

Public Sentiment Towards Vaccination After COVID-19 Outbreak in the Arab World

Rafiul BISWAS^a, Tanvir ALAM^a, Mowafa HOUSEH^a and Zubair SHAH^{a,1}

^aCollege of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

Abstract. Public perception about vaccines is imperative for successful vaccination programs. This study aims to measure the shift of sentiment towards vaccines after the COVID-19 outbreak in the Arab-speaking population. The study used vaccine-related Arabic Tweets and analyzed the sentiment of users in two different time frames, before 2020 (T1) and after 2020 (T2). The analysis showed that in T1, 48.05% of tweets were positive, and 16.47% of tweets were negative. In T2, 43.03% of tweets were positive, and 20.56% of tweets were negative. Among the Twitter users, the sentiment of 15.92% users shifted towards positive, and the sentiment of 17.90% users shifted towards negative. Public sentiment that have shifted towards positive may be due to the hope of vaccine efficacy, whereas public sentiment that have shifted towards negative may be due to the concerns related to vaccine side effects and misinformation. This study can support policymakers in the Arab world to combat the COVID-19 pandemic by utilizing tools to understand public opinion and sentiment.

Keywords. Sentiment analysis, Twitter, Arab, Arabic tweet, COVID-19

1. Introduction

Vaccinating people is thought to be one of the crucial factors that can mitigate the spread of the COVID-19 pandemic. With a large number of people being vaccinated, people can return to a more normal and likely cautious lifestyle as restrictions are eased [1]. Researchers and healthcare professionals worldwide are working to develop an effective vaccine to prevent novel coronavirus strains. However, the success of vaccination programs relies upon their public acceptance. Vaccination intentions are influenced by vaccine efficacy, side effects, mistrust in health care, political concern, health inequality in an ethnic minority and rapid development of vaccines [2]. Exposure to misinformation about the safety and effectiveness of vaccines through online social media often leads people to feel pessimistic and hesitant [3].

Twitter is one of the most popular microblogging social media platforms in the world [4]. Sentiment analysis based on tweets can measure public opinions, attitudes and emotions toward an entity [5]. However, it often becomes difficult to find out the exact meaning of the text when facing short and ambiguous Twitter messages. In this study, sentiment analysis was performed on vaccine-related Arabic tweets to observe public attitudes towards vaccines after the outbreak of COVID-19. It is hoped that this study

¹ Corresponding Author. Zubair Shah, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar. Email: zshah@hbku.edu.qa.

will support understanding the public sentiment about COVID-19 vaccines in the Arab world.

2. Methods

2.1 Data collection

Twitter Academic Research API [6] was used to access vaccines-related public tweets. Tweets were searched using key Arabic words. The search terms included “الطفولة” OR “لقاحات” OR “تطعيم” OR “تطعيمات” OR “لقاح”. Python Tweepy library was used to download the data into the PostgreSQL database. The downloaded data was processed to extract tweets and their metadata such as ID, timestamp, text, author ID, number of likes, retweets and replies, and author profile information, including followers. The IDs of tweets were used as primary key of the table in the database to avoid duplicate entries of tweets.

2.2 Data Preprocessing

The downloaded data was preprocessed before the analysis. First, retweets were removed using string “RT @|RT@” that exists in tweet. Next, tweets of verified users were removed as these users are often government entities and celebrities which may not represent public opinion. After that, punctuation, links, mentions, hashtags, and stopwords were removed. Further, text was processed to remove emojis and special characters. Next, the data was separated into T1 (01 August 2009 to 31 December 2019) and T2 (01 January 2020 to 15 February 2021). Finally, users who posted at least one tweet in both periods were included in the study data. Figure 1 demonstrates the steps included in data preprocessing.

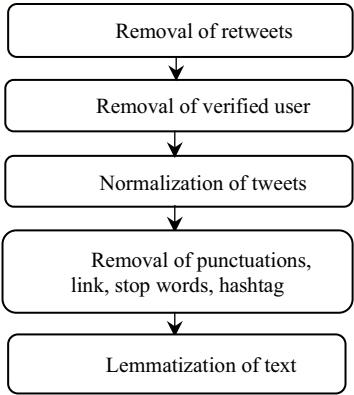


Figure 1. Data preprocessing

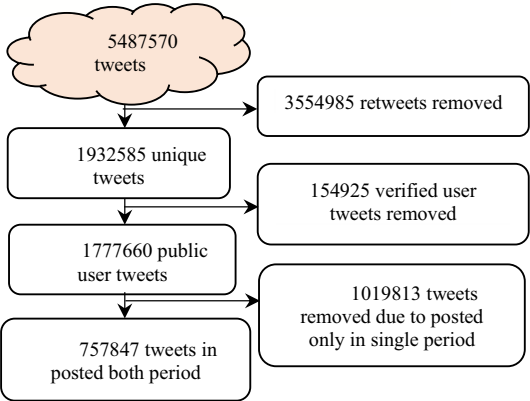


Figure 2. Flowchart of selection of tweets

2.3 Data Analysis

Python Textblob library [7] was used to find sentiment of each word within a range from -1 to +1. Tweets were labelled as positive (>0 to +1), negative (<0 to -1), and neutral

(exact 0). If all the words in a sentence are positive, the polarity is high. Similarly, if all the words in a sentence are negative, the polarity is low. If any sentence has both positive and negative words, polarity is calculated by the summing positive and negative scores. The sentiment shift of users is measured from T1 towards T2. The shift in sentiment is defined as Negative to Positive (N2P), Neutral to Positive (U2P), Negative to Neutral (N2U), Positive to Negative (P2N), Neutral to Negative (U2N) and Positive to Neutral (P2U). The positive shift in sentiment is measured by (N2P, U2P, N2U), and the negative shift in sentiment is measured by (P2U, P2N, U2N). No change in sentiment is measured by (U2U, P2P, N2N).

