

Acto1 – Preguntas 1 y 2

(Grupo1_Tarde – 1,4 puntos)

Apellidos y Nombre: Díaz-Alejo León, Stéphane

IMPORTANTE: todos los cálculos se mostrarán redondeados a dos decimales

- 1) Sea una colección de documentos con 100 documentos, identificados con los números de 1 al 100. Sabemos que los documentos relevantes para una determinada consulta son:

[3,5,18,22,35,40,41,63,80,89].

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

S1= [22, 81, 3, 26, 40, 7, 80, 5, 25, 19, 89, 76]

Se pide:

- a) Calcular la eficacia (Precision, Recall y la F-medida con $\beta=1$) para la consulta. **(0,3 puntos)**

Consulta	Precision	Recall	F-1
S1	0.5	0.6	0.55

- 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. **(0,6 puntos)**

Tabla Precision&Recall Reales

S1	1	2	3	4	5	6	7	8	9	10	11	12
Relevante	1	0	1	0	1	0	1	1	0	0	1	0
Precision	$1/1=1$	$1/2=0.5$	$2/3=0.67$	$2/4=0.5$	$3/5=0.6$	$3/6=0.5$	$4/7=0.57$	$5/8=0.63$	$5/9=0.56$	$5/10=0.5$	$6/11=0.55$	$6/12=0.5$
Recall	$1/10=0.1$	$1/10=0.1$	$2/10=0.2$	$2/10=0.2$	$3/10=0.3$	$3/10=0.3$	$4/10=0.4$	$5/10=0.5$	$5/10=0.5$	$5/10=0.5$	$6/10=0.6$	$6/10=0.6$

Tabla Precision&Recall Interpoladas

Precision S1	1	1	0.67	0.63	0.63	0.63	0.55	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

- c) Calcular la Precisión media del sistema. **(0,1 puntos)**

Precisión media = 0.4

2) Esta pregunta consta de dos apartados:

(0,4 puntos)

a) ¿Cómo sería el índice permuterm para la palabra “pato”?

pato\$, \$pato, o\$pat, to\$pa, ato\$p

b) ¿Cómo sería el mecanismo de búsqueda correspondiente a los wildcard queries “p*o” y “*ato” suponiendo que disponemos de este tipo de índice?

o\$p* y ato\$*