

Article

It's Your Turn, Are You Ready to Get Vaccinated? Towards an Exploration of Vaccine Hesitancy Using Sentiment Analysis of Instagram Posts

Mohammed Talha Alam ^{1,2}, Shahab Saquib Sohail ^{1,*} , Syed Ubaid ^{1,3} , Shakil ^{1,4}, Zafar Ali ⁵,
Mohammad Hijji ⁶ , Abdul Khader Jilani Saudagar ⁷  and Khan Muhammad ^{8,*} 

- ¹ Department of Computer Science and Engineering, School of Engineering Sciences and Technology, Jamia Hamdard, New Delhi 110062, India
 - ² Department of Machine Learning, Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi P.O. Box 54115, United Arab Emirates
 - ³ Department of Computer Science, Aligarh Muslim University, Aligarh 202001, India
 - ⁴ Faculty of Electronic and Information Technology, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warszawa, Poland
 - ⁵ School of Computer Science and Engineering, Southeast University, Nanjing 211189, China
 - ⁶ Faculty of Computers and Information Technology (FCIT), University of Tabuk, Tabuk 47711, Saudi Arabia
 - ⁷ Information Systems Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia
 - ⁸ Visual Analytics for Knowledge Laboratory (VIS2KNOW Lab), Department of Applied Artificial Intelligence, School of Convergence, College of Computing and Informatics, Sungkyunkwan University, Seoul 03063, Korea
- * Correspondence: shahabssohail@jamiahamdard.ac.in (S.S.S.); khanmuhammad@g.skku.edu (K.M.)



Citation: Alam, M.T.; Sohail, S.S.; Ubaid, S.; Shakil, A.; Ali, Z.; Hijji, M.; Saudagar, A.K.J.; Muhammad, K. It's Your Turn, Are You Ready to Get Vaccinated? Towards an Exploration of Vaccine Hesitancy Using Sentiment Analysis of Instagram Posts. *Mathematics* **2022**, *10*, 4165. <https://doi.org/10.3390/math10224165>

Academic Editor: Victor Leiva

Received: 26 September 2022

Accepted: 26 October 2022

Published: 8 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The deadly threat caused by the rapid spread of COVID-19 has been restricted by virtue of vaccines. However, there is misinformation regarding the certainty and positives outcome of getting vaccinated; hence, many people are reluctant to opt for it. Therefore, in this paper, we identified public sentiments and hesitancy toward the COVID-19 vaccines based on Instagram posts as part of intelligent surveillance. We first retrieved more than 10k publicly available comments and captions posted under different vaccine hashtags (namely, covaxin, covishield, and sputnik). Next, we translated the extracted comments into a common language (English), followed by the calculation of the polarity score of each comment, which helped identify the vaccine sentiments and opinions in the comments (positive, negative, and neutral) with an accuracy of more than 80%. Moreover, upon analysing the sentiments, we found that covaxin received 71.4% positive, 18.5% neutral, and 10.1% negative comments; covishield obtained 64.2% positive, 24.5% neutral, and 11.3% negative post; and sputnik received 55.8% positive, 15.5% neutral, and 28.7% negative sentiments. Understanding vaccination perceptions and views through Instagram comments, captions, and posts is helpful for public health officials seeking to enhance vaccine uptake by promoting positive marketing and reducing negative marketing. In addition to this, some interesting future directions are also suggested considering the investigated problem.

Keywords: COVID-19; vaccination; vaccine hesitancy; sentiment analysis; health informatics; data analytics; intelligent surveillance; infodemics

MSC: 68T01

1. Introduction

COVID-19 began to wreak havoc on all aspects of life in December 2019, and the World Health Organization (WHO) declared it a pandemic in March 2020 [1]. Following the spread of COVID-19, the public's interest in vaccine-related questions and issues grew. Vaccines have been developed by the research teams of pharmaceutical companies, research

organisations, universities, and countries around the world [2]. Vaccination is still required to prevent the spread of this infection [3], but the procedure has encountered difficulties due to public opposition [4], hesitancy [5], and fear [6]. The WHO vaccination record lists hesitancy as one of the top ten concerns [7]. Furthermore, the WHO defines it as the “delay in accepting or refusing vaccinations despite vaccine availability” [8]. Political aspirations, religious convictions, anti-vaccination propaganda, and the perception of unusual intrigues are all likely factors affecting vaccine hesitancy, rejection, or refusal [9,10].

Today, people are hesitant to receive vaccines because of false information, misrepresentation, deception plans, the spread of ideas through social media, insufficient and conflicting government statements, public disappointment, and the fear of being unaware [11]. Concerns such as “how safe is the vaccine, aftereffects, and impact” play a role in the occurrence of vaccine hesitancy. Furthermore, false information spreads faster than reliable and accurate information [12]. As a result of the subsequent spread of false information and distortion, the end results are unfavorable, including low vaccination rates, rising treatment costs, and so on [13]. As a result, COVID-19 vaccination has become challenging due to reluctance, rejection, and delays. Therefore, immediate action is needed [14].

To solve the aforementioned challenges, the following research questions (RQ) are addressed in this paper:

1. RQ1: Are people hesitant about taking vaccines or not?
2. RQ2: How do people perceive vaccines, especially covaxin, covishield, and sputnik?

Thus, addressing these research questions can influence many government-level policies, health official planning, and pharmaceutical industries strategies. To this end, we believe that this article offers potential contributions, which can be summarized as follows:

- We extracted Instagram posts to identify the public sentiments towards vaccines. To the best of our knowledge, based on our search, we found no work exclusively examining Insta posts for vaccine hesitancy during COVID-19.
- We incorporated the Selenium Python libraries for the purpose of scraping the on-line posts.
- We incorporated the VADER lexicon under NLTK to analyze the polarity scores of each comment. This gives an over 80% accuracy while identifying sentiment polarization and thus supporting intelligent systems needed for medical surveillance
- We obtained very useful and interesting findings regarding users’ perception of vaccines. These results can help health officials and government services to boost their vaccine drives and attract the masses towards vaccination in a better way. Hence, they could influence the pandemic effects.
- The proposed research methodology offers a platform for exploring various aspects of future infodemic-related research.

The rest of the paper is explained as follows. The related work is described in Section 2. Section 3 contains the details regarding the adopted methodology. In Section 4, the experimental results are shown, with the necessary discussion. We discuss various aspects of the adopted methodology and our findings in Section 5. In addition, potential future extensions of this research are discussed in this section. Finally, the work is concluded in Section 6.

2. Background

This section highlights the motivation of this research, followed by a review of the recent studies on related work.

2.1. Motivation

Instagram is a popular social media network that enables quick image and comment sharing across users’ profiles. It is a widely used and well-liked image-based social media platform that is quickly gaining popularity. It is also used as a mobile “app” for social photography. Instagram is a relatively new social media platform that gives users the chance to share photographs and videos of their experiences with others. Hashtags are

frequently added to the uploaded images. Since Instagram employs the words “following” and “follower,” and allows users to leave comments on posted photographs, it operates very similarly to Twitter. The “@” and “#” symbols are often used by users to reference other users and to group posts with the same hashtags together in the captions they add to the images, respectively. Brands can easily reach consumers and convey their unique perspectives, and businesses can utilize Instagram to market their ideas. Young people use Instagram to filter their own images and produce aesthetically appealing feeds. Instagram advertising can be a powerful tool for strengthening brand ties with customers.

Our motivation, in considering Instagram as a primary social network, is the outcome of the following search query entered in Scopus, wherein we observed 196 responses for the following query:

("Sentiment" AND "COVID" AND "Vaccine" AND "Twitter").

Surprisingly, only six results were observed using the following query:

("Sentiment" AND "COVID" AND "Vaccine" AND "Instagram").

This indicates that relatively few researchers have examined Instagram posts and comments to investigate how people perceive vaccinations. Figure 1 shows a word cloud that illustrates the outcomes of the first query. It is evident that Twitter- and Facebook-related works have been adequately reported; however, Instagram-based sentiment analysis with respect to the COVID vaccine is yet to be properly addressed.

