Department of Statistics

STATISTICS COLLOQUIUM

VICTOR VEITCH

Department of Statistics
Columbia University

Adapting Black-box Machine Learning Methods for Causal Inference

WEDNESDAY, February 5, 2020 at 4:00 PM
Jones 303, 5747 S. Ellis Avenue

ABSTRACT

I'll discuss the use of observational data to estimate the causal effect of a treatment on an outcome. This task is complicated by the presence of 'confounders' that influence both treatment and outcome, inducing observed associations that are not causal. Causal estimation is achieved by adjusting for this confounding by using observed covariate information. I'll discuss the case where we observe covariates that carry sufficient information for the adjustment, but where explicit models relating treatment, outcome, covariates, and confounding are not available. For example, in medical data the covariates might consist of a large number of convenience health measurements of which only an unknown subset are relevant, and even then in some totally unknown manner. Or, the covariates might be a passage of (natural language) text that describes the relevant information. I'll describe an approach that adapts deep learning and embedding methods to produce representations of the covariate information targeted toward the causal adjustment problem. In particular, I'll describe how to modify standard architectures and training objectives to achieve statistically efficient and practically useful causal estimates.
