

## Colloquium

# Juanjuan Fan and Richard Levine

San Diego State University

Wednesday, April 08, 2015 at 3pm in Bell Hall 143

### **Part I** Predicting Glaucoma Progression using Decision Trees and Random Forest

Classification tree method is extended to correlated binary data and applied to the problem of developing objective prognostic classification rules in ophthalmology research. The robust  $z$  statistic from generalized estimating equation (GEE) is used as the splitting statistic to measure the between-node difference while adjusting for correlation among the fellow eyes for the same patient. The proposed method is assessed through simulations conducted under a variety of model configurations and illustrated using the perimetry and psychophysics in glaucoma (PPIG) study data. Furthermore, a random forest is constructed using longitudinally collected visual field data. In addition to producing rankings of variable importance, the random forest is also able to predict glaucoma progression with much improved sensitivity and specificity for the test sample.

### **Part II** Assessing Instructional Modalities: Individualized Treatment Effects for Personalized Learning

Learning analytics/educational data mining have become buzzwords of the day as "big-data" methods increasingly allow us to assimilate and analyze data on students to understand learners and learning environments. In this talk, we will briefly introduce and discuss our take on learning analytics. We will then introduce an "individualized treatment effect" as a means to quantify individual student performance under different instructional modalities or intervention strategies. The machine learning approach is motivated and illustrated by two recent studies on pedagogical reforms in statistics instruction through the SDSU ITS Course Design Institute (CDI) and the CSU Promising Course Redesign program. The studies consider student success in an online offering of a core applied statistics course and pedagogical reforms in an excess enrollment elementary statistics bottleneck course, respectively. Each of these studies leverage SDSU student information, learning management system, and course-specific assessment data towards studying student success relative to the redesign efforts, characterizing the successful student, and creating an analytics infrastructure for identifying at-risk students.

---

For further information, please contact Dr. Emil Schwab, [eschwab@utep.edu](mailto:eschwab@utep.edu)  
Refreshments will be served in front of the colloquium room,  
15 minutes before the start of the colloquium.