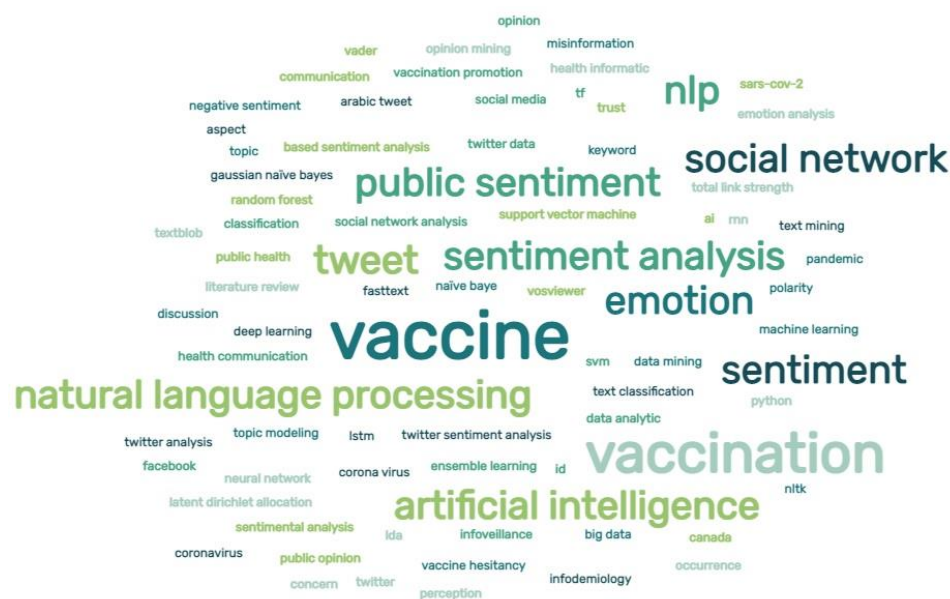


Figure 1. Word cloud for the Scopus outcomes based on the search of the authors' keywords related to research on ("COVID" and "Twitter" and "sentiment" and "vaccine").

In this regard, Wawrzuta et al. examined the arguments offered by anti-vaxxers in relation to COVID-19 vaccinations on various social media platforms [15]. Amanatidis et al. explored sentiments regarding the Pfizer (New York, NY, USA), AstraZeneca (Cambridge, UK), and Johnson and Johnson (Leiden, Netherlands) vaccines [16]. In this work, we primarily aim to explore the sentiments of the people of India. However, the three primarily used vaccinations in India, namely, covaxin, covishield, and sputnik, have not yet received enough attention in the literature. We thus included these vaccinations in our research and examined the public sentiment on them.

2.2. Related Work

Researchers have significantly contributed to the exploration of various aspects of the “COVID-19” pandemic. These studies range from those identifying symptoms of the diseases [17], sensor-based COVID-19 solutions [18,19], and epidemic analysis [20,21] to

those suggesting behavioral measures that can be used to handle social situations during periods affected by deadly infections [22]. Presently, vaccines are playing a vital role in restricting the spread of the disease. Hence, identifying the public opinion regarding vaccines is also important [23]. To this end, there are studies that incorporate sentiment analysis techniques to identify users' opinions on different aspects of pandemic; however, very few have explored the impact on vaccines. One study [24], for example, reported a systematic review on the sentiment analysis of infectious diseases and outbreaks published in the previous decade. They discovered that COVID-19 is still ambiguous, even though the cases are increasing rapidly. They concluded that by using sentiment analysis, we can prevent diseases such as COVID-19 from resurfacing. Another study [25] examined how people felt during the pandemic and discovered that countries such as India and Australia reported positive feelings about their awareness programmes, whereas Chinese people reported negative feelings about their experiences during the pandemic. Many tweets have been circulated on Twitter with various hashtags and words that reflect the emotions people experienced during the pandemic. These words include COVID-19-related attributes, such as masks and quarantines, as well as sympathy and reactions, etc. The researchers used the NRC Lexicon, which analyzes eight different emotions but excludes sarcasm and irony. The study's limitation was that the tweets were only in English.

Chakraborty et al. [26] suggested the value of emotional intelligence for identifying emotion and sentiments. They claimed that emotional intelligence can help people to filter tweets, reducing anxiety and depression, which can affect aged persons or sensitive people. In another study [27], to examine the emotional nature of numerous tweets, researchers applied the recurrent neural network (RNN) for emotional prediction, searching for connections between words and categorizing them as indicating positive or negative emotions. Rather than simple positive and negative extremes, they placed numerous texts into considerably more specific classes of emotional strength, such as weakly positive and strongly negative. They used TextBlob and RNN to evaluate a sample of 500 tweets using the keyword "COVID". The TextBlob and RNN findings were paired with a keyword-based special data scraper to teach the RNN model using these specific recently scraped datasets. As a result, they were assigned an emotional classification for certain topics and properly segmented using both the models.

A few more studies have been reported that aimed to explore users' perceptions of vaccines through the sentiment analysis of their social network activity. In one of the related studies [28], researchers analyzed the sentiments and opinions of people regarding COVID-19 vaccines from approximately 4.5 million tweets. The Twitter platform can also help to promote public health, increase vaccination adoption, and reduce vaccine skepticism and opposition. Furthermore, to encourage and enhance vaccine uptake, public health organizations could increase positive communications, diminish negative and opposing messages, and proactively suspend anti-vaccination accounts such as bots. In general, vaccine opposition and hesitancy were found to be more prevalent than vaccine interest, but opinions varied by country. Negative tweets were largely about fear, discouragement, anger, and politics, whereas positive tweets were mostly about hope, support, and faith.

Another study [29] emphasized the value of employing machine learning for big data opinion mining, as well as the need to continue refining algorithms to overcome the present obstacles, such as irony and sarcasm recognition. The authors provided a quick and low-cost analysis of vaccination sentiment polarity in a huge number of tweets spanning several years. Another study [30] using the Scopus database included a sentiment analysis as well as a bibliometric analysis of COVID-19 vaccines. Trends in COVID-19- and sentiment-analysis-related issues published by various institutions across the world were the focus of the study.

Another study used a web crawling method to investigate COVID-19 vaccine acceptance in South Korea, using three keywords: "AstraZeneca", "Vaccine", and "Pfizer". They discovered that 71 percent of the words combined with "AstraZeneca" were classified as negative, while 56 percent of the words crawled with "Pfizer" were classified

as negative [31]. A mixture of social media content analysis and social network analysis revealed the themes of the vaccination discourse on Twitter. Themes of safety, efficacy, and trust in institutions need to be examined, as tailored outreach initiatives and intervention tactics are used on Twitter to enhance COVID-19 vaccination uptake [32]. Misinformation about vaccines on social media could exacerbate vaccine phobia, slowing the progress toward vaccine-induced herd immunity and potentially increasing the number of infections connected to novel COVID-19 mutations. For these reasons, understanding vaccination hesitancy from the perspective of social media is critical. Because data access is the primary barrier to reaching this goal, the study produced a data collection that may be utilized for investigating anti-vaccine propaganda on social media and enabling a better understanding of vaccination reluctance [33].

The contribution of this article, with respect to the related work, is summarized in Table 1. It is evident that not all the surveys used sentiment analysis for the three vaccines namely, covaxin, covishield, and sputnik. In addition to this, no reported sentiment analysis has been performed using non-English posts and comments from social media to identify the public perception. Therefore, we collected public emotions for all the three vaccines discussed above, and those languages other than English were also included and translated.

Table 1. Contribution of our work with respect to the related literature.

Authors	Model	Year	Covaxin	Covishield	Sputnik	Translation	Reference
Akash Dubey et al.	NRC Emotion Lexicon	2021	✓	✓	×	×	[34]
S. Kumaresh et al.	Logistic Regression and Naïve Bayes	2021	✓	✓	×	×	[35]
Anuratha et al.	NRC Hashtag Sentiment Lexicon	2021	✓	✓	×	×	[36]
Mudassir et. Al.	ABSA, TextBlob, and VADER	2021	✓	✓	×	×	[37]
Alam et al.	LSTM and Bi-LSTM	2021	✓	✓	✓	×	[38]
Our Proposal	VADER and BERT	2022	✓	✓	✓	✓	-

3. The Proposed Methodology

There are numerous vaccines that have been approved and are currently in use around the world. The covaxin, covishield, and sputnik vaccines, on the other hand, have received a great deal of attention in India. Our study focuses on these three vaccines and examines social media interactions to determine the public opinion on each. As demonstrated in the preceding section, there are numerous approaches and methods for determining the sentiment of a statement. This section depicts the overall approach of our investigation. Figure 2 depicts the overall architecture of our study. This analysis is divided into five sections: (a) the data collection and extraction, (b) Instagram comment scraping, (c) dataset creation, (d) language translation, and (e) vaccine sentiment analysis. In the first step, we used the Selenium and ChromeDriver libraries to extract Instagram users' posts and captions containing covaxin, covishield, and sputnik hashtags. The Instagramy library in Python was then used to collect the comments associated with these posts for all three vaccination hashtags. We were able to create a dataset for each vaccine hashtag, as well as the captions extracted in the first step. The fourth stage entailed the translation of each comment and caption into a common language for further analysis using the Googletrans Python package.

