AN UNSUPERVISED CLUSTERING ALGORITHM APPLIED FOR GENE SELECTION MICRO ARRAY DATA

G. BASKAR[1] & P. PONMUTHURAMALINGAM[2]

[1]Research Scholar, Department of Computer Science, Government Arts College (Autonomous),

Coimbatore, Tamil Nadu, India

[2]Associate Professor & Head, Department of Computer Science, Government Arts College,

Coimbatore, Tamil Nadu, India

ABSTRACT

To explore data-mining techniques as an preset means of reducing the difficulty of data in the large Bioinformatics database and of discovering useful patterns and relationships in data. Data mining isn't an end point, but in one period is an overall data discovery process. It is an iterative process in which preceding processes are modified to support new hypotheses suggested by the data. The stability of feature selection has recently become a topic of strong in both the machine learning and in Bioinformatics community, Feature selection is a term frequently used in data mining for decreasing input to a manageable size for processing and analysis. Micro array data is a commonly used technique for choosing candidate gene in various cancer studies. In this paper, we proposed clustering algorithm on cancer data set, with time, accuracy and memory space.

KEYWORDS: Clustering, Feature Selection, Micro Array, Stability