Traveling Course

Council of Chapters

Course Title:
Tree-Based Machine Learning Methods for Prediction and Variable Selection

Instructors:
Hemant Ishwaran and Min Lu, University of Miami

Course Length:
Half day and Virtual Only

Abstract:
Tree-based machine learning methods o□er several benefits in data analysis, including non-linearity, robustness, scalability and handling mixed data types. This course emphasizes practical learning with hands-on code examples and result interpretations, which is essential for understanding and applying these techniques. Based on the widely popular R package "randomForesSRC", we will present methods for computing predicted outcomes, variable importance indices and other inference estimates. In addition, we will introduce a new model-independent variable selection method, called the rule-based variable priority, and present its implementation using the R package "varPro". For all these analyses, we will cover different types of outcomes including continuous, categorical, multivariate, survival and competing risk outcomes. Utilizing real-world datasets from medicine and public health, topics in these analyses will provide hands-on code, working examples and result interpretations. We will provide additional code for visualizing model results and constructing coefficient tables for interpretation, and address scenarios such as imbalanced classes, unsupervised problems, fast implementation on big data and protection of confidential data.