

EnumTree: An Enumerative Biclustering Algorithm for DNA Microarray Data

Authors : Haifa Ben Saber, Mourad Elloumi

Abstract : In a number of domains, like in DNA microarray data analysis, we need to cluster simultaneously rows (genes) and columns (conditions) of a data matrix to identify groups of constant rows with a group of columns. This kind of clustering is called biclustering. Biclustering algorithms are extensively used in DNA microarray data analysis. More effective biclustering algorithms are highly desirable and needed. We introduce a new algorithm called, Enumerative tree (EnumTree) for biclustering of binary microarray data. is an algorithm adopting the approach of enumerating biclusters. This algorithm extracts all biclusters consistent good quality. The main idea of EnumLat is the construction of a new tree structure to represent adequately different biclusters discovered during the process of enumeration. This algorithm adopts the strategy of all biclusters at a time. The performance of the proposed algorithm is assessed using both synthetic and real DNA micryarray data, our algorithm outperforms other biclustering algorithms for binary microarray data. Biclusters with different numbers of rows. Moreover, we test the biological significance using a gene annotation web tool to show that our proposed method is able to produce biologically relevent biclusters.

Keywords : DNA microarray, biclustering, gene expression data, tree, datamining.

Conference Title : ICADMA 2015 : International Conference on Advanced Data Mining and Applications

Conference Location : Barcelona, Spain

Conference Dates : August 17-18, 2015