Abstract

Information Overload is a condition where too much information that is obtained and resulted in some difficulty in understanding the problem and the information obtained. One example that can be seen is in the news currently distributed online that can be issued by a variety of different sources, so there will be a lot of news documents. To facilitate understanding of the information on the lot of news can be overcome by doing a document summarization to take the news information. Document summarization can be defined as taking important sentence or main sentence from the document without losing the meaning that contained in the document. Conditional Random Fields (CRF) is a probabilistic model to overcome the segmentation and labeling sequence data, which is influenced by the conditions in the preceding sentence and the value of the sentence itself is derived from the value of the feature, either complex features or basic features of each sentence. Summarization is done by extract-summary category, which the results of the summarization use sentence that have been extracted from the document. Accuracy of the system is measured by using ROUGE-2 (Recall-Oriented Understudy for Gisting Evaluation), by comparing summaries created by the system and summaries created by humans.

Keywords : Information Overload, Conditional Random Fields, feature, extract-summary, basic feature, complex feature, ROUGE-2