Statistics Seminar
Department of Mathematical Sciences

DATE:	Thursday, Sept.30, 2021
TIME:	1:15pm – 2:15pm
LOCATION:	Zoom meeting
SPEAKER:	Yangsheng Wang, Binghamton University
TITLE:	Reinforcement Learning Trees

Abstract

In this paper, we introduce a new type of tree-based method, reinforcement learning trees (RLT), which exhibits significantly improved performance over traditional methods such as random forests (Breiman, 2001) under high-dimensional settings. The innovations are three-fold. First, the new method implements reinforcement learning at each selection of a splitting variable during the tree construction processes. By splitting on the variable that brings the greatest future improvement in later splits, rather than choosing the one with largest marginal effect from the immediate split, the constructed tree utilizes the available samples in a more efficient way. Moreover, such an approach enables linear combination cuts at little extra computational cost. Second, we propose a variable muting procedure that progressively eliminates noise variables during the construction of each individual tree. The muting procedure also takes advantage of reinforcement learning and prevents noise variables from being considered in the search for splitting rules, so that towards terminal nodes, where the sample size is small, the splitting rules are still constructed from only strong variables. Last, we investigate asymptotic properties of the proposed method under basic assumptions and discuss rationale in general settings.