Abstract

As more and more documents are available in the form of hypertext in world wide web, proper clustering of documents are required for generating the similar results fetched by the search engine specific to a user entered search query as the documents within the cluster are similar to each other. In Agglomerative approach of clustering the clusters of documents are merged into cluster unless until all the clusters belong to a root cluster. In this paper design and comparison of two agglomerative hierarchical clustering of documents is done on parameters such as cluster generation, relevance of result get against a query and purity of clusters.

References

3. Chris Staff: Bookmark Category Web Page Classification Using Four Indexing and

Index Terms

Computer Science  Artificial Intelligence

Keywords

Agglomerative, relevance, purity, cluster, search engine.