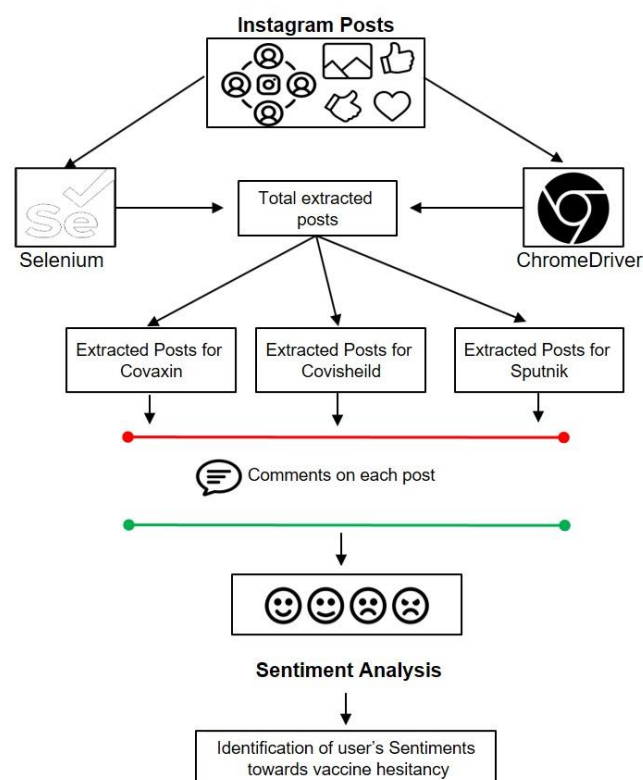


Figure 2. Block diagram for the sentiment analysis of posts and comments from Instagram.

Finally, we calculated the polarity scores for each comment and caption and identified vaccine sentiments and opinions of people using VADER through the Natural Language Toolkit. We then used a sentiment classifier with a pre-trained NLP model, BERT, to compare these sentiments and obtained an accuracy of >80%.

3.1. Data Source and Extraction

The extraction of data from Instagram was performed using the Selenium (Version 2, developed by Thoughtworks) and Chromedriver libraries. Selenium is a Python library and toolset for automating a variety of tasks in internet browsers. Web scraping is one such method for extracting important data and information. Selenium WebDriver uses ChromeDriver, which is a standalone server or separate application, to operate Chrome. It is required to perform Selenium test scripts on the Google Chrome browser [39]. We extracted the posts with three distinct vaccine-related hashtags, namely, #covaxin, #covishield, and #sputnik, that were trending in India. We searched all the public posts and captions that mentioned these hashtags, without language limitations. The search returned a total of 71k, 186k, and 146k posts for #covaxin, #covishield, and #sputnik, respectively. Out of these, 225 posts for each vaccine were captured in five scrolls. The caption and other relevant data (publication date, username, post and profile URL, number of comments, likes, view count, and timestamp) were stored for each post.

3.2. Scraping for Instagram Comments

The comments on each of the captured posts with covaxin, covishield, and sputnik hashtags were scraped with the help of the Instagramy Python package. Instagramy is a Python package that scrapes Instagram without requiring any additional dependencies [40]. This tool can scrape a user's recent post data, such as likes, comments, captions, hashtags, and other post-related information. We can also use this library to retrieve a user's biography and determine whether the profile is verified. After scraping the posts from the previous stage, this tool captured the comments, which amounted to 1k for covaxin, 5k for

covishield, and 3k comments for the sputnik hashtags. For each hashtag, these comments, together with the captions, were forwarded to create a dataset (see Figure 3).

query	error	timestamp	profileUrl	comment	likeCount	replyCount	username	commentId	commentText	ownerId	profilePicture	isReply
https://www.instagram.com/		2021-07-2	https://www.instagram.com/	Chechi i hope you are feeling good post vaccination	0	0	amrutha_j	2021-07-0	1.82E+16	1.74E+10	https://scontent-yz1-1.cdninstagram.com/	FALSE
https://www.instagram.com/		2021-07-2	https://www.instagram.com/	For the past few months the covid-19 outbreak has	0	0	ubaidahsa	2021-07-0	1.81E+16	7.8E+09	https://scontent-yz1-1.cdninstagram.com/	FALSE
https://www.instagram.com/		2021-07-2	https://www.instagram.com/	I wanted a Sputnik vaccine, but I got an Astra Zeneca	3	0	i_am.berni	2021-06-2	1.79E+16	3.91E+09	https://scontent-yz1-1.cdninstagram.com/	FALSE
https://www.instagram.com/		2021-07-2	https://www.instagram.com/	Thanks alot's sir for this information	0	0	cider5391	2021-07-1	1.79E+16	2.72E+10	https://scontent-yz1-1.cdninstagram.com/	FALSE
https://www.instagram.com/		2021-07-2	https://www.instagram.com/	Actually in Feb CovishieldAstrazenica vaccines 2 lakh	0	0	chadaramj	2021-07-1	1.79E+16	4.5E+10	https://scontent-yz1-1.cdninstagram.com/	FALSE

Figure 3. An excerpt from the “comments” dataset.

3.3. Dataset Creation

A comprehensive dataset was created, comprising the above 10k comments and 675 captions. Comments and captions in languages other than English were also included in the dataset. Following the creation of the dataset, a cleaning procedure was run on all the comments and captions to generate a more coherent version of the text retrieved. The cleaning procedure involved eliminating non-representable characters such as hashtags (#), user mentions (@), URLs, and email addresses. In addition, we replaced two or more spaces with a single space. The dataset was saved in the CSV format for further analysis (see Figures 4–6).

postUrl	profileUrl	username	fullName	commentCount	likeCount	pubDate	description	imgUrl	postId	ownerId	type	query	timestamp	videoUrl	viewCount
https://www.instagram.com/	https://www.instagram.com/	sharma.ka	Jaya Sharm	27	318	2021-07-0	India's Covaxin dose	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	2.25E+09	Photo	#covaxin	2021-07-19T19:40:18.436Z		
https://www.instagram.com/	https://www.instagram.com/	shivneepk	SHIVNEEP	25	1087	2021-06-3	India's Covaxin	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	4.46E+10	Photo	#covaxin	2021-07-19T19:40:18.436Z		
https://www.instagram.com/	https://www.instagram.com/	paulosilva	Desenhos	2	219	2021-07-1	@desenhosdonand	https://scontent-yz1-1.cdninstagram.com/	2.62E+18	4.15E+10	Photo	#covaxin	2021-07-19T19:40:18.436Z		
https://www.instagram.com/	https://www.instagram.com/	amu	nev dY... dY... dY...	0	469	2021-07-1	and #Covaxin have	https://scontent-yz1-1.cdninstagram.com/	2.62E+18	6.12E+09	Photo	#covaxin	2021-07-19T19:40:18.436Z		

Figure 4. The #covaxin post metadata with additional details such as the URL, likeCount, viewCount, and others.

postUrl	profileUrl	username	fullName	commentCount	likeCount	pubDate	description	imgUrl	postId	ownerId	type	query	timestamp	imgUrl	videoUrl	viewCount
https://www.instagram.com/	https://www.instagram.com/	mehabook	Mehabook	310	83871	2021-07-0	Finally	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	3.13E+09	Photo	#covishield	2021-07-21T07:00:05.250Z			
https://www.instagram.com/	https://www.instagram.com/	patusu_ba	Nesan Bali	11	1	2021-07-2	#covishield 1st	https://scontent-yz1-1.cdninstagram.com/	2.62E+18	6.12E+09	Photo	#covishield	2021-07-2	https://scontent-yz1-1.cdninstagram.com/		
https://www.instagram.com/	https://www.instagram.com/	therjgeet	RJ Geet dY	15	193	2021-07-2	2nd dose done	https://scontent-yz1-1.cdninstagram.com/	2.62E+18	1.5E+09	Photo	#covishield	2021-07-2	https://scontent-yz1-1.cdninstagram.com/		
https://www.instagram.com/	https://www.instagram.com/	floatingstr	Floatingstr	2	338	2021-07-0	F U L L Y V A C	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	3.58E+10	Photo	#covishield	2021-07-2	https://scontent-yz1-1.cdninstagram.com/		

