



COLOQUIO DE ESTADÍSTICA

MACHINE LEARNING

29 de octubre de 2014 / Auditorio-ILMAS / 12:00 horas
Circuito Escolar, Ciudad Universitaria

PLÁTICAS

Parallelizing MCMC for Bayesian nonparametrics

Ph.D Sinead Williamson

University of Texas at Austin, USA

Bayesian nonparametric models, such as those based on the Dirichlet process and the Pitman-Yor process, provide elegant and flexible alternatives to parametric models when the number of underlying components is unknown or growing. Unfortunately, inference in such models can be slow, and previous parallelization methods have relied on introducing approximations which can lead to inaccuracies in the posterior estimate. In this talk, I will construct auxiliary variable representations for the Dirichlet process, the Pitman-Yor process, and some hierarchical extensions, and show how these representations facilitate the development of distributed Markov chain Monte Carlo schemes that use the correct equilibrium distribution. Experimental analyses show that this approach allows scalable inference without the deterioration in estimate quality that accompanies existing methods. Joint work with Avinava Dubey and Eric Xing.