

EN 605.744 Information Retrieval  
Proposal  
Ting He  
Oct 15, 2021

Title:

Content analysis of COVID-19 vaccine side effects on Twitter

Primary Objectives:

- Identify self-report side effects and their modality from Twitter for COVID-19 Pfizer, Moderna and Janssen vaccines along with the times
- Conduct sentiment analysis for 3 vaccines along with the times
- Map to standard concept IDs in Unified Medical Language System
- Compare the distribution to those reported in Vaccine Adverse Event Reporting System

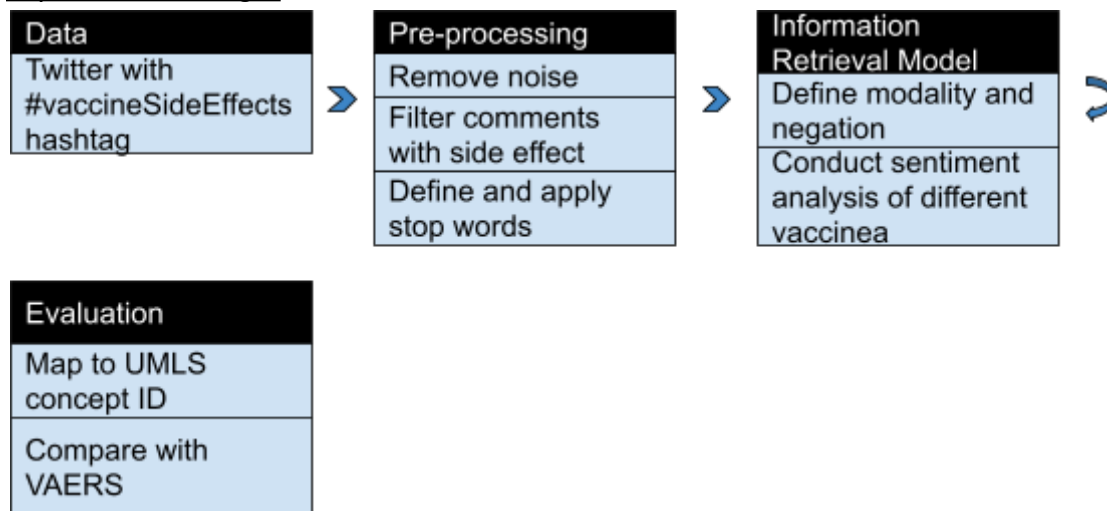
Why this is an interesting or important problem:

Many people who got vaccines might suffer mild or moderate side effects and didn't seek clinical care and reported in the official system.[5]To better understand the full spectrum of symptoms experienced by vaccinated people, I plan to look into Twitter for additional resources.

Sources of data:

Twitter using hashtag #vaccineSideEffects in USA

Experimental design:



Evaluation design:

Compare the distribution in [Vaccine Adverse Event Reporting System](#)

Relevant literature:

Research:

1. Griffith, J., Marani, H., & Monkman, H. (2021). COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework. *Journal of medical Internet research*, 23(4), e26874. <https://doi.org/10.2196/26874>
2. H. Jelodar, Y. Wang, R. Orji and S. Huang, "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach," in *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2733-2742, Oct. 2020, doi: 10.1109/JBHI.2020.3001216.
3. Chen, Q., Leaman, R., Allot, A., Luo, L., Wei, C. H., Yan, S., & Lu, Z. (2020). Artificial Intelligence (AI) in Action: Addressing the COVID-19 Pandemic with Natural Language Processing (NLP). *arXiv preprint arXiv:2010.16413*.
4. Luo, Y. (2021). Using tweets to understand how COVID-19–Related health beliefs are affected in the age of social media: Twitter data analysis study. *J Med Internet Res*, 23(2), e26302.
5. Abeed Sarker, Sahithi Lakamana, Whitney Hogg-Bremer, Angel Xie, Mohammed Ali Al-Garadi, Yuan-Chi Yang, Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource, *Journal of the American Medical Informatics Association*, Volume 27, Issue 8, August 2020, Pages 1310–1315, <https://doi.org/10.1093/jamia/ocaa116>
6. Karami, A., Zhu, M., Goldschmidt, B., Boyajieff, H. R., & Najafabadi, M. M. (2021). COVID-19 Vaccine and Social Media in the U.S.: Exploring Emotions and Discussions on Twitter. *Vaccines*, 9(10), 1059. doi:10.3390/vaccines9101059
7. Lyu, J. C., Le Han, E., & Luli, G. K. (2021). COVID-19 vaccine–related discussion on Twitter: topic modeling and sentiment analysis. *Journal of medical Internet research*, 23(6), e24435.

#### Research(discussion part):

8. Kolokythas, A. (2021). What do the aftermath of the 2010 Haiti earthquake, Hurricane Sandy, the Boston Marathon bombing, the 2013 Ebola outbreak, and the COVID-19 pandemic have in common?. *Oral surgery, oral medicine, oral pathology and oral radiology*, 132(4), 371-372.
9. Spiteri, J. (2021). Media bias exposure and the incidence of COVID-19 in the USA. *BMJ global health*, 6(9), e006798.
10. Ayers, J. W., Chu, B., Zhu, Z., Leas, E. C., Smith, D. M., Dredze, M., & Broniatowski, D. A. (2021). Spread of Misinformation About Face Masks and COVID-19 by Automated Software on Facebook. *JAMA Internal Medicine*.

#### Technical(Pre-processing part):

1. How did Twitter react to the COVID vaccine side effect?  
<https://medium.com/geekculture/how-did-twitter-react-to-the-covid-vaccine-side-effect-d8543798263f>

#### Technical(Modality and negation part):

1. Modality and Negation in Natural Language Processing.  
[https://mirror.aclweb.org/ijcnlp11/downloads/tutorial/tu3\\_present.pdf](https://mirror.aclweb.org/ijcnlp11/downloads/tutorial/tu3_present.pdf)
2. Mood & modality and dialogue sentiment

<https://towardsdatascience.com/mood-modality-and-dialogue-sentiment-b06cd36eca88>

Technical(Sentiment analysis part):

1. <https://towardsdatascience.com/extracting-patient-sentiment-for-pharmaceutical-drugs-from-twitter-2315870a0e3c>

Technical(Time series analysis part):

1. Sentiment Analysis of COVID-19 Vaccine Tweets

<https://towardsdatascience.com/sentiment-analysis-of-covid-19-vaccine-tweets-dc6f41a5e1af>

Outline planned work for the project

Task	Pull data (Twitter and VAERS)	Write into Database (Twitter only)	Pre-processing (Twitter and VAERS)	Define modality and negation	Submit Milestone report	Sentiment analysis	Evaluation	Presentation (video and written report)
Module 7, Oct 16 - 17								
Module 8, Oct 23 - 24								
Module 9, Oct 30 - 31								
Module 10, Nov 6 -7								
Module 11, Nov 13 -14								
Module 12, Nov 20-21								
Module 13, Nov 27-28								
Module 14, Dec 4-5								