Abstract

Summarization of text documents is increasingly important with the amount of data available on the Internet. The large majority of current approaches view documents as linear sequences of words and create query-independent summaries. However, ignoring the structure of the document degrades the quality of summaries. Furthermore, the popularity of web search engines requires query specific summaries. Here one method is used to create query-specific
summaries by adding structure to documents by extracting associations between their fragments. Query-oriented summarization is primarily concerned with synthesizing an informative and well-organized summary from a document collection for a given query. In the existing summarization methods, each individual sentence in the document collection is represented as Bag of Words (BOW).

Reference

- A Query Focused Multi Document Automatic Summarization Pinaki Bhaskar and Sivaji Bandyopadhyay, Department of Computer Science & Engineering, Jadavpur University, Kolkata – 700032, India
- Extractive Summarization Using Supervised and Semisupervised Learning by Kam-Fai Wong*, Mingli Wu, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong
- An Integrated Multi-document Summarization Approach based on Word Hierarchical Representation by You Ouyang, Wenji Li, Qin Lu, Department of Computing, the Hong Kong Polytechnic University.
- Document Topic Generation in Text Mining by Using Cluster Analysis with EROCK by Rizwan Ahmad, Computer Engineering Department, National University of Science & Technology, Rawalpindi.
- Ramakrishna Varadarajan School of Computing and Information Sciences, Florida International University, paper on “A System for Query-Specific Document Summarization”
- A paper on ROCK: A Robust Clustering Algorithm for Categorical Attributes by Sudipto Guha, Stanford University, Stanford, CA 94305.

Index Terms

Computer Science Ubiquitous Computing

Key words

Query-specific summarization user survey
keyword search