Figure 5. The #covishield post metadata with additional details such as the URL, likeCount, viewCount, and others.

postUrl	profileUrl	username	fullName	commentCount	likeCount	pubDate	description	imgUrl	postId	ownerId	type	query	timestamp	videoUrl	viewCount
https://www.instagram.com/	https://www.instagram.com/	yashuteja	Yash Tutej	2	332	2021-06-3	Vaxed is	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	1.93E+09	Photo	#sputnik	2021-07-19T19:07:31.516Z		
https://www.instagram.com/	https://www.instagram.com/	spacetechn	Space Tec	15	1885	2021-06-2	Sputnik	https://scontent-yz1-1.cdninstagram.com/	2.6E+18	1.13E+10	Photo	#sputnik	2021-07-19T19:07:31.516Z		
https://www.instagram.com/	https://www.instagram.com/	i_am_sput	I am Sput	16	139	2021-07-0	I never we	https://scontent-yz1-1.cdninstagram.com/	2.61E+18	1.91E+08	Photo	#sputnik	2021-07-19T19:07:31.516Z		
https://www.instagram.com/	https://www.instagram.com/	shaxayusu	dY... dY... dY...	5	1	2021-07-1	Vaksinatsi	https://scontent-yz1-1.cdninstagram.com/	2.62E+18	6.33E+09	Video	#sputnik	2021-07-1	https://scontent-yz1-1.cdninstagram.com/	4

Figure 6. The #sputnik post metadata with additional details such as the URL, likeCount, viewCount, and others.

3.4. Language Translation

The sentiment analysis was used to study a variety of vaccinations, including covaxin, covishield, and sputnik. However, we discovered that there is a requirement to translate non-English social media messages. To deal with this issue, we used the Googletrans library to translate those posts to include a wider range of perspectives in the analysis (see Figure 7). Since the dataset included texts in a variety of languages from across the globe, it was necessary to translate each text into a common language before proceeding with the analysis. With all the comments and captions cleaned, the entire dataset was translated into the English language using the Googletrans Python library. Googletrans is a Python package that implements the Google Translate API and is available for free and unlimited use. It makes calls to operations such as detect and translate using the Google Translate Ajax API. More than 4000 comments and captions were reported to be in a language other

than English; however, they were all translated into English before being subjected to further data analysis.

```
import pandas as pd
from googletrans import Translator

df = pd.read_csv(r"path")
transl = df.loc[:, 'comment']
translations = {}
translator = Translator()
for i in range(len(transl)):
    # Unique elements of the column
    print(translator.translate(transl[i], dest="en").text)
```

Figure 7. Code snippet for the comments' translation into English language [41].

3.5. Sentiment Analysis for Identifying Vaccine Hesitancy

This section describes the sentiment analysis performed on the dataset generated after the data translation phase. This analysis was carried out using the Jupyter Notebook (Version 4.4.0, developed by Fernando Pérez, Brian Granger), a server-client application that allows users to edit and execute notebook documents using a web browser, making it ideal for prototyping and sharing notebooks with visualizations. Several Python libraries were used in the investigation, including Pandas, NumPy, and Scikit-Learn. The sentiment analysis was carried out with the help of the Natural Language Toolkit (NLTK) using the Valence Aware Dictionary and sEntiment Reasoner (VADER). NLTK is a Python toolkit used for working with natural language processing (NLP) that is available with several text-processing libraries [42]. VADER, on the other hand, is a sentiment analysis tool which is widely used for exploring social network platforms to analyze users' sentiments. VADER is a powerful tool which indicates the degree of positivity or negativity of the sentiments, in addition to scoring these orientation [43].

We imported NLTK and implemented the VADER lexicon to analyze the polarity scores of each comment and caption included in the dataset. The comments were classified as positive, negative, or neutral, and their compound polarity scores, which represented the degree of positivity or negativity of a comment, were determined using VADER's Sentiment Intensity Analyzer. Based on their compound polarity scores, the comments, along with the captions, were labeled as positive or negative. To validate our findings, we subjected our dataset to a pre-trained NLP model called BERT. The dataset was split based on the Pareto principle for training and testing purposes. Then, this fine-tuned model made predictions for each comment and caption by labeling them as positive or negative. Using Scikit-Learn, a classification report was then generated for the labeled and computed polarity scores. The performance metrics, such as the precision, recall, and *F*-1 score, were calculated as shown in Table 2. A confusion matrix was then created, and the accuracy for each vaccine hashtag was determined.

Table 2. Experimental findings of sentiment polarization for different hashtags after analysis.

Comments Analysis	#Covaxin	#Covishield	#Sputnik	Comment Analysis
Positive	1369	3210	2190	Positive
Neutral	354	1225	608	Neutral
Negative	195	565	1126	Negative
Total	1918	5000	3925	Total

4. Experimental Results

The goal of this research was to identify and confirm people's scepticism about receiving COVID-19 vaccines. To accurately assess this apprehension about vaccinations, the results were compared to the predictions of a pre-trained NLP model. The initial step involved obtaining vaccine-related posts, which numbered 71,661 for covaxin, 186,811 for covishield, and 146,647 for sputnik. Following that, 675 posts were collected from five scrolls, along with their captions for further processing. This process was followed by the scraping of the comments on the posts that included these three vaccination hashtags. The search resulted in 1918, 5000, and 3925 comments for the covaxin, covishield, and sputnik hashtags, respectively. A dataset with text in multiple languages was prepared from an overall collection of 10,843 comments and 675 captions. A total of 4353 (40% of total) comments in languages other than English were identified (refer to Figure 8). This complete dataset was then translated into a common language, i.e., English, for the further analysis of the study.

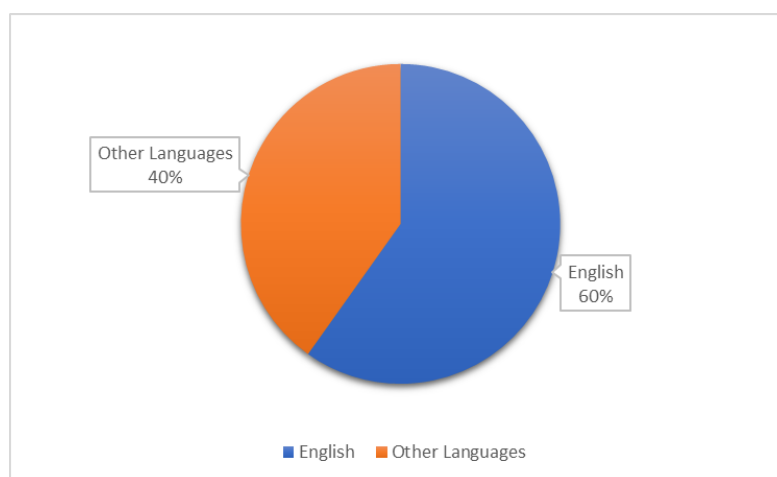


Figure 8. Distributions of Instagram comments from the perspective of English and non-English languages.

4.1. Findings

The polarity scores of each comment and caption, which were now in one language, English, were computed using NLTK's VADER lexicon. The dataset for each hashtag was imported into Jupyter Notebook using Python's Pandas library, and the comments and captions were grouped as positive, negative, or neutral using VADER's Sentiment Intensity Analyzer. For comments under the covaxin hashtag, 1369 of them were categorized as positive, 354 as neutral, and 195 conveyed negative sentiments. Similarly, for covishield, 3210 comments were classified as positive, 1225 as neutral, and 565 as negative. Finally, for sputnik, the results indicated 2190 positive comments, 608 neutral, and 1126 negative comments, as shown in Table 2. The captions, on the other hand, contrasted with the comments in that most of them were characterized as having a negative tone. Among the 675 captured captions, 463 of them were identified as negative, 99 of them were of neutral tone, and 115 captions were labelled as positive ones.

