The University of Chicago
Department of Statistics
Practice Job Seminar

JIE YANG
Department of Statistics
The University of Chicago

"Exchangeable Cluster Processes for Classification Problems"

Thursday February 16, 2006 at 12:00 pm
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Abstract: In this talk, we introduce two families of exchangeable cluster processes for classification problems. The first family called the Gauss-Ewens cluster process is generated from the Fisher discriminant model by a standard Dirichlet allocation scheme. It permits a new unit to be assigned to a class that has not been observed. The second family is based on the permanent process introduced recently. In the stochastic classification model determined by the permanent cluster process, no more than 4-5 parameters need to be estimated, regardless of the number of classes or the dimension of the feature space. The model works well even if a class occupies a non-convex region or disconnected regions in the feature space. Under the permanent model, the conditional distribution of a subsequent unit given the training data is expressed in terms of ratios of weighted permanents. We propose analytic approximations for the permanent ratios which is reasonably accurate for typical classification problems.

Information about building access for persons with disabilities may be obtained in advance by calling the department office at (773) 702-8333.