



RECUPERACIÓ DE LA INFORMACIÓ (REIN)

Formulari I

$$f(r) = \frac{k}{r^\alpha} = k * r^{-\alpha}$$

$$d = k_1 * N^\beta$$

$$tf-idf(d, t) = \frac{f_{d,t}}{\max_{j \in d} f_{d,j}} * \log \left(\frac{D}{df_t} + 1 \right)$$

Nota: el **+1** no el fem servir en els nostres càlculs manuals

$$sim(d1, d2) = \frac{d1 \cdot d2}{|d1||d2|} = \frac{d1}{|d1|} \cdot \frac{d2}{|d2|}$$

$$Recall = \frac{|\mathcal{R} \cap \mathcal{A}|}{|\mathcal{R}|}$$

$$Precision = \frac{|\mathcal{R} \cap \mathcal{A}|}{|\mathcal{A}|}$$

Regla de Rocchio:

$$q' = \alpha * q + \beta * \frac{1}{|R|} * \sum_{d \in R} d - \gamma * \frac{1}{|NR|} * \sum_{d \in NR} d \quad (\alpha > \beta > \gamma \geq 0)$$

PageRank:

$$p_i = \sum_{j \rightarrow i} \frac{p_j}{out(j)}$$

$$\sum_{i=1}^n p_i = 1$$

$$G = \lambda M + (1 - \lambda) \frac{1}{n} J \quad t.q. \quad 0 < \lambda < 1$$