Trinh Pham
Yadira Valenzuela

Lab 13: Responsible AI

#1. **a**) "*Amazon scraps secret AI recruiting tool bias against women*"

There was an apparent testing bias during AI training, as the majority of tech applicants were men. This resulted in the machine learning algorithm filtering out women candidates. The obvious fix for this matter is to have a diverse range of resumes to test and train. Given that the male-to-female ratio of applicants is uneven, we would advise them to do a 50/50 testing and training ratio. Decrease the number of male applicants and increase female applicants.

Bias detection algorithms could greatly reduce the AI's ability to detect certain words such as "women". The team can also remove unjustifiable weights in the model.

b) "*Cleaning Up ChatGPT Takes Heavy Toll on Human Workers*"

We think that they need to provide counseling for the workers. And now that the system is trained maybe create an AI that monitors the AI. Programs like ChatGPT are going to use NLP to learn from information online and it will be exposed to vulgar and violent things. An AI that monitors ChatGPTS output could be used to prevent workers from being traumatized.

c) ChatGPT should be held accountable for using articles without disbursing funds to these authors. ChatGPT makes money off of other author';s work and should be accounted for when using their articles for their program. For testing purposes, they use over millions of articles, so disbursement can depend on how the company wants to distribute amongst each author. The copyright laws can aid in this situation as well. Based on what the copyright law says, AI must follow protocol.

#2.

- We would add predictive analysis coding statements to our code. They can be used to determine the probability of an outcome for a situation and to calculate the path a car or objects are going to take. This would be useful in emergencies (linear regression and decision trees).
- Another would be a conditional statement that calls for a human to make the decision. AI is very "smart" but there are times when a human's instincts can make a better decision.

Trinh Pham

Yadira Valenzuela

We would also add a condition in which the AI takes over again if the person is unable to.

- We would also use DL models to train the AI to make decisions in different emergency situations.