

Abstract

The existence of a document grouping is necessary for many fields of data technology. The data development is more increase and the need for rapid data processing group in real time the state became more developed.

Cover Clustering Coefficient Method (C2ICM) is one of the grouping or clustering algorithm a document that uses the model probabilistically, in common terms, and the seed document as an important point in determining the initial initialization of cluster formation, where classified documents have been processed in advance in order to be processed using C2ICM algorithm and meet the needs of the dynamic development of the data is always the case time. That process of real update is also called preprocessing or indexing of documents. In Indexing, there are some steps commonly performed, one stemming. At this final project, Jelita Asian algorithm used in phase before the document gets the next step to be clustered by C2ICM algorithm. The collection document is an abstract document from the final project of IT Telkom students.

Tests performed in this final documents are to clustered by C2ICM algorithm, analyze concepts and stages C2ICM Cover Coefficient Algorithm C2ICM and analyze the results of the quality of clusters produced by fusion with Asian Jelita C2ICM using the Silhouette Coefficient. And found that the quality of the resulting cluster is included in the quality of the weak by the average silhouette values obtained.

Keywords : *C2ICM, Jelita Asian, CBR , SilhouetteCoefficient , Clustering, Indexing*