Gene Shaving Algorithm as a Supervised Method for Data Classification

Georgiana-Andreea PETRUS

Abstract

Large gene expression studies often provide millions of different pieces of data. The aim of this paper is to test a statistical algorithm for microarray data, one called Gene Shaving. The method identifies subsets of genes with coherent expression patterns and large variation across conditions. This method differs from hierarchical clustering and other widely used methods for analyzing gene expression studies in that it is a supervised method that uses known properties of the genes or samples to assist in finding meaningful groupings. In this paper it will be illustrated the use of the Gene Shaving method to analyze gene expression measurement in cancer treatment. The method identifies a cluster of genes whose expression is highly predictive of survival. The Gene Shaving method is a potentially useful tool for exploration of gene expression data and identification of interesting clusters of genes.