Yue You (UC Berkeley)

Title: Targeted Learning for Variable Importance in Precision Medicine

Abstract: Determining the relative importance of a variable is typically done as a byproduct of fitting algorithms, such as coefficients in large-scale regression (e.g., LASSO), leave-one-variable- out changes in risk (random forest) and other methods that might yield insight, but are tied to specific fitting procedures and even parametric models. We proposed a parameter within a non-parametric model that measures the importance of each variable as the amount of at- tribution of that variable towards changes in the mean outcome. Specifically, for each of the candidate competing causes of the outcome, we utilized an estimate of this attribution using a statistical approach based upon a combination of machine learning and causal inference via Targeted Learning. This approach allows for 1) variable importance comparisons at the same scale regardless of the original scale of the variable, 2) estimation not dependent on arbitrary parametric assumptions, 3) and asymptotically linear (locally efficient) estimator for which robust asymptotic inference is available. We implemented this approach to determine the variable importance of clinical measures in trauma patients in predicting the probability of mortality at different time periods (from time of injury) using data from three independent trauma studies. This approach allowed comparisons of variable importance within and between trauma cohorts and identified variables with the biggest potential ”intervention” impact for mortality. Our re- sults showed that the most important variables across all time intervals is initial International Normalized Ratio, while importance of other variables varied by time. These findings were similar across three trauma cohorts. This method can serve as an alternative to more standard variable importance procedures that lack both broad clinical interpretability and mechanisms for accurate statistical inference in the context of data-adaptive estimation.