Abstract

Microarrays are made it possible to simultaneously monitor the expression profiles of thousands of genes under various experimental conditions. Identification of co-expressed genes
and coherent patterns is the central goal in microarray or gene expression data analysis and is an important task in Bioinformatics research. In this paper, K-Means algorithm hybridised with Cluster Centre Initialization Algorithm (CCIA) is proposed Gene Expression Data. The proposed algorithm overcomes the drawbacks of specifying the number of clusters in the K-Means methods. Experimental analysis shows that the proposed method performs well on gene Expression Data when compare with the traditional K- Means clustering and Silhouette Coefficients cluster measure.

Reference

- Chen Zhang and Shixiong Xia, 2009 “K-Means Clustering Algorithm with Improved Initial center,” in Second International Workshop on Knowledge Discovery and Data Mining (WKDD), pp. 790-792.
- Daxin Jiang, Jian Pei, and Aidong Zhang, October 2005 “An Interactive Approach to mining Gene Expression Data”. IEEE Transactions on Knowledge and Data Engineering, Vol 17, No.10, pp.1363-1380.
- K. A. Abdul Nazeer and M. P. Sebastian, July 2009 “Improving the accuracy and


**Index Terms**

Computer Science

Data Mining

**Key words**

Clustering

CCIA

K-Means

Gene expression data