

Recuperación de Información

8 febrero de 2016

1. (2 puntos) Describe el proceso de indexación en un Sistema de Recuperación de Información.
2. (2 puntos) Describe el modelo de RI basado en Modelado de Lenguaje.
3. (2 puntos) Ante una consulta, un SRI encuentra 20 documentos para los que tenemos los siguientes juicios de relevancia:
rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2) rr (2)
Suponiendo que en la colección hay un total de 10 documentos relevantes, calcular la exhaustividad, la precisión, $P@10$, $AP@10$ y la $NDCG@10$ del ranking obtenido.
4. (2 puntos) Algoritmo PageRank y su uso en la Recuperación de Información web.
5. (2 puntos) Algoritmo K-medias. Aplicación al problema de agrupamiento de documentos.

Notas

Tiempo para realizar el examen: 3 horas.