

*Reliable Effective Terascale Linear Learning*

Dr. John Langford

Doctor of Learning, Yahoo Research

Friday, January 27, 2012

11:00 – 12:30

McGill University, McConnell Engineering building, room 603

Abstract: How do you learn a linear predictor on a dataset with 2 trillion nonzero features in a reasonable amount of time? Allsingle-machine algorithms fail, because the time to even stream the data through a network interface is too great. I will discuss an algorithm for doing this based on a combination of online learning, LBFGS, parallel learning, and a new communication infrastructure---Hadoop compatible Allreduce that we created. Deployed on 1000 nodes, we can learn faster than any single machine ever will be able to learn a linear predictor on this hardware, the first such learning algorithm for which this claim can be made. The code behind this is open source in Vowpal Wabbit (<http://hunch.net/~vw/>).

This is joint work with Alekh Agarwal, Olivier Chapelle and Miroslav Dudik.

---