ACMS Statistics Seminar

Jake Soloff

University of Chicago

Wednesday, April 26, 2023

154 Hurley Hall

3:30 PM – 4:30 PM

Bagging Provides Assumption-Free Stability

Algorithmic stability is a framework for studying the properties of a model fitting algorithm, with downstream implications for generalization, predictive uncertainty quantification, and other important inferential goals. Stability is often defined as the property that predictions on a new test point are not substantially altered by removing a single point at random from the training set. However, this stability property itself is typically an assumption: for highly complex predictive algorithms and/or non-smooth data distributions, stability may fail or at least be hard to quantify. In this work, we derive a finite-sample guarantee on the stability of bagging for any model with bounded outputs. Our result places no assumptions on the distribution of the data, on the regularity of the base algorithm, or on the dimensionality of the covariates. Our guarantee applies to many variants of bagging and is optimal up to a constant.

This is joint work with Rina Foygel Barber and Rebecca Willett.