Abstract

Today, needing for informations was becoming one of the most important thing. Informations come from electronic and nonelectronic source, one of this is news. Internet is a most effective medium for distributing of informations. Now news can be obtained more easily through internet, in many sites. So, the number of news which is available in become greater and it is difficult for the readers getting the right news that needed.

This final project try to group Indonesian news. Document clustering is an alternative for grouping news. Document clustering is a process to group documents, so that document which has many similarities with other document in same cluster and many dissimilarities with other document in different cluster. The algorithm that implemented in this final project is K-means clustering.

Analysis concerning to stemming that is used in preprocessing to analyse connection between stemming, and the quality of cluster by use purity. The best result from clustering is combining from stemming.

Keywords: *clustering, Indonesian news, K-means, stemming, purity.*