

# STATISTICS SEMINAR

Ruoqing Zhu, Ph.D.

University of Illinois at  
Urbana-Champaign

Thursday, March 24, 2016

Dickens Hall, Room 207, 4:00-5:00 pm

Refreshments: Dickens 108, 3:30-4:00 pm



## Greedy Learning of Tree-based Methods in Personalized Medicine

Interpretability of a treatment decision model is highly desirable in personalized medicine. Such a property is naturally offered by single tree-based methods, such as the classic CART model proposed by Breiman et. al. (1984). When high-dimensional genetic profiles are often involved in medical studies, the performance of single tree-based methods suffers greatly from this curse of dimensionality. As what we shall demonstrate in this talk, this is mostly due to the suboptimal design of the splitting rules. To combat this task, we propose a greedy splitting variable selection strategy, based on a weighted variable importance measure. Numerical studies suggest that the propose method is more likely to produce simpler tree structures and yet more accurate decision rules than existing tree-based methods. Its performance is further demonstrated on the cancer cell line encyclopedia (CCLE) data.