CS6603 HW Final Project

Haoli Sun, Lige Han and Xuguang Cao
xcao80@gatech.edu

1 STEP1

The data we are using is from UCI machine learning repository. The data is related with direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to predict if the client will subscribe a term deposit (variable y).

This data belongs to banking domain.

This data has 4521 observations.

This data has 17 variables.

Variable 'bought_product' (variable y) is the dependent variable.

There are 2 variables associated with protected class.

Age: The Age Discrimination in Employment Act of 1967 (ADEA)
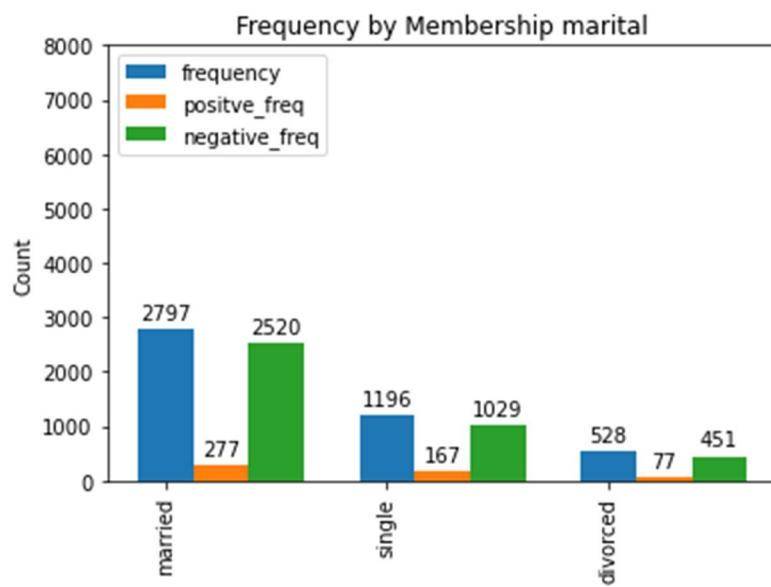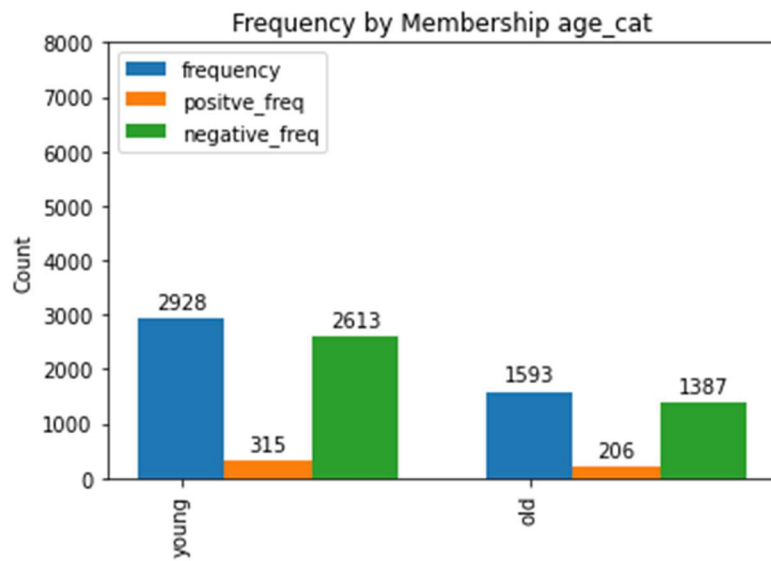
Marital: Civil Right Act of 1968.

2 STEP2

Variable "Age" and "Marital" are associated with protected classes.

Members in the protected classes:

	Age	Marital
Member_1	<= 45 years (old population)	Married
Member_2	>45 (young population)	Single
Member_3		Divorced

Histograms of each membership:

Frequency by Membership age_cat

Frequency by Membership marital

3 STEP3

The following table documents the protected class variable selected, the privileged/unprivileged groups/values, the pre-processing bias mitigation function selected, and the fairness metrics/resulting values computed in Step 3.2 and Step 3.4.

