



Contents lists available at ScienceDirect

Diabetes & Metabolic Syndrome: Clinical Research & Reviews

journal homepage: www.elsevier.com/locate/dsx

Original Article

Twitter sentiment analysis from Iran about COVID 19 vaccine

Zahra Bokaei Nezhad^{*}, Mohammad Ali Deihimi

Department of Computer Science, Shiraz University, Shiraz, Iran



ARTICLE INFO

Article history:

Received 6 December 2021

Received in revised form

7 December 2021

Accepted 9 December 2021

Keywords:

SARS-CoV-2

COVID-19

Public health

Vaccination

Sentiment analysis

ABSTRACT

Background and aims: The development of vaccines against COVID-19 has been a global purpose since the World Health Organization declared the pandemic. People usually use social media, especially Twitter, to transfer knowledge and beliefs on global concerns like COVID-19-vaccination, hence, Twitter is a good source for investigating public opinions. The present study aimed to assess Persian tweets to (1) analyze Iranian people's view toward COVID-19 vaccination. (2) Compare Iranian views toward a homegrown and imported COVID-19-vaccines.

Methods: First, a total of 803278 Persian tweets were retrieved from Twitter, mentioning COVIran Barekat (the homegrown vaccine), Pfizer/BioNTech, AstraZeneca/Oxford, Moderna, and Sinopharm (imported vaccines) between April 1, 2021 and September 30, 2021. Then, we identified sentiments of retrieved tweets using a deep learning sentiment analysis model based on CNN-LSTM architecture. Finally, we investigated Iranian views toward COVID-19-vaccination.

Results: (1) We found a subtle difference in the number of positive sentiments toward the homegrown and foreign vaccines, and the latter had the dominant positive polarity. (2) The negative sentiment regarding homegrown and imported vaccines seems to be increasing in some months. (3) We also observed no significant differences between the percentage of overall positive and negative opinions toward vaccination amongst Iranian people.

Conclusions: It is worrisome that the negative sentiment toward homegrown and imported vaccines increases in Iran in some months. Since public healthcare agencies aim to increase the uptake of COVID-19 vaccines to end the pandemic, they can focus on social media such as Twitter to promote positive messaging and decrease opposing views.

© 2021 Diabetes India. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Coronavirus (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. As of October 25, 2021, more than 226 million cases and 4 million deaths have been reported worldwide [1]. The development of vaccines against COVID-19 has been a global purpose since the World Health Organization declared the pandemic. Currently, various vaccines are approved worldwide, including messenger RNA (mRNA) vaccines developed by Pfizer/BioNTech and Moderna, ChAdOx1 vaccine from AstraZeneca/Oxford, etc. [2].

It is essential to reach an acceptable protective and sustained herd immunity level through vaccination to end the pandemic. According to scientists, this immunity happens only when a large

portion of a community (approximately 70%) gets vaccines [3]. However, attaining such a level of herd immunity could be complicated as public opinions toward vaccines can change based on different events and even vary between different COVID-19 vaccines [4].

Iran was among the primary nations where COVID-19 has spread in early 2020. Therefore the government has aimed to end the pandemic by vaccination [5]. As a result, COVIran Barekat is a homegrown vaccine developed in Iran by Shifa Pharm Industrial Group [6]. Besides the homegrown vaccine, several imported vaccines are currently used in Iran, including Sputnik Light, Sputnik V, Janssen (Johnson & Johnson), Pfizer/BioNTech, AstraZeneca/Oxford, Moderna, Sinopharm, etc. [7]. Comparing homegrown and foreign vaccines is a hot topic in Iran. Iranian people have engaged in comparing vaccines to decide which one they should get. On the other hand, they use social media such as Twitter to express their views and thoughts online [8], so COVID-19 has been a popular subject on Twitter since January 2020 [9]. Therefore, Twitter lets

^{*} Corresponding author. Eram Sq., Science and Technology Park, Shiraz University, Shiraz, Iran.

