

Que (6)  $\rightarrow$

for IR(1)  $\rightarrow$

$$\text{Average precision@k} \Rightarrow \frac{\sum_{k=1}^5 \text{prec.}(k) * \text{Rel.}(k)}{\text{Total No. relevant doc}} \quad \text{--- (1)}$$

$$\Rightarrow \frac{\text{prec.}(1) * \text{Rel.}(1) + \text{prec.}(2) * \text{Rel.}(2) + \text{prec.}(3) * \text{Rel.}(3) + \text{prec.}(4) * \text{Rel.}(4) + \text{prec.}(5) * \text{Rel.}(5)}{\text{Total No. of relevant docs.}}$$

$$\Rightarrow \frac{\frac{1}{1} * 1 + \frac{2}{2} * 1 + \frac{3}{3} * 1 + 0 + 0}{3}$$

$$\Rightarrow \frac{3}{3} = 1.$$

IR(2)  $\rightarrow$

A.P. @ k  $\Rightarrow$  { same formula from eq<sup>n</sup> (1) }

$$\Rightarrow \frac{\frac{1}{1} * 1 + \frac{1}{2} * 0 + \frac{2}{3} * 1 + 0 + \frac{3}{5} * 1}{3}$$

$$\Rightarrow \frac{1 + \frac{2}{3} + \frac{3}{5}}{3} \Rightarrow \frac{15 + 10 + 9}{15} \Rightarrow \frac{34}{45} \Rightarrow 0.76$$

Now, since  $\rightarrow$   $\boxed{AP(IR_1) > AP(IR_2)}$

So, IR<sub>1</sub> is better.

$\rightarrow$  option (a)

Que. 5 → all matches satisfy → option (d).

I → 3, II → 4, III → 2, IV → 1

Que. 4 → Model. of I.R., where, relevance feedback is not used. → (b) Ad-hoc retrieval.

Que. 1 → (Q) query ⇒ "smart+ tortoise"

(D<sub>1</sub>) Document 1 ⇒ "Tortoise with smart brains"

(D<sub>2</sub>) Document 2 ⇒ "Lazy tortoise"

$\text{Idf}(\text{smart}) = 1.65$

$\text{Idf}(\text{tortoise}) = 2$

~~$\text{Tf}(\text{smart})$~~  since →  $(Q \cap D_1) \Rightarrow \left\{ \begin{array}{l} \text{"smart"} \\ \text{"tortoise"} \end{array} \right\}$

$\frac{\text{Term}}{\text{freq.}} \rightarrow$

1

1

and,  $(Q \cap D_2) \Rightarrow \left\{ \begin{array}{l} \text{"tortoise"} \end{array} \right\}$

$\text{T.f.} \rightarrow$  1

for Doc. 1 →  $\text{Cal } (\text{Tf-IDf})_{Q \cap D_1} \Rightarrow$

$\Rightarrow \sum_{Q \cap D_1} (\text{Tf-IDf})$

$\Rightarrow (\text{Tf-IDf})_{\text{smart}} + (\text{Tf-IDf})_{\text{tortoise}}$

$\Rightarrow \left\{ 1 + \log_{10}(1) \right\} \times 1.65 + \left\{ 1 + \log_{10}(1) \right\} \times 2$

$\Rightarrow$  3.65



for Doc. (2)  $\rightarrow$  (Tf-IDf)<sub>"toroise"</sub>

$$\Rightarrow \{1 + \log_{10}(1)\} \approx 2$$

$\Rightarrow$  (2).

Ans.  $\rightarrow$  3.65 and 2

$\rightarrow$  option (b).

Que (3)  $\rightarrow$  while referring to postings in an index  $\rightarrow$   
option (b)  $\rightarrow$  Descending order of g(d). { bcoz we first start with the highest g(d) value, as it can result in greater net score }.

Que (4)  $\rightarrow$  which doesn't give satisfactory results.  $\rightarrow$   
 $\Rightarrow$  option (c) { bcoz for (a), some models only take the feedback & for (b), we make those weights as 0.5 }

Que (5)  $\rightarrow$  A thesaurus aids in  $\rightarrow$  option (d).

Que (2)  $\rightarrow$  for  $D_1 \rightarrow$  1 match, (Jill)

for  $D_2 \rightarrow$  2 matches (Jill, hurt)

while calc. dot product & ~