Cluster identification is useful for finding insights into the huge datasets for finding out the attributes, characteristics of a particular dataset. Today, many organizations have started to use their own data analytic tools for finding clusters.

This paper focuses on various algorithms for finding clusters for huge and different datasets. We have used different datasets and applied MapReduce algorithms for achieving the results. The experimental results obtained in substantial algorithmic computations provide clusters that are used for quick decision making. We present the results performed over various datasets that scales well with respect to both data set size and data set dimensionality.

References


27. Max Bodoia , “MapReduce Algorithms for k-means Clustering”.

**Index Terms**

Computer Science  
Artificial Intelligence

**Keywords**

Machine Intelligence, Dimensionality reduction.