ABSTRACT

Microarray technology has enabled the monitoring of expression levels of thousand of genes across different experimental conditions. The use of microarrays in many biological systems has resulted in a large amount of data. The most common technique to organize and analyze microarray data is to perform cluster analysis. Identification of group of genes that show similar expression patterns is a key step in the analysis of gene expression data. This report presents a comparison of the well known clustering approaches with application to gene expression data. In this report we propose a modified clustering algorithm for gene expression datasets. Experiments on this algorithm with real life datasets show that it has a better performance compared to other well known clustering algorithms, with respect to cluster evaluation measures.