James Yang

DATA 512A

10/9/22

<div align="center">Reading Reflection 2</div>

1. **In 2-5 full, grammatically correct sentences, answer the question "How does this reading inform your understanding of human centered data science?"**
   In the reading "Mixed Messages? The Limits of Automated Social Media Content Analysis", the paper emphasizes how there is a fundamental gap between NLP research and policy makers. This idea reinforces my thoughts regarding lack of knowledge from senators when addressing possibility of regulation with companies such as Facebook and Google. NLP and social media regulation is just one of the issues that human centered data science faces when integrating itself into the real world. The article brings up important limitations that NLP models face; these tools need to be trained on a clean subset of data, and when it is skewed it creates bias. This furthers my understanding of human centered data science because there is ultimately no such thing as a clean subset of data, thus bias will be formed. However, there is an equilibrium of bias that needs to be trained to create the ideal set of data.

2. **In a list of bullet points, no more than one sentence each state:**
   - **What was the question that the author tried to answer or raise as important?**
     - The paper was created to explain the capabilities and limitations of text classifiers for social media posts and other online content.
   - **What was the method used to address the question?**
     - Describing different attributes of NLP tools and their capabilities. After describing these issues and concerns, give recommendations to both policymakers and researchers on how to effectively utilize these tools in developing classifiers, censorship in practice, etc.
   - **What was the primary or most important point of the reading; the finding?**
     - The most important section is the recommendations for policymakers and researchers. NLP tools are designed to identify text attributes but can be relatively ineffective and sometimes inconclusive which means that it shouldn't be mandated in law. Researchers need to develop more accurate models using more effective analysis in multiple languages outside of English. These models will hopefully become accurate enough so that one day policymakers can integrate them into law.

3. **Using full sentences, list at least 1 question that this reading raised in your mind, and say why the reading caused you to ask this question.**
   The section depicting the NLP processing tool amplifying social bias makes me question the stereotypes that were created. The reason is because NLP models are not inherently bias, it's the data that it trains from that makes it bias. However, we will never create a perfect model because there is never going to be a perfect set of data to train upon. Humans are not perfect in what they say, and that's how almost every model is going to

output as well. Do policymakers believe that NLP can substitute as virtual police in a sense?

4. **In weeks where you have two required readings, you should explain in 2-5 sentences how the two readings are related. Sometimes the readings allow you to draw similar conclusions; other times the readings may present contrasting ideas.**

In the article revolving around AI's Islamophobia problem, it correlates with the highlighted issue in the conference paper regarding flaws found in real world NLP text models. It's a direct result of NLP generating bias to create controversial results. This model represents the training data having terrible outputs for something as innocent as "Two muslims walked into a…". The idea of garbage in, garbage out creates a very real situation of what these models get developed on. The solutions that they address in the AI article are the very same solutions that the paper addresses where these models cannot be trusted until the bias is addressed, and until these issues get addressed there is simply no way they can be integrated into a robust law enforcer or guideline of any kind.