Vaccine Vision: A deep learning approach towards identifying societal concerns regarding vaccines

Kaustav Das[1], Shruti Biswas[1]

[1]*Amity University, Kolkata.*

Abstract

In the wake of the COVID-19 pandemic, discerning public sentiments regarding vaccination stances - Pro-Vax, Anti-Vax, or Neutral - emerged as a pivotal undertaking. Leveraging machine learning on COVID-19 related tweets, this study focuses on categorizing individuals based on their vaccination perspectives. However, delving beyond the surface, the analysis uncovers a mosaic of concerns within Anti-Vax sentiments that surpass a mere dichotomy. These concerns encompass a diverse spectrum, spanning from conspiracies and political suspicions to multifaceted uncertainties. In response, this work employs a nuanced multi-label classification approach, aiming to thoroughly comprehend and classify the varied concerns explicitly articulated within Anti-Vax tweets. By scrutinizing a corpus of COVID-19 related tweets, this study endeavors to shed light on the intricate landscape of vaccine hesitancy, providing a comprehensive understanding of the multifaceted reasons underlying Anti-Vax sentiments.

Keywords

Twitter, microblogs, COVID-19, vaccine concerns, tweet, multi-label classification

1. Introduction

Vaccine hesitancy is defined as "delay in acceptance or refusal of vaccination despite the availability of vaccination services"[1]. It is seen as the primary cause of decreasing vaccine rates and the resurgence of vaccine-preventable illnesses in many countries.

To combat the COVID-19 pandemic, researchers and pharmaceutical companies came up with a number of vaccines using the S-protein of SARS-CoV-2. But regardless of the efforts made, number of concerns rose, which led to a significant decrease in the vaccination drives. In the fight against COVID-19, **vaccine hesitancy** has been identified as a significant hindrance.[2]. The reason for most of this hesitancy can be traced back to social media platforms, while social media platforms themselves didn't directly influence the general public, a lot of misinformation by unregulated users whose opinions when voiced through these platforms caused mass paranoia and eventually a distrust towards the public health services. Some factors that contributed to the hesitancy also include the fact that the vaccines being administered, reportedly came with mild to quite severe side effects. Even before the emergence of SARS-CoV-2, WHO had already highlighted vaccine hesitancy as one of the ten leading threats to global health.[3] Thus, it is quite evident that vaccine hesitancy is a significant and complex phenomenon that cannot be avoided. To aid such a situation, a thorough analysis of the general public concern

needs to be done. By virtue of social media, these concerns have been voiced by people in the form of tweets and other social media posts. Therefore, in this paper with the help of deep learning techniques and natural language processing, we have built an efficient multi-label classifier to identify some of the concerns associated with these tweets. Most of the tweets are primarily focused on vaccines associated with COVID-19 like Moderna, Pfizer, Astrazeneca etc. and most of them also voice more than one concern(label) hence the need for a multi-label classifier.

2. Dataset

With our requirements and concerns(labels) defined let's have a look at the dataset.The dataset used for this work can be credited to *"CAVES: A Dataset to facilitate Explainable Classification and Summarization of Concerns towards COVID Vaccines"*[4]. This dataset has been provided as part of the **FIRE conference** [5]. Here are some examples of tweets per label:

CAVES Dataset	
Original Tweet	**Labels**
@SenSanders We have immune systems. Vaccines are filled with animal and human DNA, disinfectants, heavy metals, chemicals. For over 30 years the US has paid out over 4 billion In compensations for deaths and injuries because the US gave Pharma liability freedom in 1986. #VaccinesKill	ingredients, pharma, side-effect
@barbieashdown @leslieh707 Oh my! One does not have to be an expert at reading body language to know he is covering something up. Depopulation is his game. No CV19 vax for me. No Moderna vax for anyone.	conspiracy
@PetenyiSandor @katka_cseh Your choice, go and get a Chinese vaccine. I will wait for vaccines which are approved by the EU. We will have enough by the summer. Let's see how much the Chinese actually will deliver, how many people willing to accept it and if Orban can organize mass vaccination. I doubt it.	country
Canadians who received Vaccine vaccine excluded from seeing Bruce Springsteen, as Broadway opens up via people are against vaccine but Dr. Bonnie Henry stands by all vaccines all are safe seems they have no idea if they are or not folks	side-effect, ineffective
@TraceyLouise @slimschutte @AefreBetty @jon_severs @StevenCross81 Ah yes, that good old fashioned vaccine that was trialled over the correct amount of time, and not forced upon every living being on the planet... Even without this new thing, if you're healthy, you'll still be fine.	unnecessary, rushed, mandatory
Big Pharma spent close to 7 Billion euros on lobbying in France...which as we know is utterly corrupted in terms of Covid response: blocked hydroxychloroquine - like all other Western countries - and embraced remdesivir and vaccines.	pharma

