Distributed Statistical Algorithms[*]

Gérard Biau

Abstract

Distributed computing offers a high degree of flexibility to accommodate modern learning constraints and the ever increasing size of datasets involved in massive data issues. Drawing inspiration from the theory of distributed computation models developed in the context of gradient-type optimization algorithms, I will present a consensus-based asynchronous distributed approach for nonparametric online regression and analyze some of its asymptotic properties. Substantial numerical evidence involving up to 28 parallel processors is provided on synthetic datasets to assess the excellent performance of the method, both in terms of computation time and prediction accuracy.