Overview of the FIRE 2023 Track: Artificial Intelligence on Social Media (AISoMe)

Soham Poddar[1], Moumita Basu[2], Kripabandhu Ghosh[3] and Saptarshi Ghosh[1]

[1]Indian Institute of Technology, Kharagpur.

[2]Amity University, Kolkata.

[3]Indian Institute of Science Education and Research, Kolkata.

Abstract
The COVID-19 pandemic showed the importance of vaccination at a large scale. However, quite often people expressed different concerns they had towards vaccines which made them hesitant to take them. Some people were concerned about the potential side-effects of vaccines, while some believed that the vaccines were not necessary due to the disease being mild. These concerns were frequently shared on social media sites such as Twitter. The FIRE 2023 AISoMe track focused on identifying these specific concern(s) that people have towards vaccines from tweets, as a 12-class multi-label classification task.

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

1. Introduction

Social media sites are rich sources of real-time information about people's opinions on various topics. The Artificial Intelligence on Social Media (AISoMe) track aims to provide datasets and shared tasks for development of AI techniques (particularly, Machine Learning and NLP techniques) for utilizing social media data for diverse practical applications.

The AISoMe 2023 track focused on a social media classification problem in the healthcare domain, which is as follows. During pandemics such as COVID-19 where complete vaccination is the primary long-term solution to fight against the disease, social media can be utilized to understand public sentiments towards vaccines [1, 2]. In particular, many people are skeptical/hesitant about the use of vaccines owing to various reasons, including the politics involved, the potential side-effects of vaccines, and the fact that vaccines have been rushed into production. We identified 11 such specific reasons (concerns about vaccines) in our prior work [3], which are listed in Table 1 along with their descriptions. Examples of tweets from each of these classes have been given in Table 2. It is important to understand the specific concerns people have towards vaccines, so that their concerns can be addressed. The AISoMe 2023 track focused on this task of labeling (classifying) a tweet with one or more of these concerns against vaccines. This is important since a person unwilling to take vaccines due to the side-effects of vaccines

Conspiracy	Deeper Conspiracy – The tweet suggests some deeper conspiracy, and not just that the Big Pharma want to make money (e.g., vaccines are being used to track people, COVID is a hoax).
Country	Country of origin – The tweet is against some vaccine because of the country where it was developed/manufactured.
Ineffective	Vaccine is ineffective – The tweet expresses concerns that the vaccines are not effective enough and are useless.
Ingredients	Vaccine Ingredients/technology – The tweet expresses concerns about the ingredients present in the vaccines (eg. fetal cells, chemicals) or the technology used (e.g., mRNA vaccines can change your DNA).
Mandatory	Against mandatory vaccination – The tweet suggests that vaccines should not be made mandatory.
Pharma	Against Big Pharma – The tweet indicates that the Big Pharmaceutical companies are just trying to earn money, or is against such companies in general because of their history.
Political	Political side of vaccines – The tweet expresses concerns that the governments/politicians are pushing their own agenda though the vaccines.
Religious	The tweet opposes vaccines due to religious reasons.
Rushed	Untested/Rushed Process – The tweet expresses concerns that the vaccines have not been tested properly or that the published data is not accurate.
Side-effect	Side Effects/Deaths – The tweet expresses concerns about the side effects of the vaccines, including deaths.
Unnecessary	The tweet indicates vaccines are unnecessary, or that alternate cures are better.
None	No specific reason stated in the tweet, or some reason other than the given ones.

Table 1
The different classes/labels (concerns or objections towards vaccines) in the CAVES dataset [3] along with their descriptions.

needs different persuasion and reasoning than someone who is hesitant to take vaccines due to the corruption in politics.

2. The Datasets and Evaluation Metrics

This section describes the train and test datasets used for the track, and also describes the metrics used for evaluating the submitted runs/methods over the test dataset.

2.1. The training / validation dataset

For training and validation, we utilize the 'CAVES' dataset from our prior work [3]. This dataset contains 9,921 anti-vaccine tweets about COVID vaccines (that were posted during 2020-21), where each tweet has been labelled with one or more of the 12 classes (given in Table 1) by human annotators. Table 2 shows some examples of tweets from this dataset, along with their labels. More details about the data collection and annotation process of the CAVES dataset can be found in the prior work [3].

Tweet Excerpt	Labels
STOP TAKING *TOXIC VAX* and *expose COVID hoax* and murders with morphine and ventillators. *there is No covid*!	ingredients, conspiracy, unnecessary
Please *don't push vaccine on us make it voluntary*. We don't trust anything to do with *Bill Gates* pushing their agenda of *vaccine chips*!!	pharma, mandatory, ingredients
The reason insurance companies won't pay out if you experience the inevitable *adverse reactions, including death* is because it is an *"Experimental Vaccine"*	side-effect, rushed
Would you want the *Russian vaccine*? If not, you shouldn't want one that's been *pushed through for political reasons* either.	political, country
Catholic leaders are advising Catholics that the COVID-19 vaccine from Johnson & Johnson is "morally compromised"	religious
I'm NOT taking your damn vaccine. Keep your conspiracy out of my veins!	none

Table 2
Examples of tweets with their labels and explanations, from the CAVES dataset. The explanations for different labels are highlighted in *italics*.

