**Charles Ryan, Research Proposal**

**What research question do you want to answer? Be specific!**

I am interested in how respondents to posts on the sub-reddit r/AmItheAsshole classify posts as “Asshole” or “Not the A-hole.” More specifically, I am interested in identifying common themes throughout responses on the sub-reddit to predict whether given posts will be classified as “Asshole” or “Not the A-hole.” Research conducted by Elle O’Brien on a dataset of 111k posts determined a “slight anti-parent trend, with Redditors being more likely to absolve authors who mention a mom or dad” (dvc.org). Additionally, posts that only mentioned a “female romantic partner “were associated with greater likelihood of being voted the Asshole” (dvc.org). I find this research intriguing and would like to expand on it by identifying the types of interactions that are more likely to result in an “Asshole” or “Not the A-hole” classification. Research question: what types of interactions result in an asshole classification?

**What are two possible answers?**

When I say types of interactions, I am thinking of declining requests, asking requests, letting people do something, etc. One possible answer is that posts that describe demands or requests of people are more likely to be classified as “A.” This hypothesis is based on the fact that roughly 70% of the sub-reddit’s participants are under 34 years old, a cohort that is known to be more timid. Another possible answer is that what matters is not the type of interaction but rather the type of interaction in relation to a type of person. For example, perhaps denying requests of spouses or parents is more likely to result in an asshole classification than denying requests of friends.

**What are the documents of text data you will analyze? How will you get these documents?**

The documents of text I will analyze are posts and comments on the sub-reddit r/AmItheAsshole. Elle O’Brien, whose research I cite above, has already put together a dataset of 111k posts, which will be valuable for understanding the structure of this data. I am interested in the comments on specific posts and what types of interactions as defined by the commenters lead to an “Asshole” classification. I will download the post ids and scores using upshift.io API and will use the Reddit API to download the content/meta-data of the posts and comments. I will also only be looking at the posts that the moderators have determined to be the best (those that have the most disagreement), which are posted to the sub-reddit r/AITAFiltered. I suspect that the filtered AITA sub-reddit will have riper comments and will provide a better data source for how people evaluate and judge the posts.

**What methods will you use to answer this question?**

First, I would clean the data so that the dataset only includes comments where an “A” or “NTA” classification is made and for which no text is redacted. A key hurdle will be determining whether there are comments on posts that classify the interaction described in the main posts. To make this determination, I will look at word frequency in the comment sections to see what classifying words are most common. Then I will use these words to create a classifier for posts to the sub-reddit based on whether they refer to these interactions. This could be done with “logistic regression using 1-gram frequencies,” which is what Elle O’Brien does in their research (dvc.org). Depending on how this goes, this will serve as the baseline estimate that other strategies for predicting the classification of posts should surpass. I will also use a chi-squared test to determine what types of interaction-classifying-words are determinative of the overall classification to a given post. Ideally, this will reveal whether a post that refers to a request is more likely to be classified as an “A” or an “NTA.”

Sources:

Imgur. “r/AmItheAsshole 2019 Subscriber Survey.” *Imgur*, imgur.com/a/POhgZsh.

Iterative. “Iterative/aita\_dataset.” *GitHub*, github.com/iterative/aita\_dataset.

O'Brien, Elle. “AITA for Making This? A Public Dataset of Reddit Posts about Moral Dilemmas.” *DVC*, dvc.org/blog/a-public-reddit-dataset.

“r/AITAFiltered - AM I THE ASSHOLE? FILTERED.” *Reddit*, www.reddit.com/r/AITAFiltered/comments/ehsc9a/am\_i\_the\_asshole\_filtered/.

“r/AmItheAsshole - 2019 Subscriber Survey Data Dump!” *Reddit*, www.reddit.com/r/AmItheAsshole/comments/dcae07/2019\_subscriber\_survey\_data\_dump/.