Figure 9 depicts the sentiment polarity distribution among the covaxin, covishield, and sputnik hashtags. For a better understanding of the analysis, the above data are represented as percentages. Out of these three different vaccine hashtags, it can be found that the comments in covaxin-related posts garnered 71.4% positive sentiments. Likewise, positive responses to the covishield and sputnik-related posts were determined to be 64.2% and 55.8%, respectively. However, of the three vaccinations, #sputnik received the most negative feedback, with 28.7% of the comments indicating vaccine hesitancy. Negative comments were also received for the #covaxin and #covishield, with 10.1 percent and 11.3 percent of the comments proving to be negative, respectively.

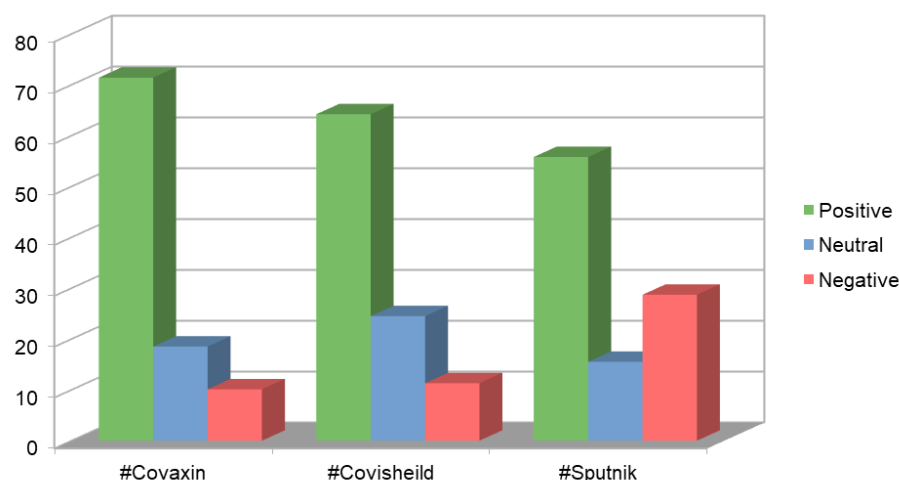


Figure 9. Sentiment polarity distributions for #covaxin, #covishield, and #sputnik vaccines (%age).

4.2. Performance Evaluation

The compound polarity score calculated with the help of the “Sentiment Intensity Analyzer” was then used to categorize the comments as positive or negative. To corroborate this result, we fed our dataset to BERT, a pre-trained NLP model for predicting sentiments. The whole dataset, which included 10,843 comments and 675 captions, was split 80:20. That is, 80% of the dataset was used as the training set, whereas 20% of the dataset was used as the testing set. The loaded BERT model was then configured and fine-tuned before making predictions of our dataset. Using Scikit-Learn, a classification report was generated for the compound polarity scores and the predicted ones.

The confusion matrix had four potential outcomes when analyzing the findings mentioned above. True positive (Tp) is the quantity of anomalous cases that were correctly identified. True negative (Tn) refers to the proportion of correctly identified normal cases. False positives (Fp) are a group of typical events that are mistakenly classified as abnormal situations. False negative (Fn) refers to a group of abnormal instances that were mistakenly thought to be normal cases. The effectiveness of each deep learning classifier is assessed overall using its sensitivity, specificity, accuracy, precision, and other evaluation metrics [44,45].

Precision—the percentage of labels that were correctly predicted positively is represented by the model precision score. Another name for precision is the positive predictive value. False positives and false negatives are traded off using the precision together with the recall.

$$Precision = \frac{Tp}{Tp + Fp}$$

Recall—the model’s accuracy in predicting positives as distinguished from actual positives is measured by the model recall score. This differs from the precision, which counts how many of the total number of positive predictions produced by the models are truly positive. Another name for recall is sensitivity, or the true positive rate. The model’s ability to recognize positive instances is demonstrated by a high recall score.

$$Recall = \frac{Tp}{Tp + Fn}$$

F-1 Score—the model score as a function of the recall and accuracy is represented by the model F-1 score. As an alternative to accuracy measurements, the F-score is a machine learning model performance statistic that equally weights the precision and recall when assessing how accurate the model is.

$$F - 1 \text{ Score} = \frac{2 \times precision \times Recall}{Precision + Recall}$$

Accuracy—the model accuracy is mathematically defined as the ratio of T_p and T_n to all the positive and negative observations, representing one of the most widely used performance metrics for machine learning classification models. In other words, the accuracy indicates the number of times our machine learning model predicted a result accurately out of all the predictions it made.

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100$$

Table 3 depicts different evaluation metrics, such as the precision, recall, F -1-score, and accuracy, for #covaxin, #covishield, and #sputnik. The highest precision, 0.846, was obtained for #covaxin, while the lowest, 0.645, was reported for #sputnik. On the other hand, #covishield was moderately precise, having scored 0.741. The statistics computed for the evaluation metrics of the recall and F -1 Score displayed comparable trends. For the recall, the values inferred were 0.785, 0.239, and 0.208 for #covaxin, #covishield, and #sputnik, respectively. Similarly, covaxin's F -1 score was 0.815, covishield's was 0.362, and sputnik's was 0.315. A confusion matrix was created, and an accuracy score for each hashtag was determined. It was observed that the highest accuracy of 84.61% was achieved for #covaxin, while #covishield and #sputnik attained an accuracy of 74.12% and 64.54%, respectively, as shown in Figure 10.

Table 3. Precision, recall, F -1 score, and accuracy of the proposed sentiment analysis method for the respective vaccines.

Vaccines	Precision	Recall	F -1 Score	Accuracy (%)
#covaxin	0.846	0.785	0.815	84.60
#covishield	0.741	0.239	0.362	74.12
#sputnik	0.645	0.208	0.315	64.54

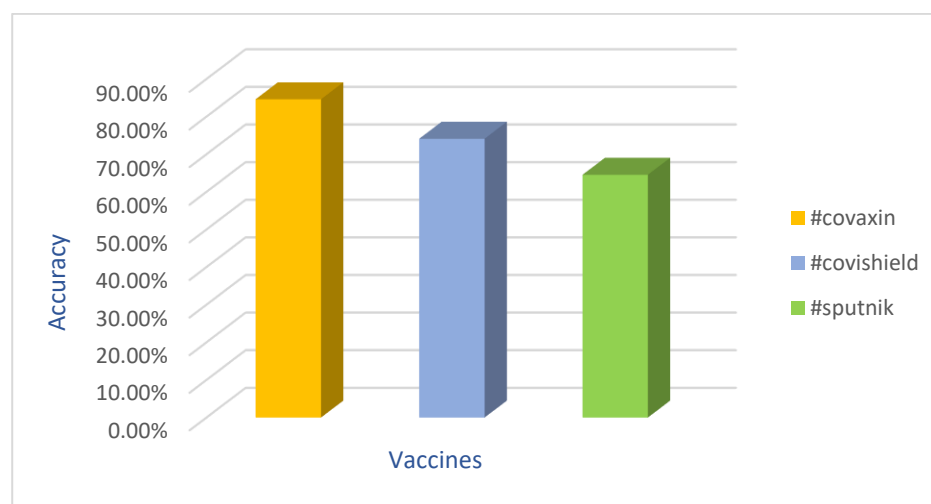


Figure 10. Comparison of the accuracy percentages for #covaxin, #covishield, and #sputnik.

5. Discussion and Research Recommendations

This study included a sentiment and opinion analysis of over 10,000 comments and captions on the COVID-19 vaccines. In our case, we considered the social network platform Instagram, which has been shown to be influential in promoting public health and, as a result, can be useful in increasing the public acceptance of vaccines by minimising public criticism. This can aid in the avoidance of vaccine apprehension. Furthermore, policymakers and government officials must exploit social networking sites, particularly

Instagram, and take appropriate actions, such as blocking fake accounts, to increase vaccine uptake. Furthermore, these steps could be critical in promoting positive vaccine perceptions and may result in a successful vaccine drive. We discovered that organizations usually publish photographs of individuals, and consumers frequently discuss the introduction of vaccinations and indulge in sharing opinions about their effectiveness.

