Tan SG, Wee HL (2020) Use of HEALTH BELIEF MODEL-BASED DEEP LEARNING Classifiers for COVID-19 social media content to examine public perceptions of physical distancing: model development and case study. JMIR Public Health Surveill 6:e20493. https://doi.org/10.2196/20493

Shamrat FM, Chakraborty S, Imran MM et al (2021) Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. Indones J Electr Eng Comput Sci 23:463–470. https://doi.org/10.11591/ijeecs.v23.i1.pp463-470

Solangi YA, Solangi ZA, Aarain S, et al (2018) Review on natural language processing (NLP) and its toolkits for opinion mining and sentiment analysis. In: 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS). pp 1–4

Soong H-C, Ayyasamy RK, Akbar R (2021) A review towards deep learning for sentiment analysis. In: 2021 International Conference on Computer Information Sciences (ICCOINS). pp 238–243

Srivastava R, Bharti PK, Verma P (2022) A review on multipolarity in sentiment analysis. In: Joshi A, Mahmud M, Ragel RG, Thakur NV (eds) Information and communication technology for competitive strategies (ICTCS 2020). Springer, Singapore, pp 163–172

Statista (2021) COVID-19 vaccines ordered by manufacturer 2021. In: Statista. https://www.statista.com/statistics/1195971/number-of-covid-19-vaccine-doses-by-manufacturer/. Accessed 20 Feb 2021

Vijaykumar S, Meurzec RW, Jayasundar K et al (2017) What's buzzing on your feed? Health authorities' use of Facebook to combat Zika in Singapore. J Am Med Inform Assoc 24:1155–1159. https://doi.org/10.1093/jamia/ocx028

WHO (2020) Coronavirus Disease (COVID-19). In: WHO - Prequalification Med. Prod. IVDs Med. Vaccines Immun. Devices Vector Control. https://extranet.who.int/pqweb/vaccines/covid-19-vaccines. Accessed 9 Aug 2021

WHO (2021) WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. https://covid19.who.int/. Accessed 13 Feb 2021

Wong MCS, Wong ELY, Huang J et al (2021) Acceptance of the COVID-19 vaccine based on the health belief model: a population-based survey in Hong Kong. Vaccine 39:1148–1156. https://doi.org/10.1016/j.vaccine.2020.12.083

World Economic Forum (2021) President Xi address: Urgening the International Community to offer necessary support to developing countries. https://www.weforum.org/events/the-davos-agenda-2021/sessions/special-address-by-g20-head-of-state-government-67e386f2d5. Accessed 11 Feb 2021

Xue J, Chen J, Hu R et al (2020) Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. J Med Internet Res 22:e20550. https://doi.org/10.2196/20550

Yadav S, Ekbal A, Saha S, Bhattacharyya P (2018) Medical sentiment analysis using social media: towards building a patient assisted system. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). European Language Resources Association (ELRA), Miyazaki, Japan