COMPARISON OF CLUSTERING ALGORITHMS: AN EXAMPLE WITH PROTEOMIC DATA

Nairanjana Dasgupta[1], Yibing Chen[1], Ananth Kalyanaraman[2] and Sayed Daoud[3]

[1]Department of Mathematics
 Washington State University
 U. S. A.

[2]Department of Computer Sciences
 Washington State University
 U. S. A.

[3]Department of Pharmaceutical Sciences
 Washington State University
 U. S. A.

Abstract

Clustering has become the norm in analyzing genomic and proteomic data. There are various algorithms for which graphs and software are easily available. However, often there is little agreement among these methods; prompting the question of which algorithm should one choose? In this manuscript, we consider eleven different methods of clustering, including partitioning, hierarchical and model based methods, to look at proteomic data from a colon cancer study. We use a pairwise index to evaluate these eleven clustering methods by comparing the clustering results. In an effort to comprehensively