ANDREW GOLDSTEIN

Department of Statistics
The University of Chicago

VEB-Boost: Variational Empirical Bayes Boosting

WEDNESDAY, December 9, 2020, at 1:00 PM
ZOOM Meeting

ABSTRACT

One of the most commonly used models in statistics and machine learning comes from the general regression setting: \( y = f(x) + e \), where \( f(x) \) is an unknown function describing the conditional mean of \( y \) given \( x \), and \( e \) is iid mean-zero Gaussian noise. In this setting, the goal is often to find an approximation to \( f(x) \), say \( g(x) \), such that for unseen \((x, y)\) pairs, \( g(x) \) is close to \( y \). One of the most successful methods in achieving this goal is boosting, in which many “weak learners” that are individually only slightly informative are combined into a “strong learner” that is able to effectively approximate \( f(x) \). However, boosting is plagued with the need to fine-tune a slew of hyper-parameters using computationally intensive cross-validation. Additionally, boosting is only suited to return point-estimates and does not provide a sense of uncertainty about those estimates.

Motivated by both boosting’s predictive power and its shortcomings, this talk outlines variational empirical Bayes boosting (VEB-Boost), a novel framework that borrows ideas from variational Bayes, empirical Bayes, and boosting in order to combine weak learners into a strong learner.