@realDonaldTrump Guess who's gonna get rich before he leaves the white house off of 71 million people he knows will take a shot in the arm if he says so? That's right. Donald Trump now wants science on his side. How much stock in Pfizer did you buy huh?	political
AUTHORITARIAN NAZI-wannabes of Democrat party are COMPLETELY ANTI VOTER ID to PREVENT FRAUD. Those very same LIBERAL TYRANTS want YOU to be DIGITALLY marked under the pretense of getting a "vaccine." NO THANKS! VOTER ID or NO VOTE! Keep both your vaccine and SATAN'S MARK!	political, religious, conspiracy
Now about this Pfizer "vaccine". Who is the government tryna give this to again. By my calculations, the people who would want it are the folks who never got Covid19. Be careful of folks giving you a vaccine that will "protect" you from a virus for which you already had it. https://t.co/zdJqKENupA.	none
@NeilClark66 Mad fascist dictator Johnson now wants to force you to take an experimental vaccine with no long-term safety profile. All this for a virus with 0.3% IFR. Still think this is about a virus? 1922 committee must step in NOW and REMOVE this communist lunatic!	mandatory, rushed, unnecessary

Table 1: Original tweet representation for each label.

2.1. Data Pre-processing

For the pre-processing step, standard text-cleaning steps were followed where lexicons like URLS, usernames, and emojis were removed. Here are the steps:

- Tweet id columns were dropped.
- The label columns were binarized/ one hot encoded using the *sklearn multi-label binarizer* so that they can be used readily for the future. machine learning steps. (https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MultiLabelBinarizer.html)
- Text standardization:
    - Removing HTML elements
    - Replacing non-standard punctuation with standard version.
    - Replacing '\r', '\n' and '\t' with white spaces
    - Removing all control characters
    - Removing duplicate white spaces
- Removing contractions from text using contractions library (https://pypi.org/project/contractions/)
- Replacing usernames and URLs with fillers 'username' and 'url'
- Removing unicode and accented characters from text
- Emojis were replaced with their meaning using the *emoji.demojize()* method of emoji library. (https://pypi.org/project/emoji/)

These steps were also taken in compliance with the standard cleaning steps taken for CT-BERT[6], for getting the best quality bert-embeddings.[1]

Cleaned Data	
Original Tweet	**Cleaned Tweet**
@PaolaQP1231 Well, I mean congratulations Covid19 for being the first ever âœthingâ to eradicate influenza. In other news, Covid vaccines will spur the rise of influenza In 2021-2022 season. Influenza will be returning for a shot at the title belt. Nov 2, 2021 on pay per view. Order today.	username well, i mean congratulations covid19 for being the first ever "thing" to eradicate influenza. in other news, covid vaccines will spur the rise of influenza in 2021-2022 season. influenza will be returning for a shot at the title belt. nov 2, 2021 on pay per view. order today.

Table 2
Original tweet vs Cleaned Tweet(with NLP)

3. Methodology

Before starting with the model training step, we must handle the data imbalance present in our dataset. Thus, to handle the imbalance a custom loss function called Distribution Balanced Loss [7] was used. **COVID Tweeter-BERT (CT-BERT)**[6]: CT-BERT is a domain-specific transformer-based model, pre-trained using a sizable corpus of tweets about COVID-19 that were posted between January 12 and April 16, 2020. It is initialized with BERT-Large[8] weights and then pre-trained/fine-tuned using 160 million tweets regarding the coronavirus. The major advantage of using CT-BERT is that all of our vaccination tweets in the dataset are concerned with COVID-19, for which CT-BERT produces domain-specific, quality embeddings. CT-BERT was our primary model and performed best for the classification task. Other models like Bi-LSTM[9], and BERT variants like ROBERTA were considered which did not perform as well.

