Duration: 105 minutes Aids Allowed: No aids allowed

| Student Number: | | 1 | I | | |] | | | | | |
|-----------------|--|-------|---|--|--|---|----|-----|-------|--------|------|
| Family Name(s): | | | | | | | | | | | |
| Given Name(s): | | | | | | | | | | | |
| UofT email: | | | | | | | _@ | mai | l.utc | oronto | o.ca |
| UTORid: | | | | | | | | | | | |

Do **not** turn this page until you have received the signal to start. In the meantime, please read the instructions below carefully.

MARKING GUIDE

| | # 1:/ 5 |
|---|-----------|
| | # 2:/ 5 |
| This term test consists of 10 questions on 18 pages (including this one), | # 3:/10 |
| printed on both sides of the paper. When you receive the signal to start, please make sure that your copy of the test is complete, fill in the identifi- | # 4:/10 |
| cation section above, and write your name on the back of the last page. | # 5:/15 |
| and use the next page if you run out of space. If you need more space for one | # 6:/ 5 |
| of your solutions, use the pages at the end of the midterm paper. <i>indicate</i> clearly the part of your work that should be marked. | # 7:/15 |
| You may use the torch and numpy modules | # 8:/10 |
| | # 9:/ 2 |
| | # 10:/ 3 |
| | TOTAL:/80 |

Question 1. [5 MARKS]

In the context of classifiers, define *calibration* and *false positive parity*. Your definitions must use conditional probabilities.

Question 2. [5 MARKS]

Describe a dataset for which a classifier can both satisfy false-positive parity and be calibrated. Describe the classifier. Explain why in your example false positive parity and calibration are satisfied, referring to the definitions of calibration and false positive parity. Your proposed dataset and classifier must be somewhat realistic: for example, it must not be the case that all the features are equal to 0 or all the outputs are equal to 0.

Question 3. [10 MARKS]

Suppose you are using logistic regression with a constant threshold (which doesn't vary with demographic) as your classifier.

Explain how *not* using a sensitive characteristic as an input can lead to unfair results (but using the characteristic would lead to fair results), according to one of the mathematical definitions of fairness we discussed in class.

Clearly state what mathematical notion of fairness you are referring to, explain why that notion of fairness can make intuitive sense, and prove that not using a sensitive characteristic can lead to unfair results.

Question 4. [10 MARKS]

Give an example that illustrates the "explaining away" phenomenon. Your example should include a set of variables that measure real things (e.g., temperature) and a fully-defined probability distribution with plausible parameter values.

Question 5. [15 MARKS]

Write down the cost function that is optimized when training GLoVe embeddings. Explain in detail the intution for arriving at that cost function.

Question 6. [5 MARKS]

In the context of training GANs, explain what the training set distribution and the generator distribution are. (Refer to the next question for the context.)

Question 7. [15 MARKS]

Derive the fact that the GAN training procedure is equivalent to minimizing the Jensen-Shannon divergence between the training distribution and the generator distribution. (You can continue your answer on the next page.)

Question 8. [10 MARKS]

Explain how node2vec embeddings are learned using negative sampling. State what the cost function is, and explain the intution behind it. You do not need to give the details about generating random walks.

Question 9. [2 MARKS]

Define perfmutation-invariant and permutation-equivariant functions.

Question 10. [3 MARKS]

Are node2vec embeddings permutation-invariant, permutation-equivariant, neither, or both? Explain your answer.

On this page, please write nothing except your name.

Family Name(s):

Given Name(s):