

Evaluación de MiniTREC tradicional

¿Se corresponden las medidas de evaluación globales de tu sistema con alguna de las columnas que se muestran en el cuadro 2, o las curvas globales de precisión-exhaustividad de la figura 1? ¿Con qué columna/curva?

El sistema del equipo 04 se corresponde con el sistema **m**, ya que coinciden sus valores tanto en el cuadro como en la gráfica.

¿Qué decisiones de diseño de tu sistema (campos indexados, analizadores elegidos, modelo de ranking, procesamiento especial de consultas, ...) justifican las medidas obtenidas? ¿Hay otros factores externos que pueden haber afectado a estos resultados?

Todas las características del sistema afectan a los resultados: campos indexados, analizador (stemming, stopwords), cada paso del procesamiento de las consultas...

Algún factor externo que puede haber afectado a los resultados son las valoraciones de los jueces, que han sido efectuadas en base a los resultados obtenidos por el sistema, pero siempre de forma subjetiva. Un cambio en la valoración de un solo documento puede generar valores distintos para las medidas de evaluación.

Según tu opinión, ¿Qué medida, o medidas, reflejan mejor la calidad de un sistema de recuperación?

Existen varias medidas que, de forma muy directa, reflejan la calidad de un sistema. Estas medidas pueden ser el score F1, la precisión a 10 o el MAP. Dependiendo del contexto, puede ser más informativo emplear unas métricas u otras, pero en este caso las tres evidencian conclusiones similares con respecto a la calidad de los sistemas evaluados.

En base a las medidas del cuadro 2, las curvas de la figura 1 y tu respuesta a la cuestión anterior, ¿Qué sistema que no sea el tuyo recomendarías utilizar?

En base a estas cuestiones, el sistema recomendado a utilizar sería el sistema **r**.