Homework 5

Ethics and Applications

COMS W4732

Computer Vision 2, Spring 2022

Chandan Suri

UNI: CS4090

Submitted on:

April 27, 2022

Question 1

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?

Answer:
Homer's plan to take pictures of plants in the Central Park in the summer to introduce into the car's control algorithm has a lot of drawbacks:

1. Firstly, just taking the pictures of the plants in Central Park wouldn't be a comprehensive data-set and wouldn't cover the whole range of green plants which could be problematic for the autonomous driving algorithm. Furthermore, considering just those plants would also introduce a bias in the algorithm towards the plants for which Homer had taken the pictures.

2. Secondly, as it's the summer season, there are a lot of plants that don't even bloom in the summer season but are mostly seen during the spring season or the winter season. This would introduce an additional bias in the system towards only detecting those types of plants that can be seen during the summers. Also, in the winter season, a lot of plants get discolored and are not all green. So, this would introduce a bias in the model towards just detecting the green plants during the summers.

3. Furthermore, in the winter season, the plants might be covered with snow which the model wouldn't be trained on as the pictures are only taken during the summer season. This parochial view of the plants might create a bias in the model to only detect the plants with green leaves and without any snow. This could be very hazardous if the car doesn't detect this.

4. Lastly, if Homer is just taking the pictures during the daytime, then the summers are a lot sunny, and the illumination patterns would differ in the images in comparison to the winter season. This kind of bias towards a particular kind of lighting and illumination can be very problematic for the autonomous car to detect the images where the illumination is different (maybe quite less).

Question 2

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?

Answer:
Lisa's plan to send drones to the locations in Manhattan after inputting the GPS coordinates and the time of the day to predict the expected profit might be fallible due to the following reasons:

1. Lisa's plan completely relies on the features like GPS locations and the time of the day to predict the profits made. This is quite a constrained assumption as there might be other factors that affect the sales of the bagels. For instance, there might be some big event in a location at some point on that day which might have affected the sales of the bagels. Also, the profit made by selling the bagels highly relies on the pedestrians which might vary according to the neighborhood and the type of community nearby. These factors aren't being taken into consideration which might introduce a bias in the model which only depends on the location coordinates and the time of the day. Lisa should consider other factors too.

2. As Lisa initially sends the drones to some locations and then retrains the model based on the sales data for those same places. This procedure might be very problematic because the model has a constrained dataset involving some places but not all in Manhattan. After retraining from those places, the drone will only go to those places without looking into some new places. This will create a bias such that the drones are going to go only to those places without considering some new options.

3. Furthermore, if the model is retrained again and again on the historical data, this might overfit the model such that the model would predict only those places because of the retraining methodology.

4. Lastly, drones might pose a threat to the pedestrians in the neighborhood. Freely moving drones in the public spaces might be problematic for the movement and might cause disturbance and various problems to the people commuting in those locations. Also, due to any malfunction, the drones might crash into people or places damaging property and posing life-threatening hazards to humans. As can be seen in this paper, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7206421/, we know that the drones might cause various issues and should be considered very carefully for any use case.

Question 3

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?

Answer:
Marge's plan to the test the cat-face recognition system based on the photos taken by the camera using her own cat's photos is certainly fallible because of the following reasons:

1. Firstly, as Marge is doing all the testing just on the basis of the photos taken of her own cat, we cannot rely on the test results of the model. Primarily, this is because the test dataset is very conservative here with only images of one cat. Doing testing on such a restricted dataset wouldn't tell us much about how the model would perform if a new type of cat or a cat with some varying features is introduced. This might fail the model and cannot be considered to be a valid test dataset for comprehensively testing the underlying model. So, testing the model with such a restricted dataset introduces a bias in the testing phase of the model such that the model is biased towards Marge's cat such that it's going to detect Marge's cat quite accurately which might not be the case for other cats.

2. Additionally, as the testing is being done just on the photos of the cats, it might be valid to assume that the training/development was also done using the photos of the cats. If this is the case, then anyone with the photo of the owner's cat might be able to bypass the security of the cat door which is a big security concern for the application. This isn't taken care of from the explanation that we have been given about the application and could pose a big threat as the security concerns of the application aren't at par with what they should be.

Question 4

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?

Answer:
Bart's model is said to be unbiased considering a lot of factors but is said to be biased with age.

1. Generally, it's known that the risk of cancer is associated with older people more rather than the younger ones. This is an actual bias that can be found in society and cannot be considered a negative impacting bias. As this kind of bias doesn't pose any kind of hazard and is actually mirroring a correlation actually found in nature, the model can be said to work perfectly even with the bias towards the age such that the model predicts the risk of cancer with a higher probability for the older folks.

2. Furthermore, predicting cancer early could be deemed as a high-risk scenario where the recall and, thus, the false negatives matter a lot. For instance, if a person doesn't have cancer and the model predicts that the person has cancer, that could be still fine as after various tests the person would know that he/she doesn't have cancer. However, if a person having cancer is predicted to not have cancer (false negative), then this might pose a great threat to the life of the involved person. So, even if Bart's model predicts some false negatives for the older folks, that can still be considered to be fine considering the high-risk scenario.

So, considering all the points above, it looks like Bart's model can be said to be working fine and might work pretty well in a real-life setting.