Flexible Machine Learning Estimation of Conditional Average Treatment Effects

Richard Post

Causal inference from observational data requires untestable identification assumptions. If these assumptions apply, machine learning (ML) methods can be used to study complex forms of causal effect heterogeneity. Several ML methods were recently developed to estimate the conditional average treatment effect (CATE). If observed features cannot explain all heterogeneity, the individual treatment effects (ITEs) can still seriously deviate from the CATE. In this talk, we will illustrate the possible difference between the ITE distribution and the individualized CATE distribution by presenting scenarios with varying conditional ITE variance. If the distribution of the ITE equals that of the CATE, the observed difference in conditional variance between treated and controls should be small. If they differ, an additional causal identifiability assumption is necessary to quantify the heterogeneity not captured by the distribution of the CATE. The conditional variance of the ITE can be identified when the ITE is independent of the outcome under no treatment given the measured features. Under this assumption, we extend the causal random forest algorithm to illustrate how ML methods may be used to estimate the conditional variance of the ITE from observational data. For the scenarios where the ITE and CATE distributions differ, we show that the extended causal random forest can appropriately estimate the variance of the ITE distribution, while the traditional causal random forest fails to do so. Finally, we will discuss the impact of the violation of the untestable identifiability assumption on the performance of the extended random forest.