Protected Class	Privileged Group	Unprivileged Group	Bias Mitigation Function	Disparate Impact	Equal Opportunity Difference	Mitiaged Disparate Impact	Mitiaged Equal Opportunity Difference
Age	Old	Young	Reweighting	1.23	0.02	9.31	0.12
Marital	Married	Not Married	Reweighting	0.70	0.04	1.15	0.01

4 STEP4

The dependent variable is column 'y', representing as the client subscribed a term deposit.

| Protected Class	Privileged Group	Unprivileged Group	Disparate Impact	Equal Opportunity Difference	Mitiaged Disparate Impact	Mitiaged Equal Opportunity Difference
Age	Old	Young	0.6	0.011	1.6	0.003

5 STEP5
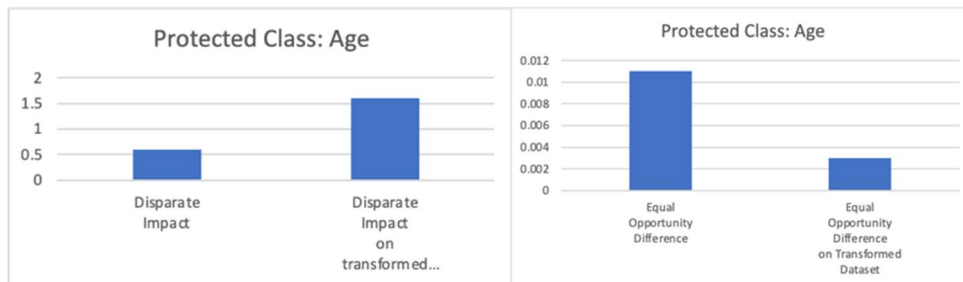
Team members: Haoli Sun, Lige Han, Xuguang Cao

The technique we think is working is Equal Accuracy because it makes sure we have a fair classifier.

Graphs of results from:

Step 3.2 and 3.4:
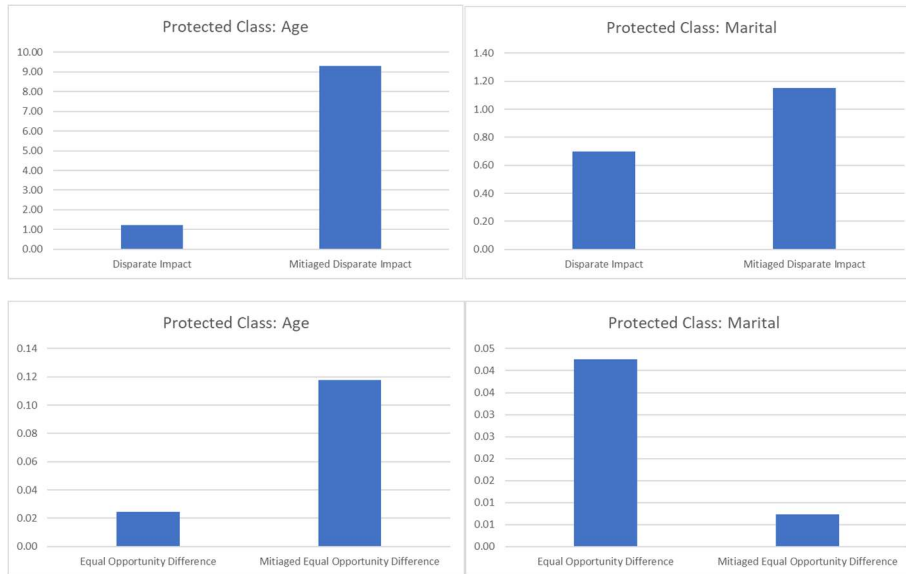
Step 4.5:

Xuguang Cao Response:

I believe the Equal Accuracy strategy seems to work here. Because Equal accuracy means that there is a similar percentage of correct predictions in each slice. The prediction for younger population and the older population is similarly accurate. The prediction male and female population is also similarly accurate. The means we have a fair classifier for each class. The unprivileged groups (young, female) are benefited from this mitigation, getting a higher rate of prediction accuracy. I do not think other groups are affected since this method is conducting overlays on top of the original prediction results. Thus, I do not believe that this method affects other groups (privileged groups). However, single threshold concerns me with increasing false positive rate. This technique although increase accuracy for unprivileged groups, it also changes the false positive and false positive rates of the prediction output, introducing new bias.

Haoli Sun Response:

I believe the equal opporunity metric is the better one comparing to another one, disparate impact, that are being used in our project because it improve the bias in the classifier and also make sure there is similar portion of predic power in each groups. I don't think it impacts other groups as the mitigation process is only applied to unpriviledged groups. And, unpriviledged group received positive impact. However, if only equal opportunity is used, the algorithm may not take into account historical and systemic challenges that certain groups have faced but equally consider them. This could result in new bias that are not accurately reflecting the reality.

Lige Han Response:

I think in this case, the reweighting is working better when equal opportunity differences are evaluated as bias metrics. It is obvious from the graphs that the mitigated disparate impact metrics changes dramaticly without a certain trend of favouring the privileged group or the unprivileged group. In contrast, the mitigated equaly opportunity differences are generally reduced closer to zero (fair). The unprivileged group received an advantege from the reweighting under the equal opportunity difference metric. The issue with evaluating fairness using only 2 types metrics is that it may just not be enough. More fairness metrics should be taken into considerations when implementing a optimized mitigation strategy.