$$tf_{ij} = f_{ij} / max_i\{f_{ij}\}$$
$$idf_i = \log_2 (N/df_i)$$
$$w_{ij} = tf_{ij}\, idf_i = tf_{ij} \log_2 (N/df_i)$$

$$\cos Sim(\vec{d}_j, \vec{q}) = \frac{\vec{d}_j \bullet \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum\limits_{i=1}^{t} (w_{ij} \cdot w_{iq})}{\sqrt{\sum\limits_{i=1}^{t} w_{ij}^2 \cdot \sum\limits_{i=1}^{t} w_{iq}^2}}$$

$$recall = \frac{Number\ of\ relevant\ documents\ retrieved}{Total\ number\ of\ relevant\ documents}$$

$$precision = \frac{Number\ of\ relevant\ documents\ retrieved}{Total\ number\ of\ documents\ retrieved}$$

$$P(r_j) = \max_{r_j \le r \le r_{j+1}} P(r)$$

$$F = \frac{2PR}{P+R} = \frac{2}{\frac{1}{R}+\frac{1}{P}}$$

$$E = \frac{(1+\beta^2)PR}{\beta^2 P + R} = \frac{(1+\beta^2)}{\frac{\beta^2}{R}+\frac{1}{P}}$$

$$Fallout = \frac{no.\,of\ nonrelevant\ items\ retrieved}{total\ no.\,of\ nonrelevant\ items\ in\ the\ collection}$$

Rochio : $\vec{q}_m = \alpha \vec{q} + \dfrac{\beta}{|D_r|} \sum\limits_{\forall d_j \in D_r} \vec{d}_j - \dfrac{\gamma}{|D_n|} \sum\limits_{\forall d_j \in D_n} \vec{d}_j$

Ide regular : $\vec{q}_m = \alpha \vec{q} + \beta \sum\limits_{\forall d_j \in D_r} \vec{d}_j - \gamma \sum\limits_{\forall d_j \in D_n} \vec{d}_j$

Ide "dec hi" : $\vec{q}_m = \alpha \vec{q} + \beta \sum\limits_{\forall d_j \in D_r} \vec{d}_j - \gamma \max_{non-relevant}(\vec{d}_j)$

Association matrix : $c_{ij} = \sum\limits_{d_k \in D} f_{ik} \times f_{jk}$ $\qquad s_{ij} = \dfrac{c_{ij}}{c_{ii} + c_{jj} - c_{ij}}$

Metric correlation matrix : $c_{ij} = \sum\limits_{k_u \in V_i} \sum\limits_{k_v \in V_j} \dfrac{1}{r(k_u, k_v)}$ $\qquad s_{ij} = \dfrac{c_{ij}}{|V_i| \times |V_j|}$

$f \propto \dfrac{1}{r}$ $\qquad f \cdot r = k$ (for constant k) $\qquad p_r = \dfrac{f}{N} = \dfrac{A}{r}$ for corpus indp. const. $A \approx 0.1$

$I_n = r_n - r_{n+1} = \dfrac{AN}{n} - \dfrac{AN}{n+1} = \dfrac{AN}{n(n+1)}$ $\qquad \dfrac{I_n}{D} = \dfrac{1}{n(n+1)}$

$V = Kn^\beta$ with constants $K,\ 0 < \beta < 1$