Abstract

Automatic Text Summarization is a computer based application that is able to summarize a text or some text with taking or filtering the most important information in it. Ability of the application is it can produce a summary which can be directly used without re-editing process. There are two types of text summarization: abstraction and *extraction*. The abstraction type can summarize text by involving paraphrase of the text. While the *extraction* type just copying the information that is considered important from the text.

On this final project, the automatic text summarization implemented *extraction* approaches using *Maximal Marginal Importance* (MMI) where this method do weighting values to sentence features in the text which previously the preprocessing is done. After each sentence has weight values, the sentences were grouped using *k-means clustering* algorithm based on the *similarity* value between two sentences. From each sentence cluster, create a sentence binary tree / more. Creation of a sentence binary tree is intended to find out at where level of sentences from each cluster is located. Where the level of the sentence was later included in the calculation value of MMI sentence. Sentence that has the highest MMI value on each cluster was chosen as the summary result. At the last process, order the index of summary sentences if it has not been sorted.

The evaluation performed using the *ROUGE* evaluation toolkit. With *ROUGE*, measurement of the accuracy of summary result (candidate summary) with reference summary based on parameter recall, precision and f-measure. "Result of evaluation shows that automatic text summarization create summary that is get similar with reference summary based on the value of parameter *ROUGE*.

Keywords : automatic text summarization, extraction, MMI, similarity, k-means clustering, ROUGE