

ACTO1 – Pregunta 3

(Grupo1_Tarde – 0,6 puntos)

Apellidos y Nombre: Díaz-Alejo León, Stéphane

(IMPORTANTE: todos los cálculos se mostrarán **redondeados a dos decimales**; se deben **justificar** las respuestas)

3) Sean una colección de documentos compuesta por los documentos doc1 y doc2 y la consulta:

doc1: El **desarrollo económico** depende de las **decisiones** del **gobierno**

doc2: El **sector económico** más importante impulsará las **decisiones económicas**

consulta: Las **decisiones económicas** debe tomarlas el **gobierno**

Los términos que considerar se han indicado en negrita, se ha llevado a cabo lematización tanto en consulta como en documentos, de forma que por ejemplo “económico” y “económicas” se consideran el mismo término. **Se pide:**

a) Completar la tabla para un esquema de pesado ltc.ltc (log-pesado, idf y coseno normalizado).

(0,4 puntos)

Term			Consulta				Doc1				Doc2			
	df_t	idf_t	$f_{t,d}$	$tf_{t,d}$	$w_{t,d}=tf \times idf$	L-Normaliz	$f_{t,d}$	$tf_{t,d}$	$w_{t,d}=tf \times idf$	L-Normaliz	$f_{t,d}$	$tf_{t,d}$	$w_{t,d}=tf \times idf$	L-Normaliz
desarrollo	1	0.3	0	0	0	0	1	1	0.3	0.71	0	0	0	0
económico	2	0	1	1	0	0	1	1	0	0	2	1.3	0	0
decisión	2	0	1	1	0	0	1	1	0	0	1	1	0	0
gobierno	1	0.3	1	1	0.3	1	1	1	0.3	0.71	0	0	0	0
sector	1	0.3	0	0	0	0	0	0	0	0	1	1	0.3	1

b) Indicar qué documento es más relevante para la consulta en base a la similitud coseno con esquema de pesado ltc.ltc. **(0,2 puntos)**

El documento 1, porque la similitud coseno es 0.71, sin embargo, con el documento 2 es 0.