E-mail addresses: zbokaei@gmail.com (Z. Bokaei Nezhad), m.a.deihimi@gmail.com (M.A. Deihimi).

health organizations track public perception of COVID-19 vaccination and helps them with better decisions and plans to increase the uptake of COVID-19 vaccines and end the pandemic. As a result, the present study assessed Iranian tweets to (1) Analyze Iranian people's view toward COVID-19 vaccination. (2) Compare Iranian views toward homegrown and imported COVID-19 vaccines.

To identify public opinions toward COVID-19 vaccination, we first need to assign a polarity of 'positive', 'negative', and 'neutral' to each retrieved tweet. To do so, we utilized a pre-trained Persian sentiment analysis model. We deployed deep learning classifiers based on CNN-LSTM Hybrid Model (Convolutional Neural Network – Long Short Term Memory) for sentiment classification, which proved high accuracy in previous works [10]. Subsequently, we created the first vaccine-related dataset in Persian for analyzing Iranian opinion toward COVID-19 vaccination.

To the best of our knowledge, the present study is the first attempt to analyze public concerns regarding COVID-19 vaccines in Iran. Fig. 1 shows the workflow of the suggested methodology.

2. Related works

Twitter has been widely used in several studies on topics related

to the COVID-19 pandemic [11]. Before the pandemic [12], were the first to analyze vaccine-related images on Twitter. They aimed to recognize the connection of the distribution of images being used in vaccine-related tweets. Using a logistic regression model, they also predicted whether or not the image had been retweeted. Ultimately they released a labeled dataset to use as a sentiment classifier for images. Before the pandemic, another research based on analyzing parent forums about health care and vaccination was done by (Tangherlini et al., 2016). They investigated the source of vaccine hesitancy from blogs and discovered some parents might use these platforms to promote vaccine hesitation and false information amongst other parents [13]. conducted a sentimental analysis model on Twitter to define the beliefs of the Indian citizens towards the COVID-19 vaccines. The authors observed that 35% of the tweets had a positive sentiment regarding COVID-19 vaccination, while 16.65% were negative.

[14] showed views for and against COVID-19 vaccination changed by community features such as assets, demographics, and religious or even family situation [2]. explored the attitudes of Twitter users towards three types of vaccines, AstraZeneca, Pfizer, and Moderna. They also investigated events that shaped different sentiments towards vaccination over time using the AFINN lexicon [3]. used Valence Aware Dictionary and sEntiment Reasoner (VADER) and proved that the pattern of defined sentiments and opinions had changed in response to vaccine-related events during the pandemic.

[15] used TextBlob to perform sentiment analysis towards COVID-19 vaccination by web scraping 154 articles on blogs and online newspapers during the COVID-19 pandemic. They concluded that over 90% of the articles contained positive sentiments towards vaccinations [16]. conducted a sentiment analysis and Latent Dirichlet Allocation topic modeling on textual data. They collected data from thirteen Reddit communities related to COVID-19 vaccines. They found that in these communities, sentiments were overall more positive than negative. Moreover, they found no significant change in people's views since December 2020.

[17] classified English and Filipino tweets using a Naïve Bayes model to conduct a sentiment analysis of COVID-19 vaccination in the Philippines. The author used the Rapid Miner data science software with 81.77% accuracy. They concluded that the majority of the tweets were enthusiastic about the idea of vaccination in the Philippines.

Since there is no research focusing on opinion about vaccination in Iran, we did a comprehensive study to investigate Iranian views toward imported and homegrown COVID-19 vaccines. This study can give public agencies an informative insight and help them further decide to increase the uptake of COVID-19 vaccines and end the pandemic.

3. Materials and methods

3.1. Data acquisition

Fig. 1 depicts an overview of the proposed model architecture. Python programming language library called "Tweepy" was connected to the Twitter academic API to collect related tweets. Due to the popularity of AstraZeneca/Oxford, Pfizer/BioNTech, Moderna, Sinopharm, and COVIran Barekat (also known as Barekat) in Iran, we ignored people's opinions related to other types of vaccines in our research. Hence, we conducted separate searches on Twitter regarding foreign and homegrown vaccines. The search keywords for the homegrown vaccine were "واکسن داخلی، واکسن برکت، واکسن داخلی" (Barekat vaccine, homegrown vaccine). Accordingly, the keywords for retrieving tweets about foreign vaccines were "واکسن آسترازنکا، واکسن اکسفورد، واکسن فایزر، واکسن مدرنا، واکسن سینوفارم، واکسن خارجی" (AstraZeneca)

