

Statistics and Data Science Seminar

RaSE: Random Subspace Ensemble Classification

Yang Feng (New York University)

Abstract: We propose a new model-free ensemble classification framework, Random Subspace Ensemble (RaSE), for sparse classification. In the RaSE algorithm, we aggregate many weak learners, where each weak learner is a base classifier trained in a subspace optimally selected from a collection of random subspaces. To conduct subspace selection, we propose a new criterion, ratio information criterion (RIC), based on weighted Kullback-Leibler divergences. The theoretical analysis includes the risk and Monte-Carlo variance of RaSE classifier, establishing the weak consistency of RIC, and providing an upper bound for the misclassification rate of RaSE classifier. An array of simulations under various models and real-data applications demonstrate the effectiveness of the RaSE classifier in terms of low misclassification rate and accurate feature ranking. The RaSE algorithm is implemented in the R package RaSEn on CRAN. This is joint work with Ye Tian.

Wednesday, November 11 at 4:00 PM in Zoom