THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

MASTER'S THESIS PRESENTATION

QINGYAO SUN
Department of Statistics
The University of Chicago

Beyond Invariance: Test-Time Label-Shift Adaptation for Addressing "Spurious" Correlations

THURSDAY, February 2, 2023, at 2:30 PM
Jones 111, 5747 S. Ellis Avenue

ABSTRACT

Spurious correlations, or correlations that change across domains where a model can be deployed, present significant challenges to real-world applications of machine learning models. However, such correlations are not always "spurious"; often, they provide valuable prior information for a prediction. Here, we present a test-time adaptation method that exploits the spurious correlation phenomenon, in contrast to recent approaches that attempt to eliminate spurious correlations through invariance. We consider situations where the prior distribution $p(y,z)$, which models the dependence between the class label y and the "nuisance" factors z, may change across domains, but the generative model for features $p(\mathbf{x}|y,z)$ is constant. We note that this corresponds to an expanded version of the label shift assumption, where the labels now also include the nuisance factors z. Based on this observation, we train a classifier to predict $p(y,z|\mathbf{x})$ on the source distribution, and propose a test-time label shift correction that adapts to changes in the marginal distribution $p(y, z)$ using unlabeled samples from the target domain. We evaluate our method, which we call ``Test-Time Label-Shift Adaptation" (TTLSA), on two different image datasets -- the CheXpert chest X-ray dataset and the Colored MNIST dataset --- and show a significant improvement over baseline methods.