



THE UNIVERSITY OF CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

SHIVANI AGARWAL

Department of Computer Science and Automation
Indian Institute of Science

Statistical Learning in Complex Prediction Spaces: What Do We Know?

MONDAY, October 20, 2014, at 4:00 PM
Eckhart 133, 5734 S. University Avenue
Refreshments following the seminar in Eckhart 110

ABSTRACT

In an increasing number of application domains, supervised learning methods are needed for making predictions in complex spaces: predicting sequences in speech recognition, predicting rankings in information retrieval and recommender systems, predicting assignments in image segmentation, and so on. How should we design statistically consistent learning algorithms for such problems? A popular learning approach is to minimize a convex surrogate loss, but how does one design surrogates that lead to consistent algorithms? While the answer to this question is well understood for simple binary classification tasks and a few other specific learning problems, a general understanding has been missing. In this talk I will describe some of our recent work on developing a unified framework for designing statistically consistent learning algorithms for complex prediction problems defined by an arbitrary loss matrix. I will introduce the notion of convex calibration dimension of a general loss matrix, which measures the smallest dimension of a surrogate space in which one can design a convex calibrated surrogate loss for a given loss matrix, and will describe some ways to design explicit convex calibrated surrogates for any given loss matrix. I will also discuss ways to achieve approximate statistical consistency when exact consistency may be computationally hard to achieve. I will conclude by describing some fundamental open questions.