3. Results

3.1 Collection of Tweets after filtering steps

The search returned nearly 5.5M (5,487,570) tweets. It was found that most of the tweets were from the Kingdom of Saudi Arabia (KSA). People from Egypt, United Arab Emirates, Jordan, and Qatar also shared vaccine information through tweets. About 3.5M (3,554,985, 64.78%) retweets were removed, which left about 2M (1,932,585, 35.72%) unique tweets. A total of 154K (154,925, 8.01%) tweets were removed as they belonged to verified users. After these exclusions, about 1.78M (1,777,660) public tweets were left. After that, tweets from 74,240 users who posted at least one tweet in both time periods were selected, which were 757,847 tweets. Finally, sentiment analysis was performed on these selected tweets. Figure 2 illustrates the overall filtering steps.

3.2 Shift of Sentiment in Arabic Tweets

Table 1. Summary of Sentiment analysis results on Arabic tweets

Period	Positive	Negative	Neutral
T1	112933 (48.05%)	70897 (16.47%)	101359 (35.59%)
T2	203204 (43.03%)	97159 (20.56%)	172063 (36.41%)

Table 1 shows the result of the overall sentiment analysis in both periods. In T1, 112,933 (48.05%) tweets were positive, 70897 (16.44%) were negative and 101,359 (35.59%) were neutral. In T2, 203204 (43.03%) tweets were positive, 97,159 (20.56%) were negative and 172,063 (36.41%) were neutral. Among the 74,240 Twitter users, the sentiment of 15.92% (n=11819) users shifted towards positive from T1 to T2, with 5645 users shifted as U2P, 1,928 users shifted as N2P and 4,246 users shifted as N2U. Whereas the sentiment of 17.90% (n =13,294) users shifted towards negative from T1 to T2, with 3142 users shifted as P2N, 6,567 users shifted as P2U and 3,585 users shifted as U2N. The sentiment of 66.17% (n =49,127) users did not change (P2P, U2U, N2N) during these two periods.

4. Discussion

This study found that almost half of the tweets regarding vaccines were positive before COVID pandemic. However, a decrease in positive sentiment and an increase in negative

sentiment were observed after the COVID-19 outbreak. Users with positive sentiment mainly posted tweets about vaccine effectiveness, and they were optimistic about the role of government and healthcare authorities. Some of the users posted negative tweets regarding the new vaccines for COVID-19, and lack of informations about the side effects of the newly developed vaccines. Negative sentiment users also claimed that the overall vaccine effectiveness has not been studied and expressed concerns to be vaccinated. Some users were neutral in expressing their thought in both periods, and continuously posted neutral tweets. These neutral tweets mainly included vaccine development and vaccine distribution information. In general, users discussed vaccine efficacy and side effects in their tweets.

This study has a few limitations. The method calculated polarity based on linguistic analyses, which may not always be reliable. The tweets were collected in February 2021. As public sentiment may change over time, data collection, analysis and continuous monitoring are necessary to keep track on the latest sentiment.

Twitter dataset used in this study and other relevant files are available at online appendix <https://github.com/rafiulbiswas/PSVA>.

5. Conclusions

From a public health perspective, vaccination is the only way to fight against the COVID-19 pandemic. This study described how vaccination intention varied within two-time frames among the Arab-speaking people based on their native language tweets. The reasons for shifting sentiment downwards need to be analyzed broadly in future studies. We believe this study would support policymakers in taking the necessary steps to be more informed about developing a successful vaccination program in the Arab world.

References

- [1] Williams L, Gallant AJ, Rasmussen S, Brown Nicholls LA, Cogan N, Deakin K, Young D, Flowers P. Towards intervention development to increase the uptake of COVID - 19 vaccination among those at high risk: Outlining evidence - based and theoretically informed future intervention content. *British Journal of Health Psychology*. 2020 Nov;25(4):1039-54.
- [2] Melton CA, Olusanya OA, Ammar N, Shaban-Nejad A. Public Sentiment and Topic Modeling Regarding COVID-19 Vaccines on Reddit Social Media Platform: A Call to Action For Strengthening Vaccine Confidence. *Journal of Infection and Public Health*.
- [3] Sallam M, Dababseh D, Eid H, Al-Mahzoum K, Al-Haidar A, Taim D, Yaseen A, Ababneh NA, Bakri FG, Mahafzah A. High rates of COVID-19 vaccine hesitancy and its association with conspiracy beliefs: A study in Jordan and Kuwait among other Arab countries. *Vaccines*. 2021 Jan;9(1):42.
- [4] Hawkins CM, Duszak R, Rawson JV. Social media in radiology: early trends in Twitter microblogging at radiology's largest international meeting. *Journal of the American College of Radiology*. 2014 Apr 1;11(4):387-90.
- [5] Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*. 2014 Dec 1;5(4):1093-113.
- [6] Twitter. Twitter API for academic research | twitter developer platform," <https://developer.twitter.com/en/products/twitter-api/academic-research>.(accessed on 29 August 2021).
- [7] Adhaamehab. (n.d.). Adhaamehab/Textblob-Ar: Arabic support for textblob. GitHub. [Online]. Available from: <https://github.com/adhaamehab/textblob-ar> (accessed on 29August 2021).