

The role of cluster analysis in the study of gene expression arrays

Andrew B. Nobel

Department of Statistics

University of North Carolina, Chapel Hill

Abstract:

Clustering methods are playing an essential role in the analysis and application of emerging gene expression technologies such as microarrays. The intense interest in these technologies is, in turn, driving many important new (or rediscovered) developments in clustering methodology. These developments include resampling techniques such as bootstrapping and bagging, assessment of cluster confidence, computation of p-values, and new approaches to variable selection. This talk will describe some of this new methodology, in the context of recent work on identifying breast cancer subtypes using cDNA microarrays. I will argue that the continuing development of clustering methodology, and the complementary study of theoretical foundations, constitute important directions for future research.