

## **Learning to Personalize from Observational Data Under Unobserved Confounding**

Nathan Kallus, Cornell University & Cornell Tech

Recent work on counterfactual learning from observational data aims to leverage large-scale data -- much larger than any experiment can ever be -- to learn individual-level causal effects for personalized interventions. The hope is to transform electronic medical records to personalized treatment regimes, transactional records to personalized pricing strategies, and click- and "like"-streams to personalized advertising campaigns. Motivated by the richness of the data, existing approaches (including my own) make the simplifying assumption that there are no unobserved confounders: unobserved variables that affect both treatment and outcome and would induce non-causal correlations that cannot be accounted for. However, all observational data, which lacks experimental manipulation, no matter how rich, will inevitably be subject to some level of unobserved confounding and assuming otherwise can lead to personalized treatment policies that seek to exploit individual-level effects that are not really there, may intervene where not necessary, and may in fact lead to net harm rather than net good relative to current, non-personalized practices.

The question is then how to use such powerfully rich data to safely improve upon current practices. In this talk, I will present a novel approach to the problem that calibrates policy learning to realistic violations of the unverifiable assumption of unconfoundedness. Our framework for confounding-robust policy improvement optimizes the minimax regret of a candidate policy against a baseline standard-of-care policy over an uncertainty set for propensity weights motivated by sensitivity analysis in causal inference. By establishing a finite-sample generalization bound, we prove that our robust policy, when applied in practice, is (almost) guaranteed to do no worse than the baseline and improve upon it if it is possible. We characterize the adversarial optimization subproblem and use efficient algorithmic solutions to optimize over policy spaces such as hyperplanes, score cards, and decision trees. We assess our methods on a large clinical trial of acute ischaemic stroke treatment, demonstrating that hidden confounding can hinder existing approaches and lead to overeager intervention and unwarranted harm, while our robust approach guarantees safety and focuses on well-evidenced improvement, a necessity for making personalized treatment policies learned from observational data usable in practice.