4. Evaluation

The primary evaluation metric for our model performance was macro-f1 The model was evaluated on the held-out validation set compromising of 1984 tweets containing tweet instances from each label class.

	macro-F1 score	micro-F1 score
Bi-LSTM(Multi-smote + word2vec)	0.79	0.82
Baseline CT-BERT(without DB-LOSS)	0.55	0.72
CT-BERT(with DB-Loss + without cleaning)	0.78	0.81
CT-BERT(with DB-LOSS)	0.88	0.89

Table 3
Performance of Classifiers on the validation set.

---

[1]CT-BERT is our final model used for classification, it will be explained later on in the methodology section

From the results seen, it can be concluded that **CT-BERT** with DB-LOSS and proper cleaning produces the best result achieving the highest macro-F1 score.

5. Results

To finally summarize and prove the effectiveness of our model we will test our model over a completely unknown distribution of vaccine-related tweets. This dataset [2] contains tweets associated with COVID-19 vaccines as well as some other non-COVID vaccines.

Here is the model performance of **CT-BERT+DB-LOSS** on the same:

Run File	Methodology	Macro-F1 score	Jaccard score
run_submission.csv	CT-BERT with DBLOSS	0.71	0.70

Table 4
CT-BERT Model performance

This shows a fairly respectable performance, considering the dataset consisted of many instances previously unseen by our model. The dataset and code for model training and setup of our experiment can be found in this repository.

6. Conclusion

This work further solidifies the effectiveness of pre-trained models like CT-BERT when combined with other techniques like Distribution Balanced Loss. Also, from a societal angle, the work further proves the extent to which such a seemingly exhaustive task of manually analyzing tweets can be automated with the amalgamation of cutting-edge A.I. systems, to further help the concerned authorities act and mitigate the concerns of the general public. Thus, yet again a deep understanding of society on a global scale can be done with the power of deep learning and NLP, proving that AI can be a fast and impactful solution to many of today's problems.

6.1. Future Work

The next step would be to make an AI system with the model built here running its background. Also, the model's performance can be further improved with the classification being done in two levels that is by first segregating the 'none' labeled tweets by classification on the first level and then classifying the rest of the tweets on the next level.

References

[1] N. E. MacDonald, et al., Vaccine hesitancy: Definition, scope and determinants, Vaccine 33 (2015) 4161–4164.
[2] D. S. Courtney, A.-M. Bliuc, Antecedents of vaccine hesitancy in weird and east asian contexts, Frontiers in psychology 12 (2021) 747721.

---

[2]This dataset has been provided by FIRE [5] as part of model evaluation for each team.

[3] E. Robertson, K. S. Reeve, C. L. Niedzwiedz, J. Moore, M. Blake, M. Green, S. V. Katikireddi, M. J. Benzeval, Predictors of covid-19 vaccine hesitancy in the uk household longitudinal study, Brain, Behavior, and Immunity 94 (2021) 41–50. URL: https://www.sciencedirect.com/science/article/pii/S0889159121001100. doi:https://doi.org/10.1016/j.bbi.2021.03.008.

[4] S. Poddar, A. M. Samad, R. Mukherjee, N. Ganguly, S. Ghosh, Caves: A dataset to facilitate explainable classification and summarization of concerns towards covid vaccines, in: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022, pp. 3154–3164.

[5] S. Poddar, M. Basu, K. Ghosh, S. Ghosh, Overview of the fire 2023 track:artificial intelligence on social media (aisome), in: Proceedings of the 15th Annual Meeting of the Forum for Information Retrieval Evaluation, 2023.

[6] M. Müller, M. Salathé, P. E. Kummervold, Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter, Frontiers in Artificial Intelligence 6 (2023). URL: https://www.frontiersin.org/articles/10.3389/frai.2023.1023281. doi:10.3389/frai.2023.1023281.

[7] T. Wu, Q. Huang, Z. Liu, Y. Wang, D. Lin, Distribution-balanced loss for multi-label classification in long-tailed datasets, 2021. arXiv:2007.09654.

[8] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2019. arXiv:1810.04805.

[9] R. C. Staudemeyer, E. R. Morris, Understanding lstm – a tutorial into long short-term memory recurrent neural networks, 2019. arXiv:1909.09586.