Senior Design Project Description

Group Name: At Risk Student Prediction
Group Member: Rui Chen, James Fantin, Kun Yi

Project Summary

Our main goal is to design an algorithm to predict student drop out. We want to make this machine learning fair against bias, have the ability to transfer learn and reduce bias when we lack private data about individuals.

Major Tasks

Question 1: How do we transfer prediction from one-school model to multi-school model?

Question 2: How do we detect bias from a single attribute to bias from the intersection of several attributes and remove the bias?

Question 3: How do we create a fair algorithm if we do not have access to private attributes about a student?

Plans

We plan to meet bi-weekly with Dr. Lan and present information about the research we are doing. By the end of the semester, we hope to reproduce existing results in the literature that we have found. We also hope to survey enough literature to become experts in student dropout prediction. In the second semester, we hope to implement the appropriate algorithms that answer each of the three questions posed above.

We will also need to acquire educational data to build our prediction models with. Our plan is to first find public datasets that have the information we desire. We will also start discussions with the registrar at the University and see about acquiring data from the University. If we acquire this data, our plan is to keep this information private and not publish the data.

At the end of the semester, we will evaluate the success of our algorithm based upon existing techniques that already exist. If we have a higher accuracy prediction and remove more bias then existing algorithms, then we will believe our algorithm has been successful. To measure our algorithm against one-school and multi-school prediction, we will build an algorithm with one school's dataset, and measure that algorithm's performance on the other school using transfer learning. If we see good performance with the transfer learning algorithm, then our algorithm will be successful.

Tools To Be Used

We will write our code in Python and possibly R.

Team Member Roles

Initially, we will all do separate work and then combine the work into the project at the end of the semester. We will all be responsible for doing research and reading articles in the field of student dropout prediction.

James will start by investigating removing bias in data without access to the private attribute information about a student.

Kun will start researching transfer prediction from one-school to multi-school using domain adaptation.

Rui will start by investigating removing bias from a single attribute and removing bias from intersectional attributes.