Classification Using Localized Spatial Depth with Multiple Localization

Subhajit Dutta
Indian Statistical Institute, India

Abstract

In the recent past, data depth has been considered by several authors as an effective methodology for supervised and unsupervised classification problems. However, most of the depth based classifiers studied in the literature require the population distributions to be elliptic and unimodal differing only in their locations in order to have satisfactory performance. Another limitation of such classifiers is that they usually require equal prior probabilities for the populations. Further, for many choices of the well-known depth function, practical implementation of depth based classifiers becomes computationally prohibitive even for moderately large dimensional data. In this talk, we propose a new classifier based on spatial depth, which can be used for high-dimensional. The main idea behind the construction of the proposed classifier is based on fitting generalized additive models to the posterior probabilities corresponding to different classes. In order to cope with possible multimodal and/or non-elliptic nature of the population distributions, we develop a localized version of spatial depth and use that with varying degrees of localization to build the classifier. Our classifier is formed by aggregation of several classifiers each of which is based on spatial depth with a fixed level of localization. This new classifier can be conveniently used for high-dimensional data, and its possess good discriminatory power for such data. Using some real benchmark data sets, the proposed classifier is shown to have competitive performance when compared with well-known and widely used classifiers like those based on nearest-neighbors, kernel density estimates, support vector machines, classification trees, artificial neural nets, etc.

(Joint work with P. Chaudhuri and A. K. Ghosh)