**CSE574 Programming Assignment 3 Report**

**Role in the Project: Machine Learning Engineers (Volunteer for NGO)**

May 8, 2020

**Group Information**

Group Number: 29

Group Members:

Xiyun Xie (51062104)

Mollie Wugetemole (50165532)

Yutong Yang (50321338)

**Proposed Model Information**

Model Choice: Neural Network

Algorithm Choice:

Secondary Optimization Criteria: Accuracy

Overall Cost:

Overall Accuracy:

This project is motivated by the potential fairness violation issues presented in COMPAS. Our goal is to propose a new model that maintain reasonable level of prediction accuracy, while, at the same time, properly apply the concept of fairness.

The stakeholders in this situation includes: 1) the defendants, as the decision made by the model will directly affect their future life; 2) the judges, as the model can either lead them into right decision or wrong decision; 3) the government, as the model will determine potential financial cost; 4) tax payer, as the tax they pay for are going to maintain the jails…. A good model will not only … but also avoid

According to the evaluation of ProPublica, COMPAS shows apparent discrimination among different racial groups []. Black defendants are more likely to …… []. This can be caused by the inherent bias in the raw data … ; This can also be a result of the selection of algorithm. e.g. different prediction model, different postprocessing methods will have different outcome.

Some of the biases are inevitable because they come from … … which we can’t control. But we can adjust the model/algorithm to make relatively fair decision on the unfair data… … To do this, we explore 3 types of prediction model, namely SVM, Naïve Bayes, NN; and apply 5 different postprocessing methods, namely… … for each of them. The criteria we are using to choose our model is mainly accuracy, but at the same time with proper fairness.

Table … shows the …

Table … shows the …

Based on the outputs of different combinations. We believe that … … is the one that most appropriate in terms of our goal…… (state sth about the overall accuracy and fairness)… also the overall cost is …(not too high)

Since ROC is an important…. []. We also create ROC curve to examine our model (Figure…)

The problem lies in our model is the disparity of … this is a tradeoff ……

**Reference**

[1] Primier

[1] ProPublica – Machine Bias

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

[2] Model Comparison

[3] ROC

What is the motivation for creating a new model to replace COMPAS? What problem are you trying to address?

Who are the stakeholders in this situation?

What biases might exist in this situation? Are there biases present in the data? Are there biases present in the algorithms?

What is the impact of your proposed solution?

Why do you believe that your proposed solution a better choice than the alternatives? Are there any metric where your model shows significant disparity across racial lines? How do you justify this?