Abstract

The growing abundance of text articles in internet requires automated tagging using key phrases. The automated key phrase generation of resources helps in the information retrieval. To generate the key phrases for texts from all possible domains, the need is an automated approach that would extract the key ideas directly from the text itself. In this paper, we have suggested a methodology that uses the noun words and phrases, their occurrence and co-occurrences to generate the keywords. The method, employing both the statistical and linguistic features, has been successful in extracting the keywords and phrases to tag a text that best summarizes its content.

References

1. Ana Mestrovic, Beliga, Sanda Martincic-Ipsic, Slobodan, "An overview of graph-based keyword extraction methods and approaches", Journal of Information and Organizational
Unsupervised Key-phrase Extraction using Noun Phrases


6. Bruce, Chad Burkey, Krulwich, "Learning user information interests through extraction of semantically significant phrases", Proceedings of the AAAI spring symposium on machine learning in information access.


8. Cornelia Caragea, Sujatha Das Gollapalli, “Extracting Keyphrases from Research Papers Using Citation Networks”, University of North Texas


Index Terms

Computer Science Information Sciences

Keywords

Unsupervised Key phrase Extraction, Automated Tagging, Key phrase extraction using Nouns, Natural Language Processing, Information Retrieval.