FATED 2022: Fairness, Accountability, and Transparency in Educational Data

Collin Lynch
North Carolina State University
cflynch@ncsu.edu

Mirko Marras
University of Cagliari
mirko.marras@acm.org

Mykola Pechenizkiy
Eindhoven University of Technology
m.pechenizkiy@tue.nl

Anna N. Rafferty
Carleton College
arafferty@carleton.edu

Steve Ritter
Carnegie Learning
sritter@carnegielearning.com

Vinitra Swamy
EPFL
vinitra.swamy@epfl.ch

Renzhe Yu
University of California, Irvine
renzhey@uci.edu

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]).

C. Lynch, M. Marras, M. Pechenizkiy, A. Rafferty, S. Ritter, V. Swamy, and R. Yu. FATED 2022: Fairness, accountability, and transparency in educational data. In A. Mitrovic and N. Bosch, editors, *Proceedings of the 15th International Conference on Educational Data Mining*, pages 848–849, Durham, United Kingdom, July 2022. International Educational Data Mining Society.

https://doi.org/10.5281/zenodo.6853079

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

[1] Bill and Melinda Gates Foundation. Ensuring all students receive a public education that equips them to succeed. https://usprogram.gatesfoundation.org/What-We-Do/K-12-Education. Accessed: 2022-02-25.

[2] B. du Boulay, A. Poulovasillis, W. Holmes, and M. Mavrikis. Artificial intelligence and big data technologies to close the achievement gap. In R. Luckin, editor, *Enhancing Learning and Teaching with Technology*, page 256–285. UCL Institute of Education Press, 2018.

[3] N. Gitinabard, R. Okoilu, Y. Xu, S. Heckman, T. Barnes, and C. Lynch. Student teamwork on programming projects what can github logs show us? In *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, pages 409–416, 2020.

[4] X. Huang, S. D. Craig, J. Xie, A. Graesser, and X. Hu. Intelligent tutoring systems work as a math gap reducer in 6th grade after-school program. *Learning and Individual Differences*, 47:258–265, 2016.

[5] C. Kung and R. Yu. Interpretable Models Do Not Compromise Accuracy or Fairness in Predicting College Success. In *Proceedings of the 7th ACM Conference on Learning @ Scale (L@S '20)*, pages 413–416, New York, NY, USA, aug 2020. Association for Computing Machinery (ACM).

[6] Z. Li, L. Yee, N. Sauerberg, I. Sakson, J. J. Williams, and A. N. Rafferty. Getting too personal(ized): he importance of feature choice in online adaptive algorithms. In *Proceedings of the 13th International Conference on Educational Data Mining, EDM 2020, Fully virtual conference, July 10-13, 2020*. International Educational Data Mining Society, 2020.

[7] Z. Liu, R. Brown, C. F. Lynch, T. Barnes, R. Baker, Y. Bergner, and D. McNamara. Mooc learner behaviors by country and culture; an exploratory analysis. 2016.

[8] C. F. Lynch. Who prophets from big data in education? new insights and new challenges. *Theory and Research in Education*, 15(3):249–271, 2017.

[9] M. Mansoury, H. Abdollahpouri, M. Pechenizkiy, B. Mobasher, and R. Burke. Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM international conference on information & knowledge management*, pages 2145–2148, 2020.

[10] T. G. Mathewson. Personalized learning can be a tool for equity or a barrier to it. https://hechingerreport.org/personalized-learning-can-be-a-tool-for-equity-or-a-barrier-to-it/. The Hechinger Report; Accessed: 2022-02-25.

[11] M. Meeter. Primary school mathematics during the covid-19 pandemic: No evidence of learning gaps in adaptive practicing results. *Trends in neuroscience and education*, 25:100163, 2021.

[12] A. Perry and N. Turner-Lee. Ai can disrupt racial inequity in schools, or make it much worse. https://hechingerreport.org/ai-can-disrupt-racial-inequity-in-schools-or-make-it-much-worse/. The Hechinger Report; Accessed: 2022-02-25.

[13] S. Ritter, M. Yudelson, S. E. Fancsali, and S. R. Berman. How mastery learning works at scale. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*, pages 71–79, 2016.

[14] A. Saxena, G. Fletcher, and M. Pechenizkiy. Hm-eiict: Fairness-aware link prediction in complex networks using community information. *Journal of Combinatorial Optimization*, pages 1–18, 2021.

[15] V. Swamy. Pedagogy, infrastructure, and analytics for data science education at scale, 2018.

[16] V. Swamy, E. Chen, A. Vankayalapati, A. Aggarwal, C. Liu, V. Mandava, and S. Johnson. Machine learning for humanitarian data: Tag prediction using the HXL standard. 2019.

[17] V. Swamy, A. Romanou, and M. Jaggi. Interpreting language models through knowledge graph extraction. *arXiv preprint arXiv:2111.08546*, 2021.

[18] R. Yu, H. Lee, and R. F. Kizilcec. Should College Dropout Prediction Models Include Protected Attributes? In *Proceedings of the Eighth ACM Conference on Learning @ Scale*, pages 91–100, New York, NY, USA, jun 2021. ACM.

[19] R. Yu, Q. Li, C. Fischer, S. Doroudi, and D. Xu. Towards accurate and fair prediction of college success: Evaluating different sources of student data. In *Proceedings of the 13th International Conference on Educational Data Mining, EDM 2020, Fully virtual conference, July 10-13, 2020*. ERIC, 2020.

[20] Z. Zakaria, J. Vandenberg, J. Tsan, D. C. Boulden, C. F. Lynch, K. E. Boyer, and E. N. Wiebe. Two-computer pair programming: Exploring a feedback intervention to improve collaborative talk in elementary students. *Computer Science Education*, pages 1–28, 2021.