Figure 11 depicts the results of our analysis, which revealed disparities in the sentiment patterns across the comment and caption data. The comment section revealed 62.42% positive, 20.16% neutral, and 17.39% negative sentiments, whereas the caption section revealed 68.50% negative, 14.80% neutral, and 16.70% positive sentiments. According to the study, there were several posts circulating hoaxes about vaccines, as evidenced by the opinions expressed in their captions. After manually inspecting the posts, we discovered a significant number of posts promoting tales of strange reactions to vaccinations, such as magnetism [46]. On the contrary, the comments on these posts expressed positive opinions about vaccinations and challenged the claims made in most of the posts.

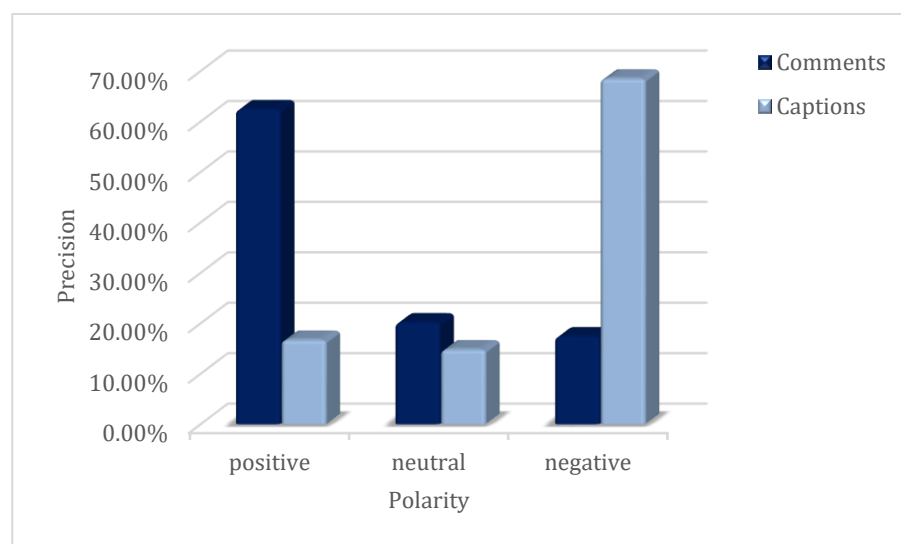


Figure 11. Comparison of the sentiments of the comments and captions in %age.

As they were the most common vaccinations in India, we compared three distinct vaccines: covaxin, developed by Bharat Biotech, covishield, manufactured by the Serum Institute of India (SII), and sputnik V produced by Russia's Gamaleya Research Institute of Epidemiology and Microbiology. As a result, it was critical to comprehend people's perspectives and apprehensions about vaccinations in India. This investigation was necessary, since most of the misleading propaganda originates from social media, and many people fall prey to it. The current study also concluded that a good percentage of people had favorable opinions about covaxin but had negative views about the sputnik vaccination. Covishield elicited a mixed response, with some pushing for the vaccination and others expressing doubts and fear.

In addition to the above findings, we conducted a literature review, and to our knowledge, this is one of the few studies that examines sentiment polarity with respect to vaccination in over 10,000 comments and captions, and the only one to do so by accounting for non-English comments and captions. The extracted information in languages other than English has been filtered out in past studies, and this has been cited as one of the study's limitations [47]. Future expansions of this research might include the directions discussed below.

5.1. Considering Irony and Sarcasm in the Study

The detection of sarcasm and irony remains difficult for both manual and machine classification approaches. Some preliminary research suggests that sarcasm alone can

contribute to a decline of up to 50% in the sentiment detection accuracy when used automatically [48]. Carvalho et al. focused on identifying a collection of surface patterns that may be used to detect sarcastic phrases in a Portuguese online newspaper, with the essential aspects being the usage of punctuation marks and emoticons [49]. Wallace et al. aimed to investigate the identification of irony using contextual data, integrating noun phrases and sentiments retrieved from comments. They advocated the use of information about the conversational threads in which the comments are included. Their strategy is predicated on the assumption that members of various user communities are prone to being sarcastic about specific topics. The researchers used a Reddit comment dataset [50]. Karoui et al. recently published a method for distinguishing between ironic and non-ironic tweets in French, proposing a two-stage paradigm. In the first half, they approached irony detection as a binary classification problem. The system then tried to correct the misclassified cases by asking Google to validate the tweets using negation. The authors employed an SVM as a classifier and obtained an F measure of 0.87 [51].

5.2. Including Results from Additional Social Media Sites

Other social networking sites, such as Facebook and Twitter, should also be viewed for a broader perspective. Additionally, more recent posts relating to other vaccine companies such as Moderna (Cambridge, MA, USA), Pfizer (New York, NY, USA), Sinovac (Beijing, China), and others should be included in the dataset to obtain more accurate findings. Zhu et al. combed through a collection of tweets categorized by phrases that may reflect how people feel about COVID vaccinations. Two alternative rating scales were utilized in this study and ranking a tweet in this way might enable the distinction between an opinion being lost and an opinion being featured on the site, which influences people's decisions and behavior [52]. The authors of another study examined a popular Facebook post regarding children's immunization and used the Linguistic Inquiry and Word Count (LIWC) tool to analyze the language used by users in a comments section. Pro-vaccination, anti-vaccination, and irrelevant (control) comments were compared in terms of the percentage of words used across a range of subjects [53]. Alhujaili et al. explored sentiment analysis techniques and methodologies that can be applied to YouTube content. They also discussed and classified these procedures, which are valuable for data mining and sentiment analysis research [54].

5.3. Incorporating Further Details into the Dataset

The accompanying pictures, emoticons, and URLs that were filtered could be added to the study, as the sentences were labelled for polarity without considering the context. For instance, Gonzalez et al. tested their hypothesis on two datasets: self-tagged tweets and manually annotated tweets. They classified a group of self-tagged tweets with the #sarcasm or #sarcastic hashtag as sarcastic instances, and some positive and negative tweets (retrieved using hashtags such as #happy, #joy, and #lucky, as well as #sadness, #angry, and #frustrated, respectively) as non-sarcastic cases [55]. In another research, the authors considered how emoji characters are used on social media and how they affect text mining and sentiment analysis. They discovered that using emoji characters in sentiment analysis produces higher sentiment scores and appears to have a greater influence on the total sentiments of positive opinions rather than negative ones [56]. If such additional information were to be provided in the dataset, the sentiment analysis findings would be more reliable and effective.

To address the limitations described above, we recommend the method in which a Sarcasm Detector [57] is combined with an Emoji Sentiment Ranking [58] approach to obtain better and more exact outcomes. Figure 12 depicts the recommended technique, which analyses more social networking sites and additional inputs, such as emojis and sarcastic phrases.

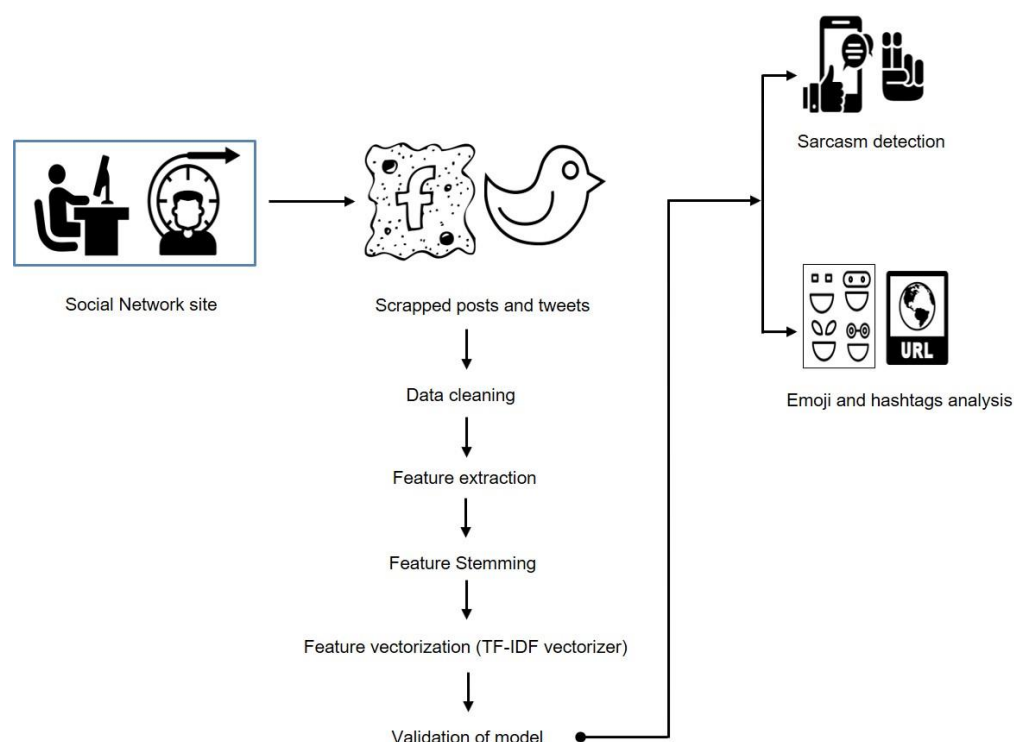


Figure 12. Research recommendation for sentiment identification by investigating emojis and sarcastic posts.

