State of the Art VSM Ranking Functions

- Pivoted Length Normalization VSM [Singhal et al 96]

$$f(q,d)=\sum_{w\in q\cap d} c(w,q)\frac{\ln[1+\ln[1+c(w,d)]]}{1-b+b\frac{|d|}{avdl}}\log\frac{M+1}{df(w)}$$

이걸로 (핵심! sublinear transformation) → 2.2강

query term component

doc. length normalizer

IDF component

- BM25/Okapi [Robertson & Walker 94]

$$b\in[0,1]$$
$$k_1,k_3\in[0,+\infty)$$

TF 변환 → sublinear trans. w/ upper bound

$$f(q,d)=\sum_{w\in q\cap d} c(w,q)\frac{(k+1)c(w,d)}{c(w,d)+k(1-b+b\frac{|d|}{avdl})}\log\frac{M+1}{df(w)}$$