Why are people hesitant to the COVID-19 Vaccine?

University of Sheffield

Charlie Grimshaw - Undergraduate Computer Science Student
Supervisor: Xingyi Song

Introduction

It is important to understand the reasons behind the public's hesitation towards COVID-19 vaccines. To the best of our knowledge, we are the first to consider vaccine hesitancy as one of the stance categories towards the COVID-19 vaccination. In this study, we create a new publicly available dataset of tweets related to the COVID-19 Vaccine, and train language models specifically to the domain and task of COVID-19 attitude tasks as well as an implementation of XCANTM, where we consider both text classification and topic modeling as well as using added VAE architecture to discover hidden topics of target categories.

Tweet Annotation Process

We used efficient data annotation strategies that use less time and human resources but achieve similar performance as the standard triple-annotation strategy.

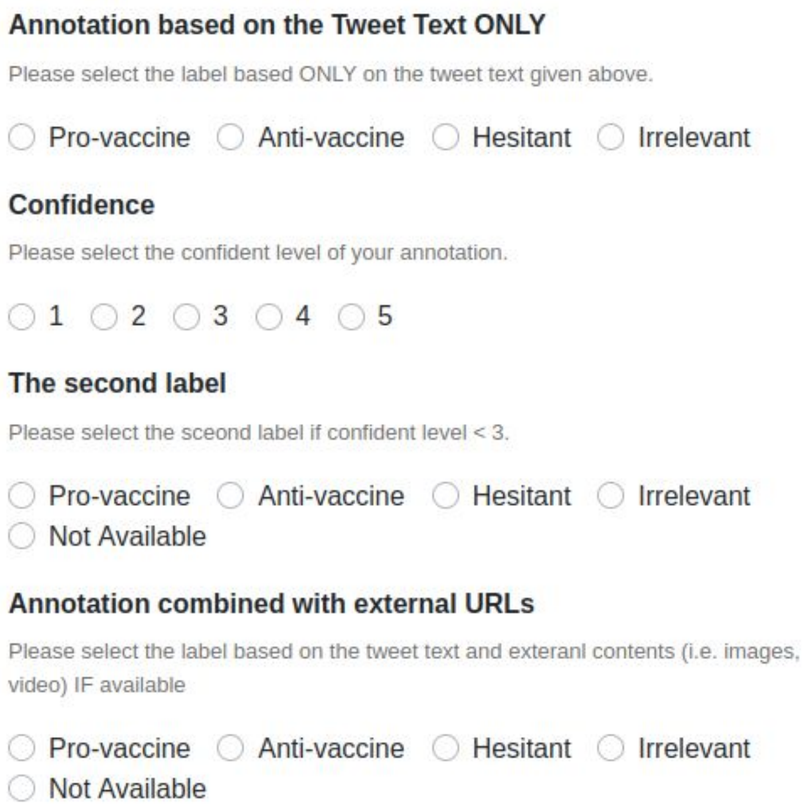

Figure 1: GATE Teamware

https://gate.ac.uk/teamware/

The participants were trained on using the GATE Teamware annotation system as shown, where they are asked to annotate each tweet.

Novel Dataset

Our new dataset is novel compared to previously published datasets as it covers a time span from the announcement of the first COVID-19 vaccine to the universal access to the booster vaccines, uses "hesitant" as a category of stance, and uses 1 or 2 annotators per tweet, in comparison to the typical 3.

Task Description

We assume a training set of n posts $T = \{(x_1, y_1), ..., (x_n, y_n)\}$ where x_i is a single tweet related to COVID-19 vaccination. $y_i \in$ {Pro-Vaxx, Anti-Vaxx, Hesitant, Irrelevant} is an associated post label. Given T, we trained an interpretable supervised neural network f that: 1. maps a new post j into one of the four categories $\hat{y} = f(x_j)$ and 2. yields faithful explanations (i.e., rationales).

