

# Data Mining Process Using Clustering: A Survey

Mohamad Saraee	Najmeh Ahmadian	Zahra Narimani
Department of Electrical and Computer Engineering Isfahan University of Technology, Isfahan , 84156-83111 <a href="mailto:saraee@cc.iut.ac.ir">saraee@cc.iut.ac.ir</a>	Department of Electrical and Computer Engineering Isfahan University of Technology, Isfahan , 84156-83111 <a href="mailto:Najmeh_2020@yahoo.com">Najmeh_2020@yahoo.com</a>	Department of Electrical and Computer Engineering Isfahan University of Technology, Isfahan , 84156-83111 <a href="mailto:z.narimani@gmail.com">z.narimani@gmail.com</a>

## Abstract

Clustering is a basic and useful method in understanding and exploring a data set. Clustering is a division of data into groups of similar objects. Each group, called cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups. Interest in clustering has increased recently in new areas of applications including data mining, bioinformatics, web mining, text mining , image analysis and so on. This survey focuses on clustering in data mining.

The goal of this survey is to provide a review of different clustering algorithms in data mining. A Categorization of clustering algorithms has been provided closely followed by this survey. The basics of Hierarchical Clustering include Linkage Metrics, Hierarchical Clusters of Arbitrary and Binary Divisive Partitioning is discussed at first. Next discussion is Algorithms of the Partitioning Relocation Clustering include Probabilistic Clustering , K-Medoids Methods, K-Means Methods. Density-Based-Partitioning , Grid-Based Methods and Co-Occurrence of Categorical Data are other sections. Their comparisons are mostly based on some specific applications and under certain conditions. So the results may become quite different if the conditions change.

**KeyWords:** clustering, partitioning, unsupervised learning, hierarchical clustering