

$$tf_{ij} = f_{ij} / \max_i \{f_{ij}\}$$

$$idf_i = \log_2 (N / df_i)$$

$$w_{ij} = t_{ij} idf_i = t_{ij} \log_2 (N / df_i)$$

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

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

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

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

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

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

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

$$\text{Rocchio: } \vec{q}_m = \alpha \vec{q} + \frac{\beta}{|D_r|} \sum_{\forall \vec{d}_j \in D_r} \vec{d}_j - \frac{\gamma}{|D_n|} \sum_{\forall \vec{d}_j \in D_n} \vec{d}_j$$

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

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

$$\text{Association matrix: } c_{ij} = \sum_{d_k \in D} f_{ik} \times f_{jk} \quad s_{ij} = \frac{c_{ij}}{c_{ii} + c_{jj} - c_{ij}}$$

$$\text{Metric correlation matrix: } c_{ij} = \sum_{k_u \in V_i} \sum_{k_v \in V_j} \frac{1}{r(k_u, k_v)} \quad s_{ij} = \frac{c_{ij}}{|V_i| \times |V_j|}$$

$$f \propto \frac{1}{r} \quad f \cdot r = k \quad (\text{for constant } k) \quad p_r = \frac{f}{N} = \frac{A}{r} \quad \text{for corpus indep. const. } A \approx 0.1$$

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

$$V = Kn^\beta \quad \text{with constants } K, 0 < \beta < 1$$