

Modelos de recuperación

Los modelos de recuperación de información son fundamentales para calcular la similitud y grado de relevancia entre la consulta y los documentos que están en el repositorio. Los modelos clásicos son: el modelo booleano, el vectorial y el probabilístico. Se han propuesto numerosas mejoras a estos modelos, con algoritmos. Desde la inteligencia artificial a las mejoras propuestas para la Web.

A lo largo del tema se muestran alternativas en la aplicación de estos algoritmos, estas alternativas tratan de mostrar cómo, dependiendo del caso, se pueden aplicar estos algoritmos con ligeras modificaciones.

En este tema también se examinan diversos métodos para determinar el peso de los términos, tanto de la consulta como del documento, a la hora de aplicar el modelo. Hay que tener en mente que no todos los términos de la consulta o del documento tienen el mismo poder discriminante, a la hora de determinar cuál debe ser el documento más relevante a la necesidad de información.

Por último, se muestran algunos de los problemas comunes a estos algoritmos y se facilita un ejercicio para que el alumno se familiarice con su aplicación.

