Abstract

Clustering is one of the most useful tasks in data mining process for discovering groups and identifying interesting patterns in the underlying data. Cluster analysis is a widely used technique today in fields like Healthcare, Sociology and Biology etc. to identify the patterns from a huge amount of data. It has many applications such as image segmentation, information retrieval, web pages grouping, market segmentation, and scientific and engineering analysis. This paper gives an overview of cluster analysis. It describes all the preliminaries required for the process and cites the main algorithms of clustering with their pros and cons.

References


Index Terms

Computer Science Information Sciences

Keywords

Clustering algorithms, Cluster Analysis, Hierarchical clustering, Partitional Clustering.