Descriptive analytics of COVID-19 vaccine adopters and hesitant among Twitter users

Tianjie Gu[a], Emmanuel Kagning Tsinda[b], Anthony Dunn[a], Adithya Madhusoodanan[c], Alain B. Zemkoho[a]

[a] University of Southampton, School of Mathematical Sciences, United Kingdom; [b] Graduate School of Medicine, University of Tohoku, Sendai 980-8575, Japan; [c] National Institute of Technology Karnataka, India.

Background

COVID-19 has seriously impacted the health of hundreds of millions of people around the world. With no doubt, vaccines are critical to getting the COVID-19 pandemic under control. However, vaccine hesitancy may slow-down the achievement of herd immunity and continue to facilitate the spread of the virus.

Identifying the sentiments underlying statements and public opinions in social media may be helpful in identifying and motivating people to get vaccinated.

Objective

This project aims to building a machine learning based model to predict the sentiments towards the COVID-19 vaccines based on the words in related tweets. Then, through analyzing the public opinions and sentiments, identify the factors effecting people adopt/reject COVID-19 vaccines and help find the way to motivate people to be vaccinated.

Methodology

Step1
- Data Preprocessing
- Classify the tweets directly by hashtags expressing positive/negative/neutral attitude towards COVID-19 vaccines

Step2
- A machine learning based classification model is built to extract COVID-19 vaccine related opinions expressed by the texts.

Step3
- The resulting classifier developed in this work achieves a high accuracy of 90%, with effective feature engineering steps including various text mining methods such as lemmatization, tokenization and TF-IDF vectorization

Step4
- A LDA model is built to to help discover the abstract topics.
- Clustering algorithms are implemented to further classify the related contents into small topics.

Positive Topics

| 01 Trust in medical institutions | 02 Worries on serious complications led by flu | 03 Vaccines are effective |
| 04 Proud of getting fully vaccinated | 05 No serious side effects after getting fully vaccinated | 06 Trust in the safety and effectiveness of vaccines |

Negative Topics

01 Side effects	02 Lack of trials	03 Dissatisfaction with Bill Gates's huge investment in vaccines
04 Mistrust in vaccines caused by the rising number of cases	05 Getting vaccines or not is a personal right	06 Distrust of the government
07 Trust in immunity system	08 Unfairness in getting access to vaccines	09 Rumor: people will die within 28 days once they got vaccinated
10 Rumor: This is a scam aiming at killing old people	11 Rumor: Dozens of people develop rare blood disorder after taking coronavirus vaccines	

Results

Figure 1
Percentage of tweets with different attitudes towards COVID-19 vaccines by country

	Negative	Positive	Neutral
United States	10%	47%	43%
India	7%	34%	59%
Britain	14%	40%	46%
Canada	11%	41%	48%
Australia	17%	32%	52%
Ireland	11%	45%	44%
South Africa	13%	39%	48%
Pakistan	6%	45%	49%
Uruguay	17%	40%	42%
Philippines	10%	42%	48%
Switzerland	8%	36%	56%
Nigeria	9%	44%	48%
Belgium	5%	41%	54%
France	9%	25%	66%
Kenya	6%	43%	51%
Germany	7%	48%	45%
Brazil	15%	43%	43%
Georgia	20%	40%	40%

Figure 2

	Negative Topic Number										
	1	2	3	4	5	6	7	8	9	10	11
United States	5%	6%	9%	10%	6%	31%	6%	9%	5%	8%	5%
India	6%	9%	8%	13%	4%	24%	4%	11%	9%	7%	4%
Britain	4%	7%	9%	8%	7%	30%	6%	10%	7%	6%	6%
Canada	6%	6%	9%	17%	6%	26%	6%	9%	5%	4%	6%
Australia	3%	8%	7%	27%	2%	21%	6%	9%	4%	5%	8%
Ireland	8%	8%	8%	11%	5%	37%	6%	8%	6%	3%	4%
South Africa	7%	8%	9%	8%	6%	28%	5%	9%	8%	6%	5%
Pakistan	8%	4%	8%	19%	4%	20%	8%	5%	11%	3%	9%
Uruguay	4%	5%	7%	11%	4%	34%	6%	13%	3%	6%	6%
Philippines	5%	15%	5%	12%	4%	19%	16%	7%	4%	7%	7%
Switzerland	2%	7%	2%	24%	6%	28%	13%	13%	2%	2%	2%
Nigeria	7%	9%	3%	14%	10%	33%	3%	5%	3%	9%	3%
Belgium	4%	4%	0%	18%	7%	36%	11%	7%	4%	4%	4%
France	4%	9%	9%	6%	0%	23%	11%	13%	8%	9%	8%
Kenya	0%	14%	6%	14%	3%	34%	3%	11%	6%	0%	9%
Germany	8%	3%	8%	5%	8%	50%	5%	3%	3%	3%	5%
Brazil	12%	6%	12%	6%	5%	28%	9%	9%	4%	4%	5%
Georgia	23%	8%	4%	5%	7%	28%	5%	4%	4%	5%	5%