6. Conclusions

The vaccines have played a vital role in combatting COVID-19. This paper aimed to identify public perceptions regarding vaccines in general and covaxin, covishield, and sputnik in particular. How users perceive the act of getting vaccinated was explored through their Instagram posts. The presented work helps identify users' opinions on, and perceptions of, vaccines, which could eventually boost the efforts of health workers and policy makers to drive their vaccine campaigns effectively, especially in the countries where the studied vaccines are in use, such as India. In addition, it also fosters a better understanding of how people react to vaccine drives. Another significant contribution of this work is the high accuracy of the language translation of comments on the Instagram posts, which can serve as a benchmark for social network-based sentiment analysis research.

In future work, it would be interesting to investigate and compare how other vaccines have been received by users, not only in India but also in other regions. Furthermore, it is critical to investigate how the demographic data of the users is distributed. Finally, to include more users in our study, we will seek data from various social media platforms, such as Facebook, Twitter, and Instagram.

To gain a better understanding of the users' perceptions, some research recommendations were suggested in Section 5. We posit that the present work will act as a platform for exploring various dimensions of the field of sentiment analysis and identifying its effects, being especially helpful for healthcare workers, policy makers, and researchers seeking to provide an adequate response to COVID-19, smart health, and medical surveillance.

Author Contributions: Conceptualization, M.T.A. and S.S.S.; methodology, M.T.A. and S.U.; software, M.T.A.; validation, K.M., S.S.S. and Z.A.; formal analysis, M.T.A. and S.; investigation, M.T.A., K.M. and S.S.S.; writing—original draft preparation, S., S.U., M.T.A. and S.S.S.; writing—review and editing, K.M., M.H. and A.K.J.S.; visualization, K.M., M.T.A. and S.S.S.; supervision, K.M.; project administration, K.M. and S.S.S.; funding acquisition, A.K.J.S. and K.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) for funding and supporting this work through the Research Partnership Program No. RP-21-07-02.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: This work was supported by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) through the Research Partnership Program No. RP-21-07-02.

Conflicts of Interest: The authors report that there are no competing interests to declare.

References

1. World Health Organization. WHO COVID-19 Dashboard. 2021. Available online: <https://covid19.who.int/cdc> (accessed on 1 November 2021).
2. Le, T.T.; Cramer, J.P.; Chen, R.; Mayhew, S. Evolution of the COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* **2020**, *19*, 667–668. [CrossRef] [PubMed]
3. Gottlieb, S. America needs to win the coronavirus vaccine race. *Wall Str. J.* **2020**, *26*. Available online: <https://www.wsj.com/articles/america-needs-to-win-the-coronavirus-vaccine-race-11587924258> (accessed on 1 September 2022).
4. Andre, F.E.; Booy, R.; Bock, H.L.; Clemens, J.; Datta, S.K.; John, T.J.; Lee, B.W.; Lolekha, S.; Peltola, H.; Ruff, T.A.; et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull. World Health Organ.* **2008**, *86*, 140–146. [CrossRef] [PubMed]
5. Lyu, J.C.; Han, E.L.; Luli, G.K. COVID-19 vaccine—Related discussion on Twitter: Topic modeling and sentiment analysis. *J. Med. Internet Res.* **2021**, *23*, e24435. [CrossRef] [PubMed]
6. Abbasi, J. COVID-19 conspiracies and beyond: How physicians can deal with patients' misinformation. *JAMA* **2021**, *325*, 208–210. [CrossRef]
7. Ball, P. Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn. *Nature* **2020**, *581*, 251–252. [CrossRef]
8. SAGE Working Group. *Report of the SAGE Working Group on Vaccine Hesitancy*; World Health Organization: Geneva, Switzerland, 2014.
9. Rutjens, B.T.; van der Linden, S.; van der Lee, R. Science skepticism in times of COVID-19. *Group Process. Intergroup Relat.* **2021**, *24*, 276–283. [CrossRef]
10. Melton, C.A.; Olusanya, O.A.; Ammar, N.; Shaban-Nejad, A. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *J. Infect. Public Health* **2021**, *14*, 1505–1512. [CrossRef]
11. Puri, N.; Coomes, E.A.; Haghbayan, H.; Gunaratne, K. Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Hum. Vaccines Immunother.* **2020**, *16*, 2586–2593. [CrossRef]
12. Balmas, M. When fake news becomes real: Combined exposure to multiple news sources and political attitudes of inefficacy, alienation, and cynicism. *Commun. Res.* **2014**, *41*, 430–454. [CrossRef]
13. Zhang, J.; Featherstone, J.D.; Calabrese, C.; Wojcieszak, M. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Prev. Med.* **2021**, *145*, 106408. [CrossRef] [PubMed]
14. Van der Linden, S.; Dixon, G.; Clarke, C.; Cook, J. Inoculating against COVID-19 vaccine misinformation. *Clin. Med.* **2021**, *33*, 100772. [CrossRef] [PubMed]
15. Wawrzuta, D.; Klejdysz, J.; Jaworski, M.; Gotlib, J.; Panczyk, M. Attitudes toward COVID-19 Vaccination on Social Media: A Cross-Platform Analysis. *Vaccines* **2022**, *10*, 1190. [CrossRef] [PubMed]
16. Amanatidis, D.; Mylona, I.; Kamenidou, I.E.; Mamalis, S.; Stavrianea, A. Mining Textual and Imagery Instagram Data during the COVID-19 Pandemic. *Appl. Sci.* **2021**, *11*, 4281. [CrossRef]
17. Menni, C.; Sudre, C.H.; Steves, C.J.; Ourselin, S.; Spector, T.D. Quantifying additional COVID-19 symptoms will save lives. *Lancet* **2020**, *395*, e107–e108. [CrossRef]
18. Cabezas, X.; Martin-Barreiro, C.; Garcia, S.; Delgado, E.; Leiva, V. A two-stage location problem with order solved using a Lagrangian algorithm and stochastic programming for a potential use in COVID-19 vaccination based on sensor-related data. *Sensors* **2021**, *21*, 5352. [CrossRef]
19. Rojas, F.; Leiva, V.; Huerta, M.; Martin-Barreiro, C. Lot-size models with uncertain demand considering its skewness/kurtosis and stochastic programming applied to hospital pharmacy with sensor-related COVID-19 data. *Sensors* **2021**, *21*, 5198. [CrossRef]
20. Cortes-Carvajal, P.D.; Cubilla-Montilla, M.; Gonzalez-Cortes, D.R. Estimation of the instantaneous reproduction number and its confidence interval for modeling the COVID-19 pandemic. *Mathematics* **2021**, *10*, 287. [CrossRef]
21. Xu, J.; Tang, Y. Bayesian framework for multi-wave COVID-19 epidemic analysis using empirical vaccination data. *Mathematics* **2021**, *10*, 21. [CrossRef]

