# Statistics Seminar

Department of Mathematics and Statistics

| | |
|---|---|
| **DATE:** | Thursday, April 10, 2025 |
| **TIME:** | 1:15pm – 2:15pm |
| **LOCATION:** | WH 100E |
| **SPEAKER:** | Tanzy Love, University of Rochester |
| **TITLE:** | Grafted and Vanishing Random Subspaces in Ensemble Classification |


**Abstract**

The Random Subspace Method (RSM) is an ensemble procedure in which each constituent learner is constructed using a randomly chosen subset of the data features. Regression trees are ideal candidate learners in RSM ensembles. By constructing trees upon different feature subsets, RSM reduces correlation between trees resulting in a stronger ensemble. Furthermore, it lessens computational burden by only considering a subset of the features when building each tree. Despite its apparent advantages, RSM has a notable drawback. In some instances, a randomly chosen subspace may lack informative features. This is especially true in situations in which the number of truly informative variables is small relative to the total number of variables. Trees that are constructed using feature subsets lacking informative features can be damaging to the ensemble. Here we present Grafted Random Subspaces (GRS) and Vanishing Random Subspaces (VRS), two novel ensemble procedures designed to remedy the aforementioned drawback by reusing information across trees. Both techniques borrow from RSM by growing individual trees on randomly selected feature subsets. For each tree in a GRS ensemble, the most important variable is identified and guaranteed inclusion into the next q feature subsets. This allows GRS to recycle a promising feature from one tree across several successive trees, effectively grafting the variable into the next q active subsets. In the VRS procedure the least important feature is guaranteed exclusion from the next q feature subsets. This creates a more enriched pool of candidate variables from which the successive feature subsets are drawn.


**Bio**

Dr. Tanzy Love is a Bayesian Statistician and an Associate Professor of Biostatistics and Computational Biology at the University of Rochester. She has also held visiting positions at Carnegie Mellon University and Stanford University. Dr. Love has published more than 100 papers in refereed books and journals and has served as

primary dissertation adviser for eight PhD students to date. Dr. Love is currently President of the Rochester Chapter of the American Statistical Association. She is also past President of The Classification Society, an international and interdisciplinary organization devoted to the scientific study of classification and clustering and a member of the International Federation of Classification Societies. Dr. Love has been teaching statistics for 25 years and enjoys finding new ways to communicate these ideas. During her spare time, Tanzy is an acrobat who enjoys aerial arts and has been featured in past Bayesian cabaret performances.