Figure 3

	Positive Topic number					
	1	2	3	4	5	6
United States	7%	7%	52%	19%	7%	8%
India	11%	5%	60%	8%	5%	10%
Britain	13%	9%	50%	14%	6%	7%
Canada	6%	5%	53%	21%	6%	7%
Australia	9%	7%	63%	7%	5%	8%
Ireland	14%	7%	50%	21%	4%	4%
South Africa	8%	5%	57%	14%	8%	8%
Pakistan	7%	6%	52%	14%	6%	15%
Uruguay	3%	8%	53%	22%	7%	7%
Philippines	4%	6%	58%	18%	5%	9%
Switzerland	9%	5%	51%	5%	4%	26%
Nigeria	14%	10%	52%	5%	5%	13%
Belgium	13%	13%	54%	6%	4%	12%
France	10%	5%	59%	19%	5%	7%
Kenya	6%	4%	62%	5%	7%	17%
Germany	8%	4%	57%	19%	5%	7%
Brazil	6%	5%	47%	26%	5%	11%
Georgia	3%	4%	56%	21%	10%	5%

- By comparing the percentage of tweets with negative/positive/neutral emotions towards COVID-19 vaccines in each country(figure1), one can find that people in Georgia, Australia and Uruguay are less willing to accept vaccination.

- By reading figure2 and figure3, it can be concluded that distrust of the government is the main reason people rejected COVID-19 vaccines and the effectiveness of vaccines is the main reason people accepted it.

- The general trend of either the number of positive/negative/neutral/total tweets keeps rising. But in August 2020, December 2020 and April 2021 the negative tweets reached the peak, this is mainly caused by the fear of side effects and mistrust of the government.

- It is interesting to find that from January 2021 to March 2021, there is an increase in positive tweets but a fall in positive tweets. In this period, tweets with positive topics such as 'got fully vaccinated' surged from zero.
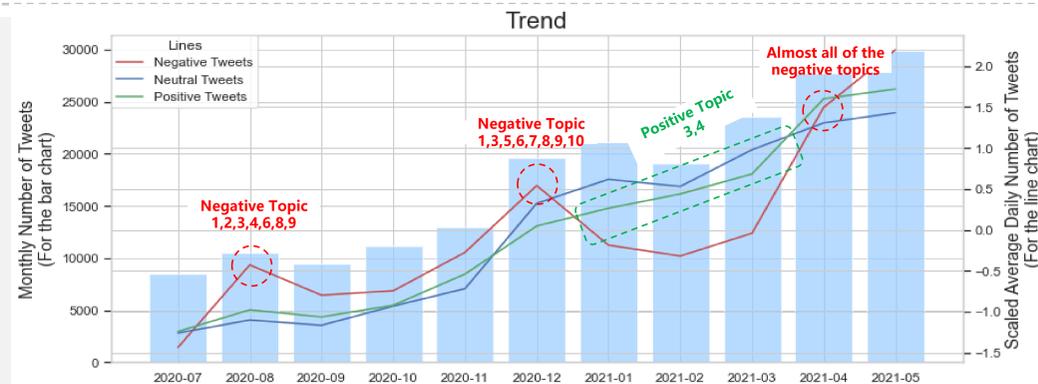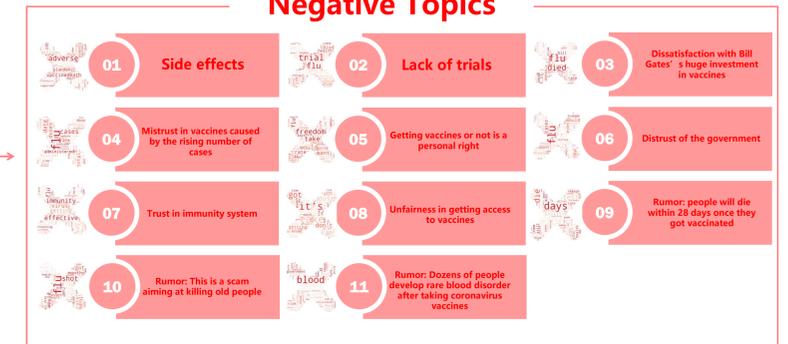

Conclusions

In general, there are more people adopting the COVID-19 vaccines than rejecting. The main factors effecting people deciding whether to accept the vaccines include the doubts about the safety and effectiveness of vaccines, unwillingness to be imposed a vaccine as compulsory and the spread of various rumors about vaccine side effects. The way to reduce vaccine hesitancy is to increase the popularity rate and ensure the fair access to COVID-19 vaccines.

Acknowledgements

References

[1] Solís Arce, Julio S., et al. "COVID-19 vaccine acceptance and hesitancy in low-and middle-income countries." Nature medicine 27.8 (2021): 1385-1394

[2] Griffith, Janessa, Husayn Marani, and Helen Monkman. "COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework." Journal of medical Internet research 23.4 (2021): e26874.