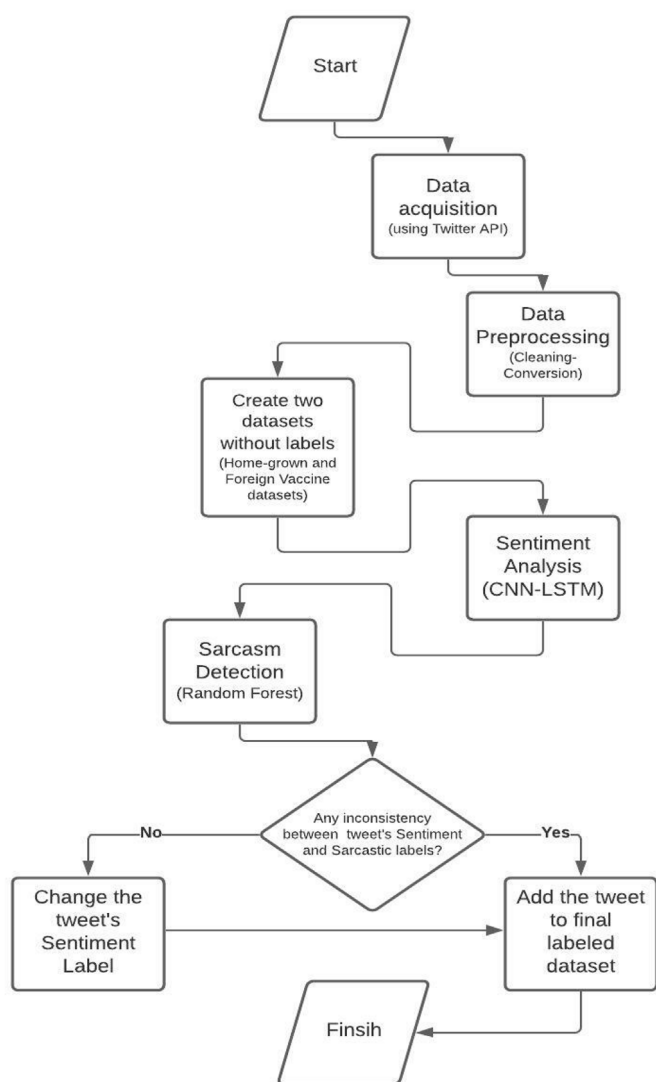


Fig. 1. Flowchart of the proposed model.

vaccine, Oxford vaccine, Pfizer vaccine, Moderna vaccine, Sinopharm, foreign vaccine). Therefore, all Persian-language tweets posted in the time frame from April 1, 2021 to 30 September 2021 that related to the keywords were retrieved.

3.1.1. Data preprocessing

In this step, we cleaned our dataset from non-Persian tweets, URLs, retweets, mentions, and some special characters such as '^ % # - +'. To clean all this trash, we used "re" Python module.

3.1.2. Persian datasets

After preprocessing step, we combined the results of foreign vaccine searches (AstraZeneca/Oxford, Pfizer/BioNTech, Moderna, and Sinopharm) to develop our first dataset. We called it *Foreign-Vaccine* dataset. Accordingly, we built our second dataset, called *Homegrown-Vaccine* dataset.

With 803278 tweets collected from Twitter, the *Homegrown-Vaccine* dataset contained 400839 (49.9%) tweets, and the *Foreign-Vaccine* dataset included 402439 (50.09%) tweets accordingly. Afterward, we applied our sentiment analysis model on each dataset separately and compared the results. In this study, we used several News Agencies reports and reviews to identify potential events that had a possible association with the different sentimental trends in Twitter.

3.2. Vaccine sentiment analysis

We used the pre-trained hybrid deep learning model proposed by Ref. [10], to assign three polarity scores ('positive,' 'negative,' and 'neutral') to each tweet. The model trained on a Persian database consisted of 11616 tweets. The proposed model was based on CNN-LSTM architecture and proved the effectiveness of using deep learning classifiers on Persian datasets. In this architecture, CNN was used as a feature extractor for LSTM on textual input data. The proposed model also used Word2vec as word embedding. Using the model, we labeled both datasets with positive (+1), negative (-1), and neutral (0) labels.

