

## Doctoral Dissertation Defense

### “Randomness in Tree Ensemble Methods”

by

**Joran Elias**

Tree ensembles have proven to be a popular and powerful tool for predictive modeling tasks. The theory behind several of these methods (e.g. boosting) has received considerable attention. However, other tree ensemble techniques (e.g. bagging, random forests) have attracted limited theoretical treatment. Specifically, it has remained somewhat unclear as to why the simple act of randomizing the tree growing algorithm should lead to such dramatic improvements in performance. It has been suggested that a specific type of tree ensemble acts by forming a locally adaptive distance metric. We generalize this claim to include all tree ensemble methods and argue that this insight can help to explain the exceptional performance of tree ensemble methods.

Committee: B. Steele (Chair), J. Graham, S. Harrar, J. Johnson (Computer Science), and D. Patterson

Thursday, May 7, 2009  
12:10 – 2:00 pm in Math 211