Statistics Seminar

Department of Mathematical Sciences

DATE:	Thursday, May 09, 2019
TIME:	12:00pm – 1:30pm
LOCATION:	WH-100E
SPEAKER:	Wenbo Wang, Binghamton University
TITLE:	Dissertation Defense - Set-valued Classification with Confidence Set Learning

Abstract

Unlike traditional single-valued classifier which gives one class label as the prediction for the true class of an observation, a set-valued classifier gives a set of class labels as its prediction. Hence, it can potentially be more useful in the situation where an accurate decision is needed but impossible to reach by single-valued classifiers.

In this dissertation, we study set-valued classification using confidence set learning framework. In this framework, one searches for different subsets, named confidence sets, of the input space. Each confidence set covers the population subclass j with some pre-determined rate. These confidence sets naturally induce a set-valued classifier. We aim to find confidence sets with the smallest overlap. Existing approaches to find confidence sets are based on the plug-in principle, which may be suboptimal for complex and high-dimensional data. We propose an empirical risk minimization method to solve this problem, both in the binary case and in the multicategory case. In particular, we make use of the support vector machine as the technical tool under this framework. Theoretically, we show that the proposed learner can control the non-coverage rates and minimize the overlap with high probability. Efficient algorithms are developed and numerical studies illustrate the effectiveness of the proposed method.

We further extend the confidence set learning framework to allow for anomaly detection. The resulting classifier is aware of possible anomalous observations and aims to detect them when possible. We give an introduction to the framework and study its connection with other existing frameworks for anomaly detection and confidence set learning.