Nonparametric Discrimination with Kernel Estimators in Microarray Analysis

Pedross FA[1], Wrulich OA[2], Pfeiffer KP[1]
[1]Department for Medical Statistics, Informatics and Health Economics
[2]Department for Medical Biochemistry
I-Med: Innsbruck Medical University

In recent years rapid advances have been made in bioinformatics. Especially sequencing of the human genome provides a considerable amount of information in disease research. Gene expression levels of two different data sets, gained by microarray analysis, are investigated. The first microarray analysis concerns different leukemia types and the second pertains to prostate carcinoma patients in comparison to healthy persons. We try to find a good discrimination between the different findings. Statistical techniques, that can handle with situations, where the number of variables beats by far the number of experiments, are needed. We use nonparametric predictive discriminant analysis with kernel estimators. For reducing the error adaptive kernel estimators will be applied. In addition we will introduce a boosting algorithm for kernel density estimators, which re-weights the data using a log -odds ratio based on the leave-one-out estimator. Boosting is justified because it can be shown that the bias can be reduced considerably. Different resampling techniques are tried to minimize misclassification rates.