Homework 5: Ethics and Applications
Computer Vision, Spring 2022
Due: April 27th at midnight EST

Each questions describes a hypothetical situation, which you might face as you start your career. Please answer each question with a short paragraph (3-4 sentences).

Each question is worth 25 points.

For this homework, you are *not* allowed to work in groups.

No late submissions will be accepted.

Q1) After learning about neural networks at Columbia, Homer landed his dream summer internship at a self-driving car startup in New York City. Although the prototype autonomous car was pretty good, there was a major bug that was preventing launch: the car would often crash into plants! So, Homer's job for the summer was to train a neural network to detect plants, which would feed into the car's control algorithm. To train the neural network, Homer decided to collect training data by walking around Central Park, and taking pictures of different plants that summer. *What is wrong with Homer's plan and why?*

Q2) After graduating from Columbia, Lisa founded her own startup. Her idea was to create an autonomous drone that would fly around Manhattan, and sell bagels to hungry pedestrians. However, Lisa wanted to maximize her profits, so she wanted to find a way to send the drones to areas where people like bagels. She decided to train a neural network that would input GPS coordinates and time of day, and output the expected profit for sending the drone there. So, she uses the neural network to estimate the best location, flies the drones there, and sells bagels. She then uses the actual sales data to train the neural network, and repeats this each day. *What is wrong with Lisa's plan and why?*

Q3) Marge thinks machine learning is just hype, so she accepts a job at a startup that manufactures smart pet products. Marge's job is to create the smart cat door, which is a door that will only open when the owner's cat shows up, and no other animal. Marge decides she can put a camera on the cat door, and create a cat-face recognition system using a variety of classical computer vision techniques. Moreover, since there is no training set, the training set cannot possibly be biased. To help her debug, she takes a few photos of her own cat to use as a test set, and refines her system until it works very well on the test set. *What is wrong with Marge's plan and why?*

Q4) Bart landed a job at a hospital after graduating, and he wanted to put his Columbia computer science degree to good use. All the doctors kept commenting to him that, if cancer can be detected early, then the outlooks for patients is much better. So, he decided to create a phone app that would screen people for cancer automatically, which he would make available for free. Since

he was at the hospital, he was able to collect a massive dataset to train a neural network that would input a medical scan and output the probability that a person is at risk for cancer. He was aware that neural networks often learn biases, so he was careful to test them for it. He empirically tested his model in many ways, for example showing that his model was not biased against race or gender. However, he found that his model was biased for age, and the model was more likely to predict older folks are at higher risk of cancer. *Why might Bart's plan be ok?*