

ABSTRACT

Automatic text summarization plays an important role in information retrieval and text classification, and may provide a solution to the information overload problem. Text summarization is a process where the size of the input document is reduced without changing the meaning of the document. There are various ways for auto text summarization. In our project we adopt extractive based summarization using sentence clustering. To find out the clusters the basic need is the sentence similarity function. In this approach each cluster head represent the topic sentence which is very much similar to the other sentences of the cluster. In this approach summarization result is not only depends on the words of the sentences, but also depends on the sentence semantic similarity. We used Assamese word net to find the semantic similarity. Our method is based on [1] where author implement the idea on Chinese language. Since the language structure of Chinese and Assamese are different we need to modify the similarity function, which is the core part of the whole approach. Besides the clustering approach we also made one experiment using another technique which is later on merged with the clustering approach to achieve more efficient result.