

Testing AI fairness in predicting college dropout rate

June 17 2021, by Tom Fleischman



Credit: Unsplash/CC0 Public Domain

To help struggling college students before it is too late, more and more universities are adopting machine-learning models to identify students at risk of dropping out.

What information goes into these models can have a big effect on how accurate and fair they are, especially when it comes to protected student characteristics like gender, race and family income. But in a new study, the largest audit of a college AI system to date, researchers find no evidence that removing protected student characteristics from a [model](#) improves the accuracy or fairness of predictions.

This result came as a surprise to René Kizilcec, assistant professor of information science and director of the Future of Learning Lab.

"We expected that removing socio-demographic characteristics would make the model less accurate, because of how established these characteristics are in studying academic achievement," he said.

"Although we find that adding these attributes provides no empirical advantage, we recommend including them in the model, because it at the very least acknowledges the existence of educational inequities that are still associated with them."

Kizilcec is senior author of "Should College Dropout Prediction Models Include Protected Attributes?" to be presented at the virtual Association for Computing Machinery Conference on Learning at Scale, June 22-25. The work has been nominated for a conference Best Paper award.

Co-authors are Future of Learning Lab members Hannah Lee, a master's student in the field of computer science, and lead author Renzhe Yu, a doctoral student at the University of California, Irvine.

For this work, Kizilcec and his team examined data on students in both a residential college setting and a fully online program. The institution in the study is a large southwestern U.S. public university, which is not named in the paper.

By systematically comparing predictive models with and without

protected attributes, the researchers aimed to determine both how the inclusion of protected attributes affects the accuracy of college dropout prediction, and whether the inclusion of protected attributes affects the fairness of college dropout prediction.

The researchers' dataset was massive: a total of 564,104 residential course-taking records for 93,457 unique students and 2,877 unique courses; and 81,858 online course-taking records for 24,198 unique students and 874 unique courses.

From the dataset, Kizilcec's team built 58 identifying features across four categories, including four protected attributes—student gender; first-generation college status; member of an underrepresented minority group (defined as neither Asian nor white); and high financial need. To determine the consequences of using protected attributes to predict dropout, the researchers generated two feature sets—one with protected attributes and one without.

Their main finding: Including four important protected attributes does not have any significant effect on three common measures of overall prediction performance when commonly used features, including academic records, are already in the model.

"What matters for identifying at-risk students is already explained by other attributes," Kizilcec said. "Protected attributes don't add much. There might be a gender gap or a racial gap, but its association with dropout is negligible compared to characteristics like prior GPA."

That said, Kizilcec and his team still advocate for including protected attributes in prediction modeling. They note that higher education data reflects longstanding inequities, and they cite recent work in the broader machine-learning community that supports the notion of "fairness through awareness."

"There's been work showing that the way certain attributes, like academic record, influence a student's likelihood of persisting in [college](#) might vary across different protected-attribute groups," he said. "And so by including [student characteristics](#) in the model, we can account for this variation across different student groups."

The authors concluded by stating: "We hope that this study inspires more researchers in the learning analytics and educational data mining communities to engage with issues of algorithmic bias and fairness in the models and systems they develop and evaluate."

Kizilcec's lab has done a lot of work on [algorithmic fairness in education](#) , which he said is an understudied topic.

"That's partly because the algorithms [in education] are not as visible, and they often work in different ways as compared with criminal justice or medicine," he said. "In education, it's not about sending someone to jail, or being falsely diagnosed for cancer. But for the individual [student](#), it can be a big deal to get flagged as at-risk."

More information: Renzhe Yu et al, Should College Dropout Prediction Models Include Protected Attributes?, *Proceedings of the Eighth ACM Conference on Learning @ Scale* (2021). [DOI: 10.1145/3430895.3460139](#)

Provided by Cornell University

Citation: Testing AI fairness in predicting college dropout rate (2021, June 17) retrieved 24 April 2024 from <https://phys.org/news/2021-06-ai-fairness-college-dropout.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.