CSE 574: ML Assignment 3 Report

1.2.1 (2 points) **- Based on what we discussed in class about this dataset and the task of stance annotation (here, for attitudes towards vaccines), on what percent of the tweets that you annotate do you think the two annotators will agree? There is no right answer here, obviously, but provide a justification for your response.**

40%. We are sure there might be some tweets that might look different from different perspectives within our team. We cannot really guess what the other annotator will be thinking while labeling the tweets.

1.2.2 (3 points) **- Based on what we discussed in class about this dataset, what percent of the tweets that you annotate do you think will be labeled *pro-vaccine*? Justify your answer.**

We think that there might be 60% of the tweets that will be labeled as pro-vaccine. Since most of the people are vaccinated right now, we think that it would reflect in a random sample of the tweets.

2.1 **- What was your overall percent agreement, and how did this compare to your estimate from 1.2.1? Why were these similar/different?**

The percent agreement that we got is ~47% which was surprisingly close to our estimate in 1.2.1. However, 1.2.1 was mostly an educated guess of the diverse perspectives in our team and how ambiguous the tweets would turn out to be, and turns out it was a good guess!

2.2 **- Was this annotation task harder or easier than you expected? Why, and what was the hardest part?**

Yes, the annotation task was much harder than we expected. Some tweets were not so straightforward and required extra context to gauge their stance. Differentiating between a neutral tweet and a tweet that gives us no information to detect the stance was also very ambiguous at times. Some tweets were in different languages, and we needed to translate them first, and then only we were able to annotate them.

2.3 **- Does this change your perspective on how people talk about health-related content on social media? Why or why not?**

Our initial perspective was that we'd see a lot of opinionated views and maybe even some misinformation in the tweets. After annotating the tweets, it's safe to say that our opinion hasn't changed much.

2.4 **- What did you learn from this annotation task that will change your perceptions of how NLP models are trained moving forward?**

We understand now that NLP models can have a training sample labeled in multiple separate ways depending on the background and context of the person labeling them. Also, depending on how the disagreements are dealt with (whether they are factored in the training or if a collective agreement is reached somehow) can have a significant impact on the prediction of the model.

2.5 **- How do you think your team's agreement (say, in terms of Krippendorf's Alpha) will compare to the other teams? Why do you think so?**

Assuming everyone got a set of tweets that had equally ambiguous tweets we assume everyone's Krippendorf's Alpha will be in the same ballpark (say +/-10%). However, the other team's agreement might also depend on their strategy to decide which tweet falls in the neutral category and which did not have enough information to detect the stance. Our strategy was to annotate those tweets with an X where an information (news headline) was being shared and annotating other tweets with a neutral stance only when we could clearly see that the person did not care either way.