Programming Assignment #3 - Manual Data Annotation

Team: svm
Shreya Anil Vinchankar - 50420008
Vaibhavi Rajendra Vichare - 50367860
Mahesh Bhosale - 50418912

Part 1: Performing Manual Annotation (60 points)

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 of your response.**
⇒
We would say around 80%, we hope that we would do at least better than random guessing. So, each one of us is to assign the label randomly, for a particular tweet we would have 2/8 (25%) chance of all agreeing, which is good. But if we use the majority voting, then we would have 6/8 (75%) chance of agreement for a particular annotator. Therefore, we expect the lower limit is 75%. Now, for the fact that all the annotators are from India (most populous country and one of the worst affected by COVID) and being international students (the importance of vaccination in international travel), we would expect that there would be miser leeway in excuses of not taking a vaccine shot or anything against it. So the general view/glance of a tweet for all annotators would be similar if not exact.

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.**
⇒
Again, as explained in the earlier question, there would be miser leeway in excuses of not taking a vaccine shot or anything against it. So the general view/glance of looking at a tweet for all annotators would lean towards sanity of vaccines. And hopefully people on twitter are well educated and well informed and would tweet about supporting the vaccines. So we expect pro-vaccine tweets to be around 70-80%.
Note: We have assumed that tweets were not selectively chosen (pro or anti or neutral towards vaccines), and it really represents the twitter population.

Part 3 : Reflecting (20 points)

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 overall percent agreement is 77.33.

There is a difference in the estimates from 1.2.1 to actual percentage agreement. The disagreement could be because of reasons like some of the tweets were not opinions but actually the news or information on covid which made it difficult to categorize the tweet as for or against covid-19 vaccines.

For Example:

"With inflation looming and US bond yields rising sharply, I am topping up on the cyclicals and a few stocks that could benefit from vaccine rollouts and economies reopening. Cyclicals imply value."

This tweet talks about other profits because of vaccine rollouts like stock benefits and not actually about the benefits of vaccines. This one can be misinterpreted as if it were in favor of vaccines.

Another one could be how a person interprets a tweet with minimum information on the tone of the user.

.

For Example:

"** NO Mask !!!
** NO Vaccine !!!
** NO Control !!!
** Deep State Out !!!"

This tweet fails in determining the tone of the user. It can either be interpreted as govt has no control over the vaccines and mask and at the same time, it can also be seen as if the user does not want the any control over vaccines and mask and is annoyed on the same

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

⇒ Yes, it was harder than expected. In language, the interpretation of the tweet can be made in numerous ways based on the past experience and demographics of the annotator. So it's really not deterministic and opinion could be easily turned around with counter arguments which would equally make sense.

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

⇒ Yes, we have seen that people make lots of fun about some of the serious health related issues. To some extent, this is fine but it's not ok if it is spreading misinformation although in a funny way.

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

⇒

We learned that everyone has a different approach to interpreting tweets. As the given dataset contained limited information about tweets, because of that it becomes difficult to reach to the final conclusion. It would be better if the NLP model trained on data which contains more information.

For Example:

In a given dataset, if we had more information about tweets like, sentiment of that particular tweet, then it would be easier for us to categorize that tweet into a particular label.

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?**

⇒

We think that Krippendorf's alpha and average percentage agreement would be similar (say +-10%) as compared to our values, because the demographics of annotators would be similar. For example, it would be harder to find a science student against the vaccine, so everyone's interpretation would not change from each other. But, some sarcastic tweets can be interpreted differently, but there were not many such tweets.