

$$\sum_{\substack{k_i \in q \\ k_i \in dj}} \log \left( \frac{N+0,5}{n_i+0,5} \right)$$

Documentos que contém  $k_i$ .

A)

$$q = \{\text{logan}; \text{ORORO}; \text{xmen}\}$$

$$D_1 = \{\text{logan}; \text{ORORO}; \text{xmen}\} = 2,439$$

$$\begin{aligned} & \rightarrow \log \left( \frac{5+0,5}{3+0,5} \right) + \log \left( \frac{5+0,5}{3+0,5} \right) + \log \left( \frac{5+0,5}{2+0,5} \right) \\ & \rightarrow 0,651 + 0,651 + 1,137 \end{aligned}$$

$$D_2 = \{\text{logan}\} = 0,651$$

$$\log \left( \frac{5+0,5}{3+0,5} \right)$$

$$D_3 = \{\text{ORORO}\} = 0,651$$

$$\log \left( \frac{5+0,5}{3+0,5} \right)$$

$$D_4 = \{\text{logan}; \text{ORORO}; \text{xmen}\} = 2,439$$

$$D_5 = \underline{\underline{0}}$$

$$\log \left( \frac{5+0,5}{3+0,5} \right) + \log \left( \frac{5+0,5}{3+0,5} \right) + \log \left( \frac{5+0,5}{2+0,5} \right)$$

B) Em comparação ao exercício anterior,  
 Nota-se que os documentos 1 e 4  
 ainda são os mais bem ranqueados, o 2 e 3  
 em seguida, e o 5 com o valor zero.  
 Todavia, no probabilístico, o 1 e 4  
 ficam empatados entre si, e o 2 e 3  
 Nenhum fator que sirva, pois não há  
 desempenho, como o tamanho como parâmetros de  
 no modelo vetorial.

C)

$2,15$

$2,25$

$2,35$

$2,45$

$$D_1 = \log \left( \frac{(2+0,5)(5-3-2+2+0,5)}{(2-2+0,5)(3-2+0,5)} \right)^{\text{logar}} = \frac{\log(8,33)}{3,05}$$

$$\log \left( \frac{(2+0,5)(5-3-2+2+0,5)}{(2-2+0,5)(3-2+0,5)} \right)^{\text{otim}} = \frac{\log(8,33)}{3,05} +$$

$$\log \left( \frac{(2+0,5)(5-2-2+2+0,5)}{(2-2+0,5)(2-2+0,5)} \right)^{\text{xmen}} = \frac{\log(35)}{5,13}$$

11,23

$$D_2 = \log \left( \frac{(2+0,5)(5-3-2+2,0,5)}{(2-2+0,5)(3-2+0,5)} \right) = \underline{\underline{3,05}}$$

$$D_3 = \log \left( \frac{(2+0,5)(5-3-2+2+0,5)}{(2-2+0,5)(3-2+0,5)} \right) = \underline{\underline{3,05}}$$

$$D_4 = * \text{lg} 0,1 \text{ ou } D_4 = \underline{\underline{11,23}}$$

$$D_5 = 0,1$$

O  $\alpha_i$  de logan, ororo e xmen são a quantidade de documentos relevantes em que eles aparecem, ou seja, 2, 2 e 2 respectivamente.

Comparado com o resultado da questão A, nota-se que o ranking continua o mesmo, com  $D_1$  e  $D_4$  empatados em primeiro,  $D_2$  e  $D_3$  empatados em segundo e  $D_5$  zerado. A diferença é o aumento no numero de todos os resultados.