Abstract

In Data clustering (a sub field of Data mining), k-means and hierarchical based clustering algorithms are popular due to its excellent performance in clustering of large data sets. This paper presents two different comparative studies which includes various Data Clustering algorithms for analyzing best one with minimum clustering error. The foremost objective of this paper is to divide the data objects into k number of different clusters with homogeneity and the each cluster should be heterogeneous to each other. However, these both algorithms (K-Mean and Hierarchical) are not free with the errors. In this paper, firstly various distance has been considered for these two algorithms for comparing and analyzing the best distance methods to solve the existing problems.

References

Means Clustering with Effect of Distance Functions for Bank Dataset. &quot; IJITR1. 3 (2013): 233-235.
- Density-based clustering algorithms DBSCAN and SNN by Adriano Moreira, Maribel Y. Santos and Sofia Carneiro.
- J. L. Bentley, Multidimensional Binary Search Trees Used for Associative Searching.
- Euclidean Distance in http://people.revoledu.com/kardi/tutorial/Similarity/EuclideanDistance. html.

Index Terms

Computer Science
Information Sciences

Keywords
K-Means Hierarchical Euclidean Distance Manhattan Distance Filtering cluster Density Based clustering algorithms on clustering error.