

ABSTRACT

Text mining can be defined as a process of information gathering in which a user interacts with a set of documents using the tools of analysis are the components of data mining.

One important task in text mining is clustering. Clustering is the process of partitioning a set of objects into clusters. Similar objects will be placed in the same cluster and different clusters will be placed in different clusters.

This final task attempt to classify documents using agglomerative nesting (AGNES). The document is a document used in Indonesian language news, because the needs of information are increase. The linkages between the news is measured based on the similarity between documents. The similarities between the news is calculated from the frequency of occurrence of words in the document using the cosine similarity.

After the testing with several min Similarity, the results shows that agglomerative nesting can be used to group Indonesian news. This is evident from the results of the average f-measure generated approximately 97.4%.

Key words : Text mining, clustering, agglomerative nesting, minSimilarity