

Collaborative Treatment Effect Estimation in Distributed Research Networks

Date: 16 January 2025 (Thursday)

Time: 9:30am - 10:30am



Seminar Link: https://cityu.zoom.us/j/83109806687

ABSTRACT

Distributed research networks (DRNs) are increasingly utilized to conduct multi-site observational studies, enhancing the efficiency and generalizability of causal inference. However, challenges such as data-sharing restrictions and presence of small sites complicate collaborative treatment effect estimation. This talk introduces two recent works addressing these challenges.

The first part of the presentation will focus on collaborative estimation of the average treatment effect (ATE). In a clinical DRN, the data partners may range from large health centers to small community clinics. A significant challenge arises from those "small sites" with limited sample size, in which the small sample size results in inaccurate estimation of propensity score models, poor adjustment for confounders, and biased estimates for ATE. To mitigate such biases, we propose a robust network federated transfer learning framework to improve the estimation efficiency for propensity score models across all sites. Our framework performs knowledge transfer across sites such that all sites can borrow useful information from others, which is particularly beneficial for small sites. We provide theoretical guarantee that our framework avoids "negative transfer", ensuring the improvement in propensity score estimation and ATE estimation. The second part focuses on collaborative estimation of heterogeneous treatment effects (HTEs) for different subpopulations to assist targeted interventions. In real-world biomedical applications, there is often no clear definition for the subpopulations, so an important task is to use data-driven methods to identify clinically meaningful subpopulations based on patient covariates distributed in different sites. To tackle this federated clustering problem, we develop a heterogeneous mixture model equipped with a novel one-shot distributed EM algorithm. This algorithm enables efficient distributed inference with only one round of cross-site communication. We provide theoretical guarantee that our one-shot estimator achieves full-sample efficiency, which allows for efficient identification of subpopulation and downstream HTE estimation.

DS SEMINAR



Dr. Yudong WANG GUEST SPEAKER'S PROFILE

Yudong WANG is a postdoctoral research fellow in the Department of Biostatistics, Epidemiology, and Informatics at the University of Pennsylvania, working with Professor Yong Chen. He obtained his PhD in 2023 from the National University of Singapore, advised by Professor Zhisheng Ye. His research aims to advance healthcare AI by developing novel statistical and machine learning approaches to support efficient data-driven decision-making in healthcare systems. His specific research interests include distributed inference, transfer learning, semiparametric methods, and learning health systems.

Enquiries: ds.go@cityu.edu.hk

All are welcome