

## **At Risk Student Prediction**

Rui Chen, James Fantin, Kun Yi

Chao Lan

Computer Science Department

University of Wyoming

Virtual Poster Presentation and Demo

Szechwan,China | Rock Springs,Wyoming,USA | Guangxi,China

Student dropout is a major concern of Universities. Nationwide, 28% of students dropout after the first year and at the University of Wyoming, the rate is 22%. Multiple machine learning models have been developed to try to predict when students will drop out so universities may offer at risk students assistance. An issue with existing algorithms on this subject is they do not ensure the predictions are fair amongst students. This has major impacts to educational institutions who must follow Title IX of the Education Amendments Act and treat all students fairly regardless of gender. Another concern is the privacy of student data which is tightly protected by FERPA. While existing fair random forest algorithms protect fairness of demographic data, none are distributed to ensure individual privacy. To address these issues, we design the first distributed fair random forest algorithm. We assume a third party holds private demographic information and is responsible for designing fair models. This third party communicates with a data center that builds a model without compromising the privacy of individuals. In addition, we investigate whether intersectional bias between multiple attributes exists in the data. To detect intersectional bias, we build a detector based on a standard decision tree and conclude that the top k features in the tree are more likely to have intersectional bias than the bottom features in the tree.