SELECTING GENES IN DNA MICROARRAY EXPERIMENTS

<u>M. del C. Romero¹</u>, J. A. Di Rienzo²

¹Universidad Nacional del Centro de la Provincia de Buenos Aires, Buenos Aires, Argentina ²Universidad Nacional de Córdoba, Córdoba, Argentina

Email: romero@econ.unicen.edu.ar

The use of FILTERS in the context of supervised classification is a common practice for attribute selection. These techniques produce a scoring of attributes according to their estimated RELEVANCE in discriminating between classes ("treatments"). A DNA microarray experiment is a paradigmatic example for the application of attribute selection techniques. The recognition of differentially expressed genes across the "treatments" is one of the main goals of these assays. Although the application of filters produces a sorted list of the most relevant genes, the problem of how many genes will be retained remains along with the problem of minimizing false discoveries. We evaluate, on simulated gene expression matrix the filter RELIEF, and explore a number of non-parametric strategies in order to decide on the number of genes to be selected from the sorted list generated by the filter. We compare these results with those of selecting genes according to p-values, corrected to minimize false discovery rates, obtained from classical statistical tests.