Fig. 2 illustrates the percentage of tweets in each sentiment class. As shown in the figure, the positive sentiments towards foreign vaccines accounted for 43% of tweets ($n = 173,048$), followed by the negative sentiments for 45% and the neutral sentiments for 12%, respectively. On the other hand, Fig. 3 indicates that the positive sentiments towards the homegrown vaccine accounted

for 40% of the tweets ($n = 160,335$), followed by the negative sentiments for 40% and the neutral sentiments for 20%. We found that the number of positive and negative opinions towards home-grown vaccines became the same.

4. Result and discussion

4.1. Overall sentiment towards COVID-19 vaccines

A total of 803278 tweets have been retrieved between April 1, 2021 and September 30, 2021. The frequency of the collected tweets regarding COVID-19 vaccines over six months is illustrated in Fig. 4 for each vaccine group separately. As shown in Fig. 4, there was a spike in the number of foreign vaccines tweets during the second week of April, with approximately 17000 tweets. A possible interpretation of this could be the declaration banning the import of UK and US COVID-19 vaccines in Iran during those times. Another notable rise in the weekly number of tweets can be observed in August's second and third weeks. That was the time when Iran's government officially announced that permission for AstraZeneca/Oxford, Pfizer/BioNTech, and Moderna was issued to import to Iran.

On the other hand, Fig. 3 reveals a rise in the number of homegrown vaccine tweets in August's first and second weeks, with about 17100 tweets. During those times, a public awareness asked the government to submit Barekat vaccine's application to WHO for evaluation. The public awareness raised many people's hesitation towards the homegrown vaccine. Another significant increase regarding Barekat vaccine was observed in May's second and third weeks when Iran's government officially declared successfully developing Barekat vaccine.

4.2. Negative sentiment towards COVID-19 vaccines

The distribution of negative sentiments towards COVID-19 vaccines is illustrated in Fig. 4. As shown in Fig. 5, there was no statically significant difference in the negative sentiment of tweets towards foreign vaccines from April to late July. However, Pfizer's side effects in Iran showed a notable rise in negative tweets between late August and September. This upward trend could also be related to some specific tweeter's accounts that tried to make negative opinions towards particular groups of vaccines [3]. During other months of study, the average negative sentiment towards foreign vaccines had no significant change.

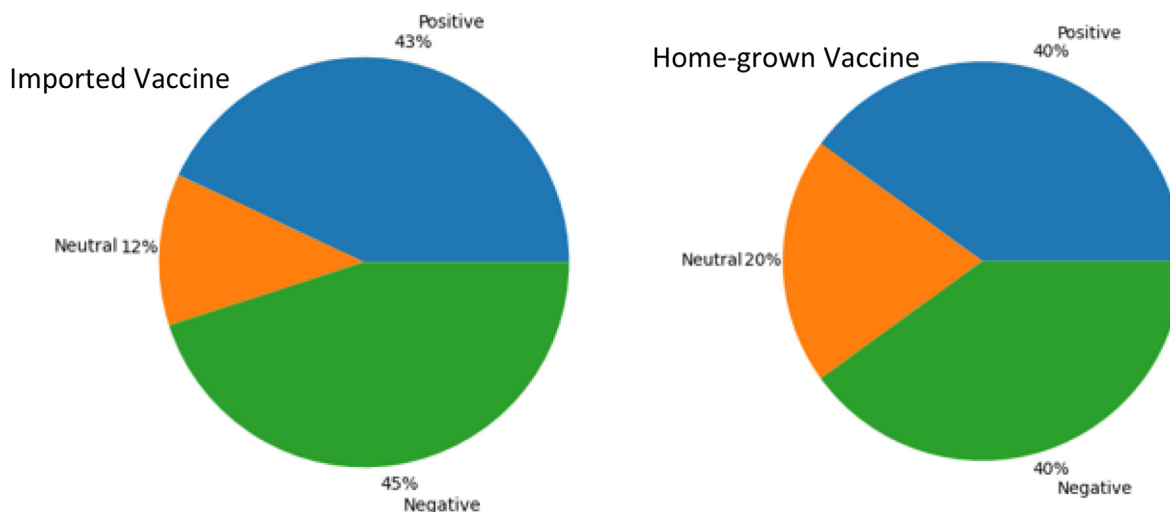


Fig. 2. Percentage of sentiment polarities towards imported and home-grown vaccines.

