✓ **Congratulations! You passed!**
TO PASS 80% or higher

[Keep Learning]

GRADE
100%

Analyzing Bias

LATEST SUBMISSION GRADE

100%

---

1. One measure of fairness is demographic parity, which means that the positive percentage is the same for all groups. Suppose Ants are 60% of the population, and Bunnies are 40% of the population. Your model predicts a positive result for 10 Bunnies. For how many Ants should the model predict a positive result to satisfy demographic parity?

1 / 1 point

- ○ 8
- ○ Cannot be determined
- ○ 10
- ● 15

> ✓ **Correct**
> Correct! There are 50% more Ants than Bunnies, so there should be 50% more positive predictions for Ants than Bunnies according to demographic parity.

2. One measure of fairness is equality of odds, which means that, all else being equal, the model is equally likely to make a positive prediction for a member of either class. Suppose that being able to run fast is meaningful to the prediction, and other properties are not considered. Your model predicts a positive result for 3 of the 30 Cats which can't run fast, and 20 of the 40 Cats which can run fast. If there are 10 Dogs which can't run fast, and 20 which can run fast, how many positive predictions should there be for equality of odds?

1 / 1 point

- ○ 14
- ● 11
- ○ 3
- ○ 15

> ✓ **Correct**
> Correct! 1 of 10 Cats that can't run fast had a positive result, and 1 of 2 Cats that can run fast had a positive result. Thus, for equality of odds, 1/10 of the 10 Dogs that can't run fast should have a positive result, and 1/2 of the 20 Dogs that can run fast should have a positive result. This is a total of 11 Dogs.

3. Another measure of fairness is equal opportunity, which means that those members with a property should have equal probability of a positive prediction, but not necessarily those without that property. Suppose that the ability to fly is meaningful to the prediction, and other properties are not considered. Your model predicts a positive result for 10 of the 80 Elephants which can't fly, and 15 of the 20 Elephants which can fly. If there are 40 Flamingos which can't fly and 60 which can fly, which of these are valid numbers of positive predictions under equal opportunity:

1 / 1 point

- ● 40 flightless Flamingos, 45 flying Flamingos
- ○ 12 flightless Flamingos, 13 flying Flamingos
- ○ 10 flightless Flamingos, 20 flying Flamingos
- ○ 0 flightless Flamingos, 50 flying Flamingos

> ✓ **Correct**
> Correct! As 3 of 4 flying Elephants had a positive prediction, 3 of 4 flying Flamingos should have one too.

4. One measure of fairness is "fairness through unawareness" which suggests that a model is fair if the protected attribute differentiating Class A and Class B is not given to the model. Which of the following are you able to assume?

1 / 1 point

- ☐ The model will give roughly the same predictions for members of Class A and Class B.
- ☐ At least one of equal opportunity, equality of odds, or demographic parity must be satisfied.
- ☐ The model would change its predictions if explicitly given the protected attribute — i.e. the attribute is not implicit in the input data.
- ☑ None of the above.

> ✓ **Correct**
> Correct! You cannot assume any of the above. A model may use dimensions that vary with the protected attribute in order to produce results that differ between the two groups in many ways, including ways that are incompatible with any of the three listed definitions of fairness, or ways that implicitly extract the protected

attribute.

5. Suppose that your friend implemented a standard DCGAN on a dataset they compiled and manually labeled 1,000 generator outputs for some feature which they care about. Your friend then changes their generator to also output the nearest label in z-space (by Euclidean distance) whenever it generates an image. What are some potential sources of bias here?

1 / 1 point

☑ Since your friend single-handedly labeled the images, it is possibly they introduced some of their own bias to the model.

> ✓ **Correct**
> Correct! Lack of diversity in labelers can contribute to bias.

☑ Since the generated images are sampled images from a Gaussian distribution, the use of Euclidean distance is likely to bias the label output with respect to distance from the center of the distribution

> ✓ **Correct**
> Correct! Using mismatched distance metrics can lead to bias. For example, with a one-dimensional Gaussian, Euclidean nearest neighbors will be biased towards the center of the distribution, which will most likely be the majority group.

☑ It cannot be assumed that their method of compilation was representative.

> ✓ **Correct**
> Correct! Without further information, their method of compilation is potentially a source of bias.

☐ None of the above.

6. Your friend tells you that they have a conditional GAN, where they added an extra loss to the generator to encourage it to reproduce the ground truth image given the conditions. They do this by penalizing the distance from the generated image to the original image and are asking you about fairness.

1 / 1 point

Your friend wants to know whether they should use the absolute value of the pixel differences between the two images as the penalty, or the square of the differences. Images of one group of people, which are a smaller fraction of the dataset, have an average pixel brightness of 0.3 while the images of the other group have an average pixel brightness of 0.9. What are some reasonable answers you could give your friend?

☑ Try both and then ask impartial people from both groups to evaluate it.

> ✓ **Correct**
> Correct! It is reasonable to ask for external feedback, as there may be otherwise unforeseen consequences to algorithmic decisions.

☐ Don't worry about it since algorithmic decisions don't cause bias.

☑ Using the quadratic (L2) loss would penalize the model more for the lighter group, so using L1 distance will likely correct for some of the dataset inequality.

> ✓ **Correct**
> Correct! The quadratic penalty could realistically cause the model to prioritize the group that is closer to the boundary, a potential source of bias.

☐ None of the above.

7. By taking this course and respective specialization, you agree to do no harm with the knowledge you acquire. GANs give you great power but with it comes responsibility. Be considerate and consider the implications of your work and application of GANs. Your response will be considered as an agreement that you will not use GANs for malicious purposes.

4 / 4 points

☑ I agree

> ✓ **Correct**

☐ I disagree