22. Van Bavel, J.J.; Baicker, K.; Boggio, P.S.; Capraro, V.; Cichocka, A.; Cikara, M.; Crockett, M.J.; Crum, A.J.; Douglas, K.M.; Druckman, J.N.; et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **2020**, *4*, 460–471. [CrossRef]
23. Machingaidze, S.; Wiysonge, C.S. Understanding COVID-19 vaccine hesitancy. *Nat. Med.* **2021**, *27*, 1338–1339. [CrossRef]
24. Alamoodi, A.H.; Zaidan, B.B.; Zaidan, A.A.; Albahri, O.S.; Mohammed, K.I.; Malik, R.Q.; Almahdi, E.M.; Chyad, M.A.; Tareq, Z.; Albahri, A.S.; et al. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Syst. Appl.* **2021**, *167*, 114155. [CrossRef]
25. Dubey, A.D. Twitter Sentiment Analysis during COVID-19 Outbreak. Available online: <https://ssrn.com/abstract=3572023> (accessed on 9 April 2020).
26. Chakraborty, K.; Bhatia, S.; Bhattacharyya, S.; Platos, J.; Bag, R.; Hassanien, A.E. Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. *Appl. Soft Comput.* **2020**, *97*, 106754. [CrossRef]
27. Nemes, L.; Kiss, A. Social media sentiment analysis based on COVID-19. *J. Inf. Telecommun.* **2021**, *5*, 1–15. [CrossRef]
28. Yousefinaghani, S.; Dara, R.; Mubareka, S.; Papadopoulos, A.; Sharif, S. An Analysis of COVID-19 Vaccine Sentiments and Opinions on Twitter. *Int. J. Infect. Dis.* **2021**, *108*, 256–262. [CrossRef]
29. Piedrahita-Valdés, H.; Piedrahita-Castillo, D.; Bermejo-Higuera, J.; Guillem-Saiz, P.; Bermejo-Higuera, J.R.; Guillem-Saiz, J.; Sicilia-Montalvo, J.A.; Machío-Regidor, F. Vaccine hesitancy on social media: Sentiment analysis from June 2011 to April 2019. *Vaccines* **2021**, *9*, 28. [CrossRef]
30. Sarirete, A. A Bibliometric Analysis of COVID-19 Vaccines and Sentiment Analysis. *Procedia Comput. Sci.* **2021**, *194*, 280–287. [CrossRef]
31. Lee, H.; Noh, E.B.; Park, S.J.; Nam, H.K.; Lee, T.H.; Lee, G.R.; Nam, E.W. COVID-19 Vaccine Perception in South Korea: Web Crawling Approach. *JMIR Public Health Surveill.* **2021**, *7*, e31409. [CrossRef]
32. Boucher, J.C.; Cornelson, K.; Benham, J.L.; Fullerton, M.M.; Tang, T.; Constantinescu, C.; Mourali, M.; Oxoby, R.J.; Marshall, D.A.; Hemmati, H.; et al. Analyzing social media to explore the attitudes and behaviors following the announcement of successful COVID-19 vaccine trials: Infodemiology study. *JMIR Infodemiol.* **2021**, *1*, e28800. [CrossRef]
33. Muric, G.; Wu, Y.; Ferrara, E. COVID-19 Vaccine Hesitancy on Social Media: Building a Public Twitter Dataset of Anti-vaccine Content, Vaccine Misinformation and Conspiracies. *arXiv* **2021**, arXiv:2105.05134. [CrossRef]
34. Dubey, A.D. Public Sentiment Analysis of COVID-19 Vaccination Drive in India. Available online: <https://ssrn.com/abstract=3772401> (accessed on 24 January 2021).
35. Kumares, S. Sentiment analysis of COVID-19 vaccine in a social media platform using machine learning techniques. *Synd.-J. Manag.* **2022**, *290*, 1056–1057.
36. Anuratha, K.; Sujeetha, S.; Nandhini, J.M.; Priya, B.; Paravthy, M. # Vaccine: Using Hashtags from Indian Tweets to Capture and Analyse the Sentiments of People on Vaccination for COVID-19 Pandemic. In *Recent Trends in Intensive Computing*; IOS Press: Amsterdam, The Netherlands, 2021; pp. 88–92.
37. Mudassir, M.A.; Mor, Y.; Munot, R.; Shankarmani, R. Sentiment Analysis of COVID-19 Vaccine Perception Using NLP. In Proceedings of the 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2–4 September 2021; pp. 516–521.
38. Alam, K.N.; Khan, M.S.; Dhruva, A.R.; Khan, M.M.; Al-Amri, J.F.; Masud, M.; Rawashdeh, M. Deep Learning-Based Sentiment Analysis of COVID-19 Vaccination Responses from Twitter Data. *Comput. Math. Methods Med.* **2021**, *2021*, 4321131. [CrossRef] [PubMed]
39. BrowserStack. How to Run Selenium Tests on Chrome Using ChromeDriver. Available online: <https://www.browserstack.com/guide/run-selenium-tests-using-selenium-chromedriver> (accessed on 1 November 2021).
40. Python Package Index. Python Package for Instagram User, Posts, Hashtags and Locations Data. Available online: <https://pypi.org/project/instagramy/> (accessed on 1 November 2021).
41. Python Package Index. Free Google Translate API for Python. Available online: <https://pypi.org/project/googletrans/> (accessed on 1 November 2021).
42. Zahidi, Y.; El Younoussi, Y.; Azroumahli, C. Comparative Study of the Most Useful Arabic-supporting Natural Language Processing and Deep Learning Libraries. In Proceedings of the 2019 5th International Conference on Optimization and Applications (ICOA), Kenitra, Morocco, 25–26 April 2019; pp. 1–10.
43. Kaggle. Vader Lexicon. Available online: <https://www.kaggle.com/nltkdata/vader-lexicon> (accessed on 1 November 2021).
44. Himeur, Y.; Sohail, S.S.; Bensaali, F.; Amira, A.; Alazab, M. Latest Trends of Security and Privacy in Recommender Systems: A Comprehensive Review and Future Perspectives. *Comput. Secur.* **2022**, *118*, 102746. [CrossRef]
45. Sohail, S.S.; Siddiqui, J.; Ali, R. A comprehensive approach for the evaluation of recommender systems using implicit feedback. *Int. J. Inf. Technol.* **2019**, *11*, 549–567. [CrossRef]
46. McGill. Can Vaccines Make Our Body Magnetic? Available online: <https://www.mcgill.ca/oss/article/covid-19/can-vaccines-make-our-body-magnetic> (accessed on 1 December 2021).
47. Wong, L.P.; Lin, Y.; Alias, H.; Bakar, S.A.; Zhao, Q.; Hu, Z. COVID-19 anti-vaccine sentiments: Analyses of comments from social media. *Healthcare* **2021**, *9*, 1530. [CrossRef]

48. Sykora, M.; Elayan, S.; Jackson, T.W. A qualitative analysis of sarcasm, irony and related# hashtags on Twitter. *Big Data Soc.* **2020**, *7*. [CrossRef]
49. Reyes, A.; Rosso, P.; Buscaldi, D. From humor recognition to irony detection: The figurative language of social media. *Data Knowl. Eng.* **2012**, *74*, 1–12. [CrossRef]
50. Kreuz, R.; Caucci, G. Lexical influences on the perception of sarcasm. In Proceedings of the Workshop on Computational Approaches to Figurative Language, Rochester, NY, USA, 26 April 2007; pp. 1–4.
51. Karoui, J.; Benamara, F.; Moriceau, V.; Aussenac-Gilles, N.; Belguith, L.H. Towards a contextual pragmatic model to detect irony in tweets. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL 2015), Beijing, China, 26 July 2015; p. 644.
52. Zhu, W.; Hu, T. Twitter Sentiment Analysis of COVID Vaccines. In Proceedings of the 2021 5th International Conference on Artificial Intelligence and Virtual Reality (AIVR), Kumamoto, Japan, 23–25 July 2021; pp. 118–122.
53. Faasse, K.; Chatman, C.J.; Martin, L.R. A comparison of language use in pro-and anti-vaccination comments in response to a high profile Facebook post. *Vaccine* **2016**, *34*, 5808–5814. [CrossRef]
54. Alhujaili, R.F.; Yafooz, W.M. Sentiment Analysis for Youtube Videos with User Comments. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25 March 2021; pp. 814–820.
55. González-Ibáñez, R.; Muresan, S.; Wacholder, N. Identifying sarcasm in twitter: A closer look. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; pp. 581–586.
56. Shiha, M.; Ayvaz, S. The effects of emoji in sentiment analysis. *Int. J. Comput. Electr. Eng. IJCEE* **2017**, *9*, 360–369.
57. Towards Data Science. Sarcasm Detection: Step towards Sentiment Analysis. Available online: <https://towardsdatascience.com/sarcasm-detection-step-towards-sentiment-analysis-84cb013bb6db> (accessed on 15 December 2021).
58. Kralj Novak, P.; Smailović, J.; Sluban, B.; Mozetič, I. Sentiment of emojis. *PLoS ONE* **2015**, *10*, e0144296. [CrossRef]