Example: Rocchio Feedback

Consider the following 3 documents (one document per line) and the query Gift that Mary love:

D1. John thinks a book is a good gift

D2. John who reads a book loves Mary

D3. Mary who reads a book loves John

Say the documents 2 and 3 have been annotated as relevant, and the document 1 as not-relevant. Let us now focus on 4 terms in the dictionary: book, gift, love and Mary.

– Give the Rocchio-modified query (using $\alpha = 1$, $\beta = 0.75$ and $\gamma = 0.15$).

$$\vec{q_m} = \alpha\vec{q_0} + \beta\frac{1}{|D_r|}\sum_{\vec{d_j}\in D_r}\vec{d_j} - \gamma\frac{1}{|D_{nr}|}\sum_{\vec{d_j}\in D_{nr}}\vec{d_j}$$

where $\alpha = 1$, $\beta = 0.75$ and $\gamma = 0.15$

Consider the following three documents (one document per line) and the query "Pseudo relevance feedback":

D1. Pseudo relevance feedback increases positive results.

D2. Pseudo feedback improves Positive feedback.

D3. Rocchio is an information retrieval researcher who developed the Rocchio Algorithm.

Say documents 1 and 2 have been annotated as relevant and document 3 as irrelevant. Let us now focus on four terms in the dictionary: Feedback, Information, Positive, Relevance, and Rocchio.

– Give the Rocchio-modified query (using α = 1, β = 0.75 and γ = 0.5).

$$\vec{q_m} = \alpha \vec{q_0} + \beta \frac{1}{|D_r|} \sum_{\vec{d_j} \in D_r} \vec{d_j} - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d_j} \in D_{nr}} \vec{d_j}$$

where $\alpha = 1$, $\beta = 0.75$ and $\gamma = 0.5$

Consider the following vectors of three documents (one document per line) and the query:

D1. [2, 0, 3, 1, 0, 0]

D2. [0, 5, 0, 1, 2, 1]

D3. [3, 1, 2, 2, 0, 1]

D4. [0, 3, 0, 3, 3, 2]

Q. [1, 0, 1, 1, 0, 0]

Say documents 1 and 3 have been annotated as relevant and 2 and 4 as irrelevant.

– Give the Rocchio-modified query (using $\alpha = 1$, $\beta = 0.75$ and $\gamma = 0.15$).

$$\vec{q_m} = \alpha\vec{q_0} + \beta\frac{1}{|D_r|}\sum_{\vec{d_j}\in D_r}\vec{d_j} - \gamma\frac{1}{|D_{nr}|}\sum_{\vec{d_j}\in D_{nr}}\vec{d_j}$$