ABSTRACT

The increasing impact of machine learning and algorithmic decision making on education has brought about growing opportunities and concerns. Evidence has shown that these technologies can perpetuate and even magnify existing educational and social inequities. Research on fair machine learning has aimed to develop algorithms that can detect and, in some cases correct, bias, but this effort within the educational data mining community is still limited.

FATED 2022 hopes to spur discussion around algorithmic fairness and bias detection as specifically applied in an educational context. Submissions and panels will be invited to discuss: (a) collection and preparation of benchmark datasets for bias detection and correction tasks, (b) evaluation protocol definition and metric formulation appropriate for bias and fairness in educational tasks, and (c) countermeasure design and development for biased and unfair circumstances. These specific topics will be complemented by a more general discussion of the education-specific challenges for fair machine learning in education, bringing together perspectives from both industry and academia. This workshop builds on the FATED workshop held at EDM 2020, and we expect the workshop to make connections among already interested researchers and provide a foundation for those who want to engage in this area.

Part of the vision of creating adaptive educational technologies and building machine learning systems for education is reducing inequality (e.g., [2]), and data-driven practices are often viewed as a way to make education more equitable (e.g., [1]). While some interventions have been found to decrease achievement gaps (e.g., [4]), there is increasing concern that these systems may instead increase achievement gaps and perpetuate existing inequities [10, 12]. For example, such systems might make targeted support more available only to students with greater access to technology, or be associated with lower learning gains in more disadvantaged schools (as seen in [11]).

In this workshop, we hope to bring an education-specific lens on broader questions related to fair ML by spurring discussion around:

- **Data Set Collection and Preparation.** By spurring discussion about what educational datasets are particularly ripe for use as benchmarks for detecting and/or correcting bias and what characteristics of an educational dataset make it most useful for measuring or detecting algorithmic bias, this workshop aims to increase awareness about what datasets are available and encourage future research to include results on benchmark datasets.

- **Evaluation Protocol and Metric Formulation.** This workshop encourages discussion about what evaluation protocols and metrics are most suitable for empirical research on fairness and bias across common types of educational machine learning and EDM tasks.

- **Detection and Countermeasure Design.** FATED 2022 provides a forum for discussion about what features of the questions that we address in educational machine learning and the datasets that we use pose particular challenges for detecting and/or addressing algorithmic bias. Further, the workshop will provide an opportunity for researchers to share their work on algorithmic bias detection and correction specifically in education-related context.

Around these themes, FATED 2022 will showcase papers that focus on datasets, evaluation protocol, research, reproducibility, and recently published work (encore papers). By stimulating these discussions, the organizers hope to build community among researchers in this area, including interested EDM researchers who are not yet involved in these topics and fair ML researchers who may wish to engage with the field of education. Surrounding literature from the workshop organizers focuses on educational technology [3, 20, 6, 13, 15], student behavioral patterns [7, 8], algorithmic fairness [19, 18, 5], explainability [17], and responsible analytics for social good [14, 9, 16].
1. REFERENCES