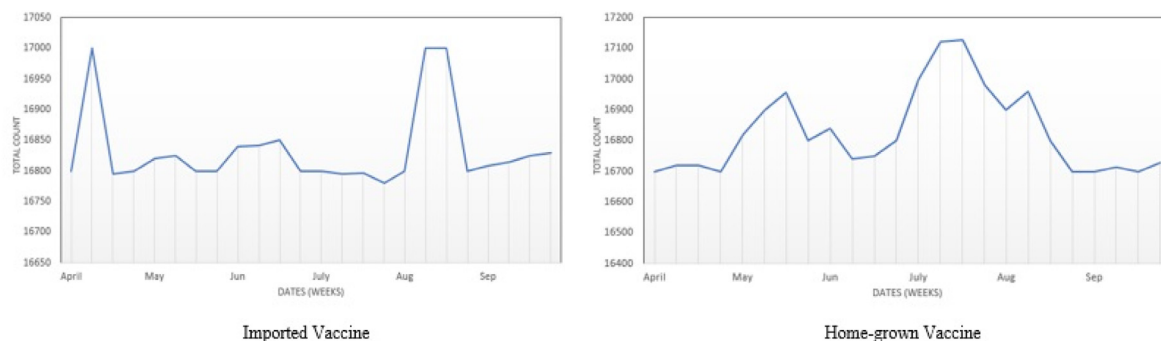


Fig. 3. The frequency of the collected tweets regarding COVID-19 imported and Home-grown vaccines over six months.

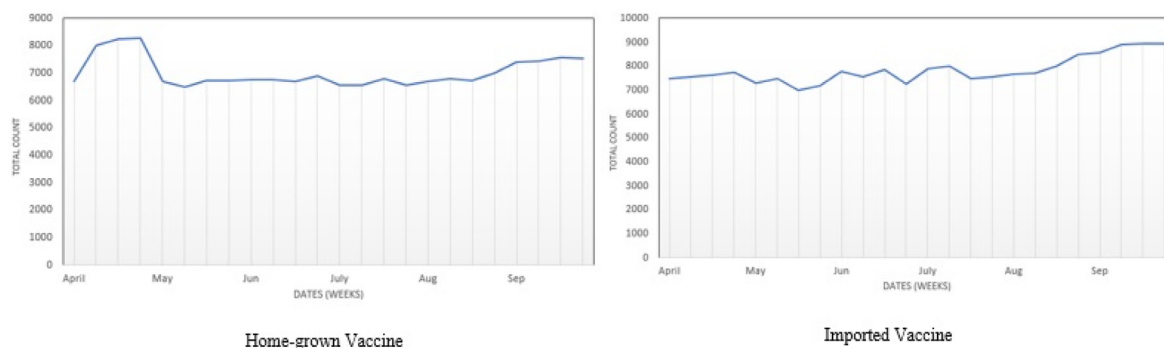


Fig. 4. The distribution of negative sentiments towards COVID-19 imported and home-grown vaccines.

