Targeting degree-awarding gap across ethnicities through means of OUAnalyse predictions

How to cite:


© 2021 Vaclav Bayer

https://creativecommons.org/licenses/by/4.0/

Version: Poster
TARGETING DEGREE-AWARDING GAP ACROSS ETHNICITIES THROUGH MEANS OF OUNAALYSE PREDICTIONS

INTRODUCTION

- UniversitiesUK and AdvanceHE report a 13% degree-awarding gap for Black, Asian, and Minority Ethnic (BAME). This degree-awarding gap is also present at the Open University.
- OUNAALyse generates predictions aiming to identify students at-risk of not submitting their assignments. Last year pilot showed 7% higher chances to pass modules for students when their tutors used OUNAALyse.
- We investigate if the predictions provided by existing Learning Analytics (LA) models are fair and serve the majority and minority ethnic groups with the same effectiveness.

METHODOLOGY

- Analysis of predictions made by LA models in the 14 largest modules from the year 2019/20.
- Evaluation (Baseline) made on data of 32,538 unique students. When disaggregated by ethnicity: White (28,535), Black (1,078), Asian (1,195), Rest (Mixed, Other, Refused, Unknown) (1,730).

METRICS

- False Positive Rate (FPR) - students erroneously predicted to Not Submit (NS).
- False Negative Rate (FNR) - students erroneously predicted to Submit (NS), more severe error as students most likely don’t receive needed support.
- AUC - model’s overall accuracy.

RQ1: Do existing LA prediction models work equally effectively for all ethnicities?

- A separate evaluation of each ethnicity is made.

- Comparison of the minority groups evaluation with the majority group (White eth.) evaluation (green-better score, red-worse score).

Existing LA models contain inequalities in accuracy and error rates across different ethnicity groups. Different methods can help to reduce inequalities on different levels, but the solution is not systematic, and therefore, different adaptations and definitions of fairness are needed.

RQ2: Fairness through unawareness

- The “ethnicity” protected attribute is excluded during the model training process.

- A comparison with Baseline evaluation is made to discover the influence of unawareness on the metrics (green-improvement, red-deterioration).

RQ3: Do the LA population-specific prediction models perform better?

- The models are trained and evaluated only on a specific population of students, then compared to evaluation of corresponding Baseline population.

CONCLUSION

- Existing LA models discover inequalities in terms of accuracy and fairness across ethnicities (RQ1).
- Removing the protected attributes from model training (RQ2) or building separate models for particular ethnicity (RQ3) seems to enhance accuracy and fairness for some ethnicity groups but does not systematically make the models more accurate and fair for everyone.
- Different settings bring different results. Therefore more research in terms of different adaptations and definitions of fairness is needed to ensure that the technology solution we build does not perpetuate existing educational gaps.

Black, Asian and Minority Ethnic (BAME) students at the Open University put more effort and spend more time studying, they are, however, less likely to complete, pass or achieve an excellent grade compared to White students.

Existing LA models contain inequalities in accuracy and error rates across different ethnicity groups. Different methods can help to reduce inequalities on different levels, but the solution is not systematic, and therefore, different adaptations and definitions of fairness are needed.

Vaclav Bayer, Martin Hilosta, Miriam Fernandez
vaclav.bayer@open.ac.uk
https://orcid.org/0000-0001-8953-6335