Abstract

Summarization is the process of reducing a text document with a computer program in order to create a summary that retains the most important points of the original document. Technologies that can make a coherent summary take into account variables such as length, writing style and syntax. The main idea of summarization is to find a representative subset of the data, which contains the information of the entire set. Text summarization is commonly used to handle summaries of email threads, action items from a meeting and simplifying text by compressing sentences used to manage knowledge and also to help Internet search engines. This paper gives comparative study of various text summarization techniques used for Indian regional languages and also discusses in detail two main Types of Text summarization techniques these are extractive and abstractive.

References

1. Sunitha C, Dr. A Jaya and Amal Ganesh “Abstractive Summarization Techniques in
Indian Languages”, Peer-review under responsibility of the Organizing Committee of ICRTCSE 2016 doi: 10.1016/j.procs.2016.05.121, International Conference of recent trends in computer science.


18. Vishal Gupta and Gurpreet Singh Lehal, “Automated Punjabi text extraction summarization system”, Manuscript received January 12, 2010; revised March 22, 2010;


24. tf-idf https://en.wikipedia.org/wiki/Tf%E2%80%93idf

Index Terms

Computer Science

Artificial Intelligence

Keywords

NLP, text summarization, text summarization techniques, extractive, abstractive, features, Rich Semantic graph, TF-IDF, NLG.