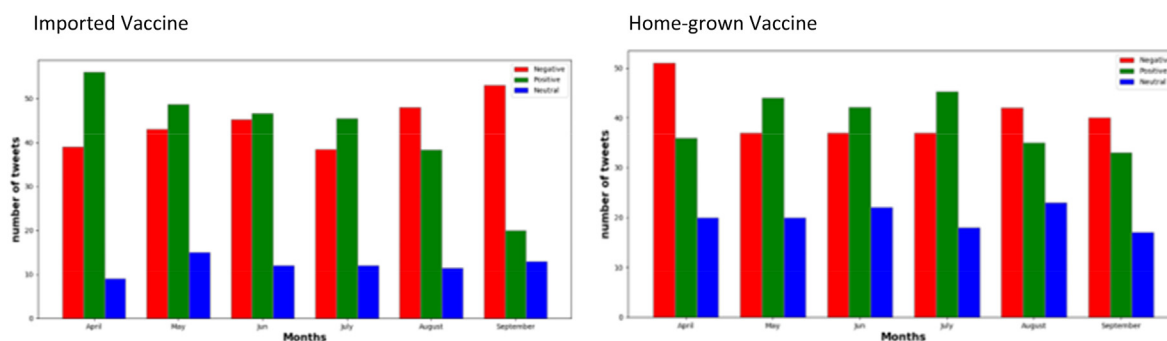


Fig. 5. The distribution of vaccine opinions towards COVID-19 imported and Home-grown vaccines over six months.

On the other hand, as shown in Fig. 5, the negative sentiment towards Barekat vaccine dramatically increased at the beginning of April. As mentioned before, some news claimed about banning the import of the UK and US COVID-19 vaccines in that period. Such reports could correlate with spikes in the negative sentiment toward Barekat vaccine. However, the negative sentiment toward Barekat vaccine remained steady until late August. In fact, between late August and September, an increase in negative sentiments towards the homegrown and foreign vaccines came about during the same period. A possible explanation could relate to the reports claiming Iran's government mandated Barekat vaccine due to Pfizer reported side effects. Hundreds of people opposed Barekat vaccine mandates. However, since no mandate has happened, the negative sentiment toward Barekat decreased in late September.

Fig. 5 depicts the number of tweets with positive, negative, and neutral sentiments per month for each dataset. The figure shows that positive tweets towards foreign vaccines were the dominant

sentiments for approximately four months of study. In comparison, the amount of positive and negative sentiments towards the homegrown vaccine were almost the same. By analyzing sentiments between both groups of COVID-19 vaccines, our results indicated that while the negative sentiment related to foreign vaccines increased, the positive sentiment regarding them decreased no more than 10% for the first four months of the study. However, the negative sentiments towards foreign vaccines saw their first peak in late August and September (for approximately 15%). At the same time, the positive sentiment regarding them decreased dramatically by about 25%. A possible explanation is that these vaccines initially were perceived as "safe" amongst Iranian people until August. In late August, the announcement of some foreign vaccines side effects significantly decreased their positive sentiment. At the same time, a 12% decrease in the positive sentiment towards the homegrown vaccine was observed.

Interestingly, in early April, the announcement of banning the

UK and US vaccines led to a rise in the positive sentiments toward foreign vaccines. In contrast, it caused a decrease in positive sentiments towards the homegrown vaccine as well. In addition, there was no significant difference in the number of neutral tweets for each vaccine group during the study.

5. Conclusion

In this study, we analyzed the sentiments of 803278 Persian tweets concerning COVID-19 vaccines. Tweets retrieved mentioning the homegrown and imported vaccines including COVIran Barekat, Sinopharm, Pfizer/BioNTech, AstraZeneca/Oxford, and Moderna vaccines between April 1, 2021 and September 30, 2021. We then grouped the retrieved tweets into two separated datasets: The *Homegrown-vaccine* dataset (for COVIran Barekat) and the *Foreign-Vaccines* dataset (for the rest of mentioned vaccines). We used a deep learning model based on CNN-LSTM architecture for sentiment analysis.

By comparing the sentiments between two groups of datasets for six months, our results indicated (1) a subtle difference in the number of positive sentiments toward the homegrown and foreign vaccines, and the latter had the dominant positive polarity. In fact, sentiments regarding vaccination remained positively stable throughout the first four months of study. However, we observed a slight decrease in the users' desire to take the vaccine when reports on vaccines' side effects increased in early August and September. (2) The negative sentiment regarding homegrown and imported vaccines seems to be increasing in some months. (3) We also observed no significant differences between the percentage of overall positive and negative opinions toward vaccination amongst Iranian people.

Additionally, we concluded that the issue of the mandatory homegrown vaccine directly led to negative opinions toward it. Moreover, banning the import of foreign vaccines caused positive sentiments toward them and negative views toward the homegrown vaccines accordingly.

