

Targeted integrative learning with applications in suicide risk prediction

Kun Chen

University of Connecticut
E-mail: kun.chen@uconn.edu

Abstract: In many scientific problems, the goal is to make inference on a specified “target population” of interest. For example, in a single-arm clinical trial, the target population can be defined by the treated patients and the key is to find out what happens to them if they were not treated; in a suicide risk study, the target population may be consist of patients who received care from a specific healthcare provider. Yet, the data available may go way beyond the target population. As such, a crucial question is how to best integrate all the information to improve the inference for the target. In this talk, we consider two scenarios. For the scenario of “integrated data”, we propose a distance-segmented regression (DSR) framework, in which a distance metric is used to measure how close each sample is to the target and is assumed to guide the conditional association between the outcome and predictors. For the scenario of “non-integratable” data, we propose a transfer learning model, in which the target population and the external database are linked through subject similarities. Applications in suicide risk prediction with medical claim data will be discussed.