Statistical inference with Adaptively Collected Data

Abstract:
Modern experimentation, data collection and modeling is often adaptive – the insight gleaned from prior data informs and influences the data collected in the future. While this adaptivity is practically useful, the correlations it induces in data are problematic for inference. Indeed, classical methods for computing uncertainty estimates like confidence intervals and p-values can be misleading in the presence of such correlations.
In this talk, I will focus on a simple model of adaptive data collection in batches. I will discuss online debiasing: a flexible algorithmic procedure that 'debiases' estimators computed on such data, thereby allowing to compute confidence intervals and p-values in the standard fashion. The procedure is efficient computationally and the resulting uncertainty estimates are near-optimal.

This is based on joint work with: Adel Javanmard, Lester Mackey, Mohammad Mehrabi, Vasilis Syrgkanis, Matt Taddy.