It is worrisome that the negative sentiment toward homegrown and imported vaccines increases in Iran in some months. Since public healthcare agencies aim to increase the uptake of COVID-19 vaccines to end the pandemic, they can focus on social media such as Twitter to promote positive messaging and decrease opposing views. Identification of the reasons behind the negative sentiment on social media can help reduce the impact of misinformation.

One of our study's limitations is that collected tweets included just a short time of vaccine availability. Further work can focus on vaccine-related tweets after September, when most people were actively receiving vaccines. Furthermore, the present study did not explore the attitude of Twitter users towards each vaccine separately. We aim to work on this case as a potential future work later. We aim to identify more vaccine sentiments and compare their progression by time, post engagement metrics such as retweets, favorites, replies, and account characteristics to enhance our work.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical approval statement

Hereby, I, Zahra Bokaei Nezhad, consciously assure that for the manuscript **Twitter Sentiment Analysis from Iran about COVID 19 Vaccine** the following is fulfilled:

- 1) This material is the authors' own original work, which has not been previously published elsewhere.
- 2) The paper is not currently being considered for publication elsewhere.
- 3) The paper reflects the authors' own research and analysis in a truthful and complete manner.
- 4) The paper properly credits the meaningful contributions of co-authors and co-researchers.
- 5) The results are appropriately placed in the context of prior and existing research.
- 6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper reference.
- 7) All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] World Health Organization [Online]. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-20-april-2021>; 2020.
- [2] Marcec R, Likic R. Using twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. *Postgrad Med* 2021;1.
- [3] Yousefinaghi 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;108:256–62.
- [4] Chen Lyu J, Le Han E, Luli GK. COVID-19 vaccine-related discussion on twitter: topic modeling and sentiment analysis. *J Med Internet Res* 2021;23(16):1–12.
- [5] Mallapaty S [Online]. Available: <https://www.nature.com/articles/d41586-021-02216-z>. [Accessed 17 August 2021].
- [6] Abdoli A, Aalizadeh R, Aminianfar H, Kianmehr Z, Teimoori A, Azimi E, Emami-pour N, Eghtedardoost M, Siavashi V, Jamshidi H, Hosseini-pour M, Taqavani M, Jalili H. Safety and potency of BIV1-CovIran inactivated vaccine candidate for SARS-CoV-2: a preclinical study. *Rev Med Virol* 2021;1–20.
- [7] McGillUniversity. COVID-19 vaccine tracker. 2021 [Online]. Available: <https://covid19.trackvaccines.org/country/iran-islamic-republic-of/>.
- [8] BokaeiNezhad Z, Deihimi M. Sarcasm detection in Persian. *J Inform Commun Technol* 2020;20(1).
- [9] Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: public sentiment analysis on twitter data and vaccinated population prediction in the USA. *Appl Sci* 2021;11(13).
- [10] BokaeiNezhad Z, Deihimi MA. A combined deep learning model for Persian sentiment analysis. *IJUM Eng J* 2019;129–39.
- [11] Glowacki EM, Wilcox GB, Glowacki JB. Identifying #addiction concerns on twitter during the COVID-19 pandemic: a text mining analysis. *Subst Abuse* 2021;42(1):39–46.
- [12] Chen T, Dredze M. Vaccine images on twitter: analysis of what images are shared. *J Med Internet Res* 2018;20(4):e130.
- [13] Praveen SV, Ittamalla R, Deepak G. Analyzing the attitude of Indian citizens towards COVID-19 vaccine - a text analytics study. *Diabetes Metab Syndr* 2021;595–9.
- [14] Lyu H, Wang J, Wu W, Duong V, Zhang X, Dye TD, Luo J. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. *Intell Med* 2021;1–34.
- [15] Bhagat KK, Mishra S, Dixit A, Chang C-Y. Public opinions about online learning during COVID-19: a sentiment analysis approach. *Sustainability* 2021;13(6):1–12.
- [16] Melton CA, Olusanya OA, NarimanAmmar, ArashShaban-Nejad. 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 Publ Health* 2021;14(10):1505–12.
- [17] Villavicencio C, Villavicencio C, Inbaraj XA, Jeng J-H, Hsieh J-G. Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using Naïve Bayes. *Information* 2021;12(5):204.