Speaker: Yang Feng, Columbia University

Title: Neyman–Pearson classification algorithms and NP receiver operating characteristic (NP-ROC)

Abstract:

In many binary classification applications, such as disease diagnosis and spam detection, practitioners commonly face the need to limit type I error (i.e., the conditional probability of misclassifying a “normal”, or class 0, observation as “abnormal”, or class 1) so that it remains below a desired threshold. To address this need, the Neyman–Pearson (NP) classification paradigm is a natural choice; it minimizes type II error (i.e., the conditional probability of misclassifying a class 1 observation as class 0) while enforcing an upper bound, $\alpha$, on the type I error. Although the NP paradigm has a century-long history in hypothesis testing, it has not been well recognized and implemented in statistical classification schemes. Common practices that directly limit the empirical type I error to no more than $\alpha$ do not satisfy the type I error control objective because the resulting classifiers are still likely to have type I errors much larger than $\alpha$. As a result, the NP paradigm has not been properly implemented for many classification scenarios in practice. In this work, we develop the first umbrella algorithm that implements the NP paradigm for all scoring-type classification methods, including popular methods such as logistic regression, support vector machines and random forests. Powered by this umbrella algorithm, we propose a novel evaluation metric for classification methods under the NP paradigm: NP receiver operating characteristic (NP-ROC) bands, an extension of the popular receiver operating characteristic (ROC) curves. NP-ROC bands will serve as a new effective tool to evaluate, compare and select binary classifiers that aim to control type I error. We demonstrate the use and properties of the NP umbrella algorithm and NP-ROC bands, available in the R package nproc, through simulation and real data case studies.