Kullback-Leibler (KL) Divergence Retrieval Model

Query Likelihood

$$f(q,d) = \sum_{\substack{w_i \in d \\ w_i \in q}} c(w,q)[\log \frac{p_{Seen}(w_i \mid d)}{\alpha_d \, p(w_i \mid C)}] + n \log \alpha_d$$

KL-divergence (cross entropy)

$$f(q,d) = \sum_{w \in d, p(w \mid \theta_Q) > 0} [p(w \mid \hat{\theta}_Q) \log \frac{p_{seen}(w \mid d)}{\alpha_d p(w \mid C)}] + \log \alpha_d$$

Query LM

$$p(w \mid \hat{\theta}_Q) = \frac{c(w,Q)}{\mid Q \mid}$$