Pritchard, David A.; Liu, Yufeng
Composite quantile-based classifiers. (English) Zbl 07260684

Summary: Accurate classification of high-dimensional data is important in many scientific applications. We propose a family of high-dimensional classification methods based upon a comparison of the component-wise distances of the feature vector of a sample to the within-class population quantiles. These methods are motivated by the fact that quantile classifiers based on these component-wise distances are the most powerful univariate classifiers for an optimal choice of the quantile level. A simple aggregation approach for constructing a multivariate classifier based upon these component-wise distances to the within-class quantiles is proposed. It is shown that this classifier is consistent with the asymptotically optimal classifier as the sample size increases. Our proposed classifiers result in simple piecewise-linear decision rule boundaries that can be efficiently trained. Numerical results are shown to demonstrate competitive performance for the proposed classifiers on both simulated data and a benchmark email spam application.

MSC:
62-XX Statistics
68-XX Computer science

Keywords:
classification; high-dimensional data; quantile-based classifier; supervised learning

Full Text: DOI