

Abstract

Automatic Text Summarization is a process of extracting important information from one or more text document. The purpose is to produce a collection of texts with shorter length, so they can be used by user by utilizing the computer-based application.

In this Final Test, Trimmer Method is used to do extraction by removing some part of the sentences with some way so it will not remove the important information on the sentence. The method on this Final Tes is implemented to Indonesia news article. For detecting the element of each sentence, first the sentence is parsed into its parts by changing it's form into a parse-tree. This process is applied by using PC-PATR as the sentence parsing system. Trimmer Method is implemented by applying some rules into each sentences. Sentence weighting use the concept of Cosine Similarity to decide which sentence should have the highest weight and be the candidate of the summary. This concept is applied by the Java Lucene Class that is also used in this Final Task. To avoid redundancy and also sort the candidate, a Dynamic Weighting Score , the Redudancy Score is used.

Testing is done by using Rouge evaluation toolkit. The result shows that this method hasn't show a good accuracy value. Besides that, it also shows that without any rate compression from user, the system has already able to produce the whole summary result with less word amount than its reference. It needs improvement on some rules and also deeper research by using another kind of text document to improve the accuracy of the summary result.

Keywords: Text Summarization, *Trimmer*, *Cosine Similarity*, *Redudancy Score*, *parse-tree*