Results

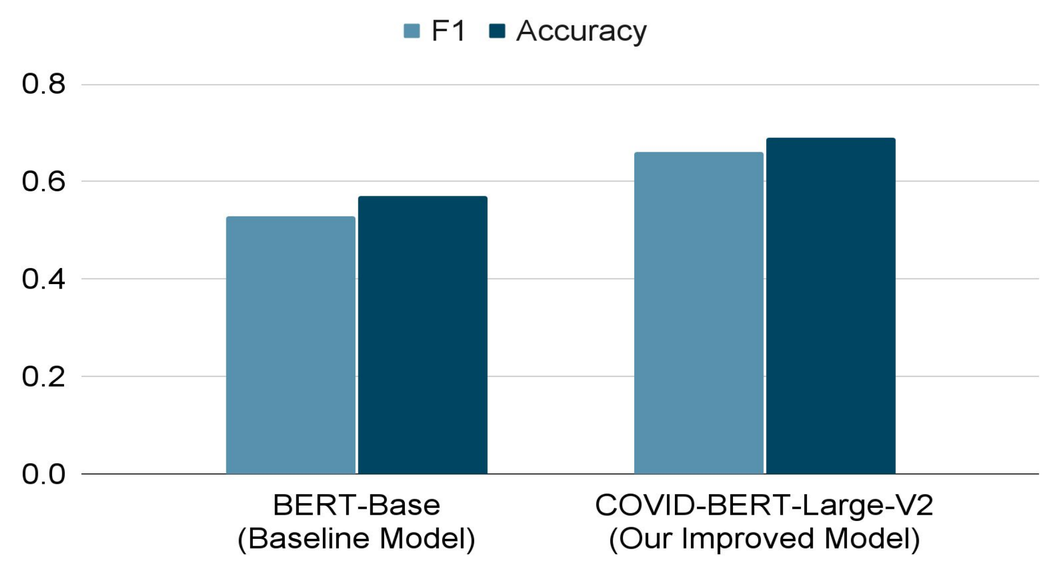

Figure 2: Results of Model Performance

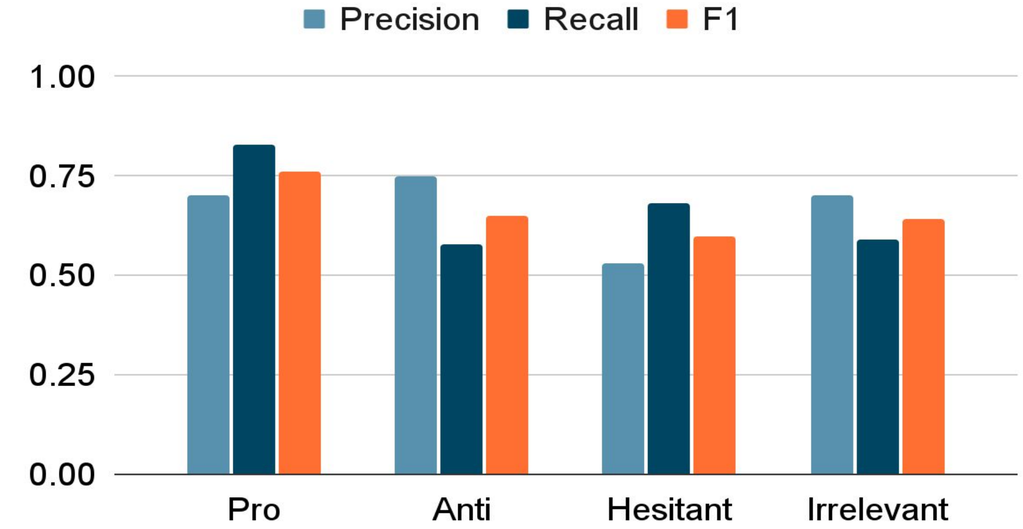

Figure 3: Results from our model, split by class

Top class regularized topics:

pro: [vaccination effects people effective covid prevent pandemic children government development]

anti: [high mandatory amp feel class anti booster information absolutely benefits]

hesitant: [vaccines effects hesitancy day allergies contracting kids masks stop covid]

irrelevant: [people feel children age conversations fear coronavirus double protective corona]

Conclusions

The results conclude that domain-adaptive pretraining and task-adaptive pretraining greatly improve the capabilities of our language model, as well as by using VAE architecture, we can clearly see the top class topics that belong to the categories. We believe the new dataset will benefit future research into this task as well as helping with the overall fight against disinformation.

References

BERT : Kenton, J.D.M.W.C. and Toutanova, L.K., 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT* (pp. 4171-4186).
COVID-BERT: Müller, M., Salathé, M. and Kummervold, P.E., 2020. Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter. *arXiv preprint arXiv:2005.07503*.

Acknowledgements

Thank you to Xingyi Song, Yida Mu and the rest of the Natural Language Processing Research Group at the University of Sheffield for mentoring me through this study, as well as giving me the opportunity to participate in my first research project.