A Semiparametric Perspective on Unsupervised Domain Adaptation: Doubly Flexible Estimation and Efficient Label Shift Adaptation



Jiwei Zhao, PhD

Associate Professor, University of Wisconsin-Madison

In this talk, we consider the unsupervised domain adaptation setting under the label shift assumption. In the first part, we estimate a parameter of interest in population Q by leveraging information from P, where three ingredients are essential: (a) the common conditional distribution of X given Y, (b) the regression model of Y given X in P, and (c) the density ratio of the outcome Y between the two populations. We propose an estimation procedure that only needs some standard nonparametric technique to approximate the conditional expectations with respect to (a), while by no means needs an estimate or model for (b) or (c); i.e., doubly flexible to the model misspecifications of both (b) and (c). In the second part, we pay special attention to the case that the outcome Y is categorical. In this scenario, traditional label shift adaptation methods either suffer from large estimation errors or require cumbersome post-prediction calibrations. To address these issues, we propose a moment-matching framework for adapting the label shift, and an efficient label shift adaptation method where the adaptation weights can be estimated by solving linear systems. We rigorously study the theoretical properties of our proposed methods. Empirically, we illustrate our proposed methods in the MIMIC-III database as well as in some benchmark datasets including MNIST, CIFAR-10, and CIFAR-100.

Tuesday, February 13, 2024, 3:30-4:30PM Eastern

2308 McGavran-Greenberg Hall

Zoom Link: https://unc.zoom.us/j/94304183888?pwd=SW5kQmJ1bkVVOFN5U2ZiQzR0MThnQT09

Meeting ID: 943 0418 3888

Passcode: 050276

