

## عنوان مقاله:

A New Approach to Improve the Accuracy of the TF\_IDF Ranking Algorithm in Text Retrieval

## محل انتشار:

چهارمین کنفرانس بین المللی وب پژوهی (سال: 1397)

تعداد صفحات اصل مقاله: 8

## نویسندگان:

Azize Nemati - Graduate Student, Department of Computer Engineering, Golestan University Gorgan

Soheila Karbasi - Assistant Professor, Department of Computer Engineering, Golestan University, Gorgan

## خلاصه مقاله:

Today, the World Wide Web is considered as the largest source of data with the help of Web search engines, as one of the most useful tools for extracting information. Due to the web growth, providing information related to user queries by search engines is very difficult. Also, the effectiveness of the information retrieval systems is largely dependent on term-weighting. Therefore, search engines use different web mining techniques to rank search results. For this purpose, various ranking algorithms are presented. In this research, the weighting algorithm TF\_IDF is used to rank the documents. By introducing the entropy parameter related to the number of user query words in the text of the documents, the accuracy of the ranking of the documents in the information retrieval is evaluated. The remarkable points obtained from the surveys on standard questions provide a new approach to increasing the efficiency of text search systems, which the responses from subsequent experiments demonstrate its validation. The proposed approach in this paper uses the Standard Web collections and the results show that it can significantly increase the accuracy of retrieval in terms of the volume of test data collection

## کلمات کلیدی:

Information retrieval, Web mining, TF\_IDF weighting model, Document scoring, Entropy

## لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/773309>

