

Title:

Pre-Validation and Inference in Microarrays

Author(s):

Robert Tibshirani and Bradley Efron

Technical Report number (Dept. of Statistics, Stanford, Univ.):

221

Date:

June 2002

Abstract:

In microarray studies, an important problem is to compare a predictor of disease outcome derived from gene expression levels to standard clinical predictors. Comparing them on the same dataset that was used to derive the microarray predictor can lead to results strongly biased in favor of the microarray predictor. We propose a new technique called “pre-validation” for making a fairer comparison between the two sets of predictors. We study the method and analytically and explore its application in a recent study on breast cancer.