

{tag}

{/tag}

IJCA Special Issue on International
Conference on Reliability, Infocom Technology and Optimization

© 2013 by IJCA Journal

ICRITO

Year of Publication: 2013

Authors:

Jitendra Nath Singh

Sanjay Kumar Dwivedi

{bibtex}icrito1306.bib{/bibtex}

Abstract

The vector space model is one of the classical and widely applied information retrieval models to rank the web page based on similarity values. The retrieval operations consist of cosine similarity function to compute the similarity values between a given query and the set of documents retrieved and then rank the documents according to the relevance. In this paper, we are presenting different approaches of vector space model to compute similarity values of hits from search engine for given queries based on terms weight. In order to achieve the goal of an effective evaluation algorithm, our work intends to extensive analysis of the main aspects of Vector space model, its approaches and provides a comprehensive comparison for Term-Count

Model, Tf-Idf model and Vector space model based on normalization.

Refer

ences

- Salton, G; Wong, A; Yang, C. S. : A vector space Model for automatic indexing, Communications of the ACM, Volume 18 and Issue 11:1975.
- Sanjay Kumar Dwivedi, Jitendra Nath Singh, and Rajesh Gotam Information Retrieval Evaluative Model, FTICT 2011: Proceedings of the 2011, International conference on Future Trend in Information & Communication Technology, Ghaziabad, India: 2011.
- Yi Shang Longzhuang Li: Precision Evaluation of Search Engines, World Wide Web: 2002.
- D. L. Lee, H. Chuang, and K. Seamons. Document ranking and the vector space model, IEEE Transactions on Software, 14(2): 1997.
- Chris Buckley. The importance of proper weighting methods, In M. Bates, editor, Human Language Technology. Morgan Kaufman: 1993.
- Longzhuang Li, Yi Shang A new statistical method for performance evaluation of search engines. ICTAI: 2000.
- Longzhuang Li, Yi Shang A new method for automatic performance comparison of search engines. World Wide Web: 2000.
- Chu, H. & Rosenthal: "Search engines for the World Wide Web: A comparative study and evaluation methodology". In Proceedings of the 59th Annual Meeting of the American Society for Information Science, Baltimore, 1996.
- Gerald Salton and Chris Buckley. Term weighting approaches in automatic text retrieval. Information Processing and Management, 24(5): Is-sue 5. 1988.
- G. Salton and C. Buckley, "Improving Retrieval Performance by Relevance Feedback," J. Amer. Soc. for Information Science, Vol. 41, No. 4, 199
- Jitendra Nath Singh & Sanjay Kumar Dwivedi: Analysis of Vector Space Model in Information Retrieval. Proceedings (IJCA) on National Conference on Communication Technologies & its impact on Next Generation Computing 2012 CTNGC (2):14-18:2012.
- Lee, D. I. ; Huei Chuang; Seamons, K. ; "Document ranking and the Vector-space model," Software, IEEE, vol. 14, no. 2, Pp. 67-75, Mar/April. 1997

Index Terms

Computer Science

Data Mining

Keywords

Vector Space Model Information Retrieval Tf-idf Term- Frequency Cosine Similarity