2.2. The evaluation dataset

For evaluation, we introduce a new dataset, developed in a similar fashion as the CAVES dataset. This dataset contains 486 tweets labelled into the same 12 classes. However, these tweets are *not* only about COVID vaccines but also about other types of vaccines (e.g., MMR vaccine, Flu vaccine), from both the COVID-era as well as pre-COVID times.

2.3. Evaluation method

The participating teams were asked to develop models for the multi-label classification task, which were trained on the CAVES dataset and whose performance will be measured over the evaluation dataset described above. Each participating team were able to submit up to 3 runs, e.g, from models with different hyperparameters. They were also free to use other attributes of the tweets (apart from the text) if they wanted, along with other publicly available datasets for training their models.

The submitted runs by the participants were ranked based on their performances on the evaluation dataset. The standard classification metric of **Macro-F1 score** on the 12 different classes was used for evaluation.

3. Methods - Submitted runs

In the AISoMe track, 22 teams participated this year, and as many as 48 runs were submitted. Most of the teams used NLP pre-processing techniques and a few teams used TF-IDF Vectorizer to extract features. Among the classification techniques, fine-tuned transformer models such as BERT, RoBERTa and Covid-Twitter-BERT (CT-BERT) [4] are utilized mostly by the participating teams. Some teams also employed LLM-based models such as GPT 3.5 and GPT2LMheadmodel. Neural network-based classifiers (MLP) and traditional classifiers (such as Multinomial Naïve

Team Id	Overview of method	Macro-F1
AKCSIT	Fine tuned CT-BERT	0.71
Datawlz	Fine tuned CT-BERT	0.71
IISERBPR-NLP	Fine tuned BERT with best threshold	0.70
DSIRC	Fine tuned CT-BERT	0.67
Cognitive Coders	DeBERTa Large Fine-tuned	0.67
TextTitans	BERT-large uncased	0.66
PICT CL LAB Group 1	RoBERTa based model	0.65
SSN_IT_Team01	RoBERTa based model	0.65
LLM-geeks	Intersection of the predictions from DeBERTa and RoBERTa	0.63
SSN_IT_Team02	RoBERTa based model	0.57
Data Warriors	LLM based model (GPT 3.5)	0.55
Alpha Intellect AI	BERT based uncased	0.54
S3 Endeavour	GPT2LMheadmodel	0.46
PICT CL Lab	Support Vector Machine (SVM) model	0.45
ZSL	Decision Tree Classifier + Multi Output Classifier	0.43
C3	RoBERTa based sentence classification	0.41
APS AI&ML	Multinomial Naive Bayes, Multi-Output Classifier	0.39
Social Media Data Analysis Team	Classifier chain with Support Vector Machines	0.38
OpenVax	Multi-Layer Perceptron (MLP) model	0.37
RANJAN A-MONKA-RESEARCH	CNN-BiLSTM model with GLOVE embeddings	0.29
Swastik Anupam	TFIDF-Neural Net	0.25
IIIT_SURAT	SVM models within the Classifier Chain	0.07

Table 3
Comparison among some of the submitted runs in the classification task. Runs are ranked in decreasing order of Macro F1-score. We are reporting only the best-performing run of each team

Bayes and Support Vector Machines, Multi-Output Classifiers) are also used by some of the teams. The summary of the techniques is reported in Table 3. It is observed than fine-tuned CT-BERT models have outperformed all traditional and other neural classifiers for our task.

4. Conclusion and Future Directions

The FIRE 2023 AISoMe track compared the performance of various methods for identifying the specific anti-vaccine concerns from tweets. We hope that the test collections developed in this track will be utilized by the research community in the development of better models for this important task in future. It can be noted that the CAVES dataset also contains explanations for the class labels, as well as summaries for the different anti-vaccine classes (details in [3]). These data can also be utilized for tasks such as explainable tweet classification and tweet summarization in future.

Acknowledgments

References

[1] S. Poddar, M. Mondal, J. Misra, N. Ganguly, S. Ghosh, Winds of change: Impact of covid-19 on vaccine-related opinions of twitter users, in: Proceedings of the International AAAI Conference on Web and Social Media, volume 16, AAAI Press, 2022, pp. 782–793.

[2] L.-A. Cotfas, C. Delcea, I. Roxin, C. Ioanăş, D. S. Gherai, F. Tajariol, The longest month: analyzing covid-19 vaccination opinions dynamics from tweets in the month following the first vaccine announcement, Ieee Access 9 (2021) 33203–33223.

[3] 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 International ACM SIGIR Conference, 2022.

[4] M. Müller, M. Salathé, P. E. Kummervold, Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter, arXiv preprint arXiv:2005.07503 (2020).