

Statistics Seminar
Department of Mathematical Sciences

DATE:	Thursday, February 25, 2016
TIME:	1:15pm to 2:15pm
LOCATION:	WH 100E
SPEAKER:	Xingye Qiao, Binghamton University
TITLE:	Stabilized Nearest Neighbor Classifier and Its Statistical Properties

Abstract

The stability of statistical analysis is an important indicator for reproducibility, which is one main principle of scientific method. It entails that similar statistical conclusions can be reached based on independent samples from the same underlying population. In this paper, we introduce a general measure of classification instability (CIS) to quantify the sampling variability of the prediction made by a classification method. Interestingly, the asymptotic CIS of any weighted nearest neighbor classifier turns out to be proportional to the Euclidean norm of its weight vector. Based on this concise form, we propose a stabilized nearest neighbor (SNN) classifier, which distinguishes itself from other nearest neighbor classifiers, by taking the stability into consideration. In theory, we prove that SNN attains the minimax optimal convergence rate in risk, and a sharp convergence rate in CIS. The latter rate result is established for general plug-in classifiers under a low-noise condition. Extensive simulated and real examples demonstrate that SNN achieves a considerable improvement in CIS over existing nearest neighbor classifiers, with comparable classification accuracy. We implement the algorithm in a publicly available R package `snn`. This is a joint work with Wei Sun and Guang Cheng.

From:

<http://www2.math.binghamton.edu/> - **Department of Mathematics and Statistics, Binghamton University**

Permanent link:

<http://www2.math.binghamton.edu/p/seminars/stat/160225>

Last update: **2016/05/02 02:45**

