

1135-VT-3062 **Daniel Brumley***, dbrumley1@uco.edu, and **Tyler Cook**. *Pathway and Gene Selection with Guided Regularized Random Forests*.

Many approaches have been developed in order to model a biological outcome based on microarray data. Much focus has recently been given to incorporating gene interactions via genetic pathway information available in online databases. The additional knowledge of gene relationships may help researchers better understand the biological processes under investigation. In this talk, we outline a method for pathway and gene selection based on guided regularized random forests (GRRF) that allows for the ranking of both pathways and genes in classification problems. In GRRF, variable importance scores from a random forest guide a regularization procedure to identify a subset of significant predictors. Simulation studies, as well as an analysis of a breast cancer dataset, show that our methodology is successful in identifying a compact set of important pathways and genes with a low prediction error rate. (Received September 26, 2017)