

Causal Inference in Educational Data Mining

Third Annual Half-Day Workshop

Adam C. Sales, Neil T. Heffernan
Worcester Polytechnic Institute
asales@wpi.edu
nth@wpi.edu

ABSTRACT

The goal in crafting intelligent tutoring systems, educational games, MOOCs, and other computerized learning tools, is to improve student learning. To that end, EDM research typically focuses on methods to identify, measure, and predict learner behaviors or outcomes. Causal research seeks to estimate the impacts of different factors on these behaviors or outcomes—not only predicting who will wheel-spin, experience frustration, or successfully learn a new skill, say, but determining causes these? Causality lies at the heart of both learning science, which seeks to understand how inputs in an educational system affect the system's outputs, and of policy, which seeks to design educational systems that improve learning.

The field of causal inference, which spans statistics, philosophy, economics, computer science, and other more traditional academic disciplines, has itself experienced rapid and exciting developments in the recent past. The new science of causality encompasses new ways of estimating effects under challenging circumstances, such as possible confounding, but also new questions—how do impacts vary between learners? What mechanisms drive causal effects? How may we construct optimal individualized policies for specific learners?

This workshop is intended to raise awareness of the ubiquity and importance of causal questions in EDM, some of the exciting methods available to address those questions, and some of the open questions of causal inference in EDM. It will include invited discussions of ongoing projects addressing causal questions, and short talks about relevant work in progress, including work in any stage of development.

Lastly, the workshop will give an opportunity for EDM researchers to submit open problem related to causality in EDM research. In five minute presentations, researchers will briefly present problems they have encountered in research, or that they just think are interesting, but that they do not yet know how to solve. Each presentation will be followed by an open-ended discussion among the workshop participants, hopefully suggesting ways to solve, or at least better refine the problem. This sub-workshop will hopefully give the presenting researchers constructive suggestions, and spur collaborations. In general, the workshop will be organized to stimulate discussion among participants, including, hopefully, constructive suggestions for open problems.

This workshop will be the third annual EDM workshop of its type—now jointly hosted along with International Conference on Artificial Intelligence in Education. It will follow a similar structure to the previous two workshops, but with all new material.

A. Sales and N. Heffernan. Causal inference in educational data mining. In A. Mitrovic and N. Bosch, editors, *Proceedings of the 15th International Conference on Educational Data Mining*, pages 852–852, Durham, United Kingdom, July 2022. International Educational Data Mining Society.

© 2022 Copyright is held by the author(s). This work is distributed under the Creative Commons Attribution NonCommercial NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.
<https://doi.org/10.5281/zenodo.6853045>