



6-12-2020

EVALUACIÓN SISTEMA DE RECUPERACIÓN TRADICIONAL

Recuperación de Información

Néstor Monzón González (735418)
Cristina Oriol García (755922)

Nuestro sistema de recuperación de información tradicional presenta las siguientes medidas globales, promediando las de cada una de las necesidades:

<i>precision</i>	<i>0.416</i>
<i>recall</i>	<i>0.375</i>
<i>F1</i>	<i>0.384</i>
<i>prec@10</i>	<i>0.640</i>
<i>MAP</i>	<i>0.591</i>

Los puntos de precisión-recall interpolados de nuestro sistema se pueden visualizar en la siguiente gráfica:

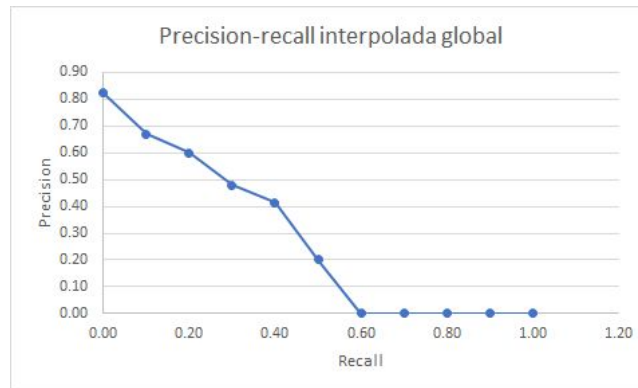


Figura 1: precision-recall interpolada global (para los cinco needs) de nuestro sistema

Como se puede observar en los resultados obtenidos es bastante similar al sistema F presentado. Todas las medidas son idénticas, excepto la precisión (que quizás se haya calculado de forma ligeramente distinta).

Las decisiones de diseño que pueden haber influido en los resultados son principalmente los pesos asignados a los distintos tipos de consulta dando demasiada importancia a algunos campos y ‘escondiendo’ otros y que no conseguimos crear en el sistema algún tipo de mecanismo para la detección de nombres propios y lugares.

En general, las medidas más completas son el F1_score y el MAP, ya que relacionan globalmente la precisión el recall. No obstante, la mejor medida de un sistema depende del uso que se le vaya a dar a este. Por ejemplo, para un buscador web quizás es más adecuada la precision@10, ya que rara vez un usuario tendrá en cuenta más de 10 resultados.

Por tanto, y teniendo en cuenta el resto de valores, el mejor sistema es probablemente el j, contando con el F1_score más alto y el segundo MAP. Además, presenta las precisiones más altas para cada recall (llevando a una mayor área bajo la curva).