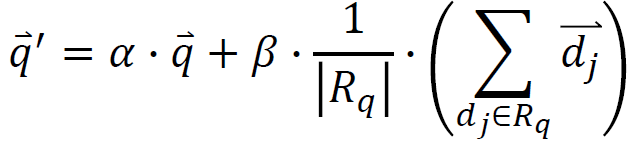
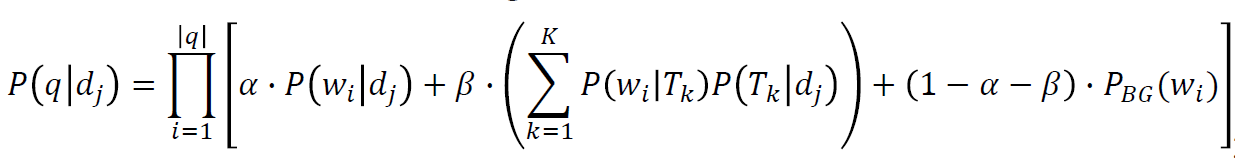
如果把PLSA直接丟上去就會過baseline，覺得不好好利用真是太浪費了，所以就依照Rocchio計算新query的simillarity方法改了PLSA的simillarity公式如下

Rocchio:



PLSA: 改過的PLSA:

P(q’|dj) = q\_ratio/|q|\*P(q|dj) + Rq\_ratio/|Rq|\*P(Rq|dj)

Rq = 前r篇relevance document全部接在一起(相當於一個很長的document)

|q|= query長度(字數)

|Rq|=前r篇relevance document字數加總

q\_ratio = 可調整參數

Rq\_ratio = 可調整參數