TF-IDF

- description : TF-IDF
- author : 오션
- email : shlim@repia.com
- lastupdate : 2020-06-24

용어

TF-IDF : Term Frequency - Inverse Document Frequency (텀 프리퀀시 - 인버스 다큐먼트 프리퀀시), 단어 빈도와 역문서 빈도

의미

1. 정보 검색이나 텍스트 마이닝에서 주로 사용하는 것으로, 해당 엔티티(entity - 실체, 여기서는 '단어'의 의미)가 가지는 중요도를 파악할 수 있다.
2. TF - 단어빈도(수) : 특정 단어가 문서내에 얼마만큼의 빈도로 등장하는지를 나타낸다.
3. IDF - 역문서빈도(수) : 전체 문서 갯수를 해당 엔티티가 포함된 문서의 갯수로 나누는 것이며, 해당 문서군의 특징을 보여준다.

Ref

TF-IDF(Term Frequency - Inverse Document Frequency)/단어 빈도와 역문서 빈도

참조

빈도 : 같은 현상이나 일이 반복되는 도수, 빈도수
inverse : 역의, 정반대의
entity : 엔티티, 실체, 독립체

오션, TF-IDF, Term, Frequency, -, Inverse, Document, Frequency

From:
http://125.132.25.164/dokuwiki/ - **문제를 잘 정의하는 것은 문제를 절반 해결한 것이다.** - 2023.12

Permanent link:
http://125.132.25.164/dokuwiki/doku.php?id=wiki:miscellaneous:tf-idf&rev=1593007280

Last update: **2022/03/10 19:52**