

EVALUACIÓN DE SISTEMAS DE RI

Sea una colección de documentos con 40 documentos, identificados con los números de 1 al 40. Sabemos que los documentos relevantes para una determinada consulta son [2,5,7,14,15,16,19,23,34,39].

Un sistema de recuperación de información devuelve el siguiente resultado para la consulta:

S1= [5, 11, 2, 19, 14, 3, 35, 34, 33, 16, 1, 8]

Se pide:

- a) Calcular la eficacia (Precision, Recall y la F-medida con $\beta=1$) para la consulta.

Marcamos en negrita los documentos que son relevantes:

S1= [5, 11, **2, 19, 14**, 3, 35, **34**, 33, **16**, 1, 8]

DEFINICIONES:

Precision= nº de docs relevantes recuperados/ nº de docs recuperados = 6/12= 0,5

Recall= nº de docs relevantes recuperados/ nº de docs relevantes en la colección= 6/10= 0,6

$$F_1 = \frac{2PR}{P+R}$$

F-1= (2*0,5*0,6)/(0,5+0,6)=0,6/1,1=0,55 (redondeado a 2 decimales)

- b) Completar las Tablas de Precision y Recall (expresando la operación de división realizada y el resultado redondeado en dos decimales, p.e. 2/3 = 0,67) e Interpoladas.

Tabla Precision&Recall Reales

S1	1	2	3	4	5	6	7	8	9	10	11	12
Relevante	Y	N	Y	Y	Y	N	N	Y	N	Y	N	N
Precision	1/1= 1	1/2= 0.5	2/3= 0.67	3/4= 0.75	4/5= 0.8	4/6= 0.67	4/7= 0.57	5/8= 0.62	5/9= 0.55	6/19= 0.6	6/11= 0.54	6/12= 0.5
Recall	1/10= 0.1	1/10= 0.1	2/10= 0.2	3/10= 0.3	4/10= 0.4	4/10= 0.4	4/10= 0.4	5/10= 0.5	5/10= 0.5	6/10= 0.6	6/10= 0.6	6/10= 0.6

Tabla Precision&Recall Interpoladas

Precision S1	1	1	0.8	0.8	0.8	0.62	0.6	0	0	0	0
Recall	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

DEFINICIÓN: Precisión interpolada en 11 niveles de recall: Para cada valor de recall estándar i desde 0.0 a 1.0 con incrementos de 0.1, se toma la precisión máxima obtenida en cualquier valor de recall real mayor o igual a i.

- c) Calcula la precisión media.

DEFINICIÓN: Para una consulta simple la Precisión Media es el promedio del valor de precisión obtenido después de que cada documento relevante sea recuperado en la lista ordenada de documentos recuperados.

P_MEDIA_S1=(1+0.67+0.75+0.8+0.62+0.6+0+0+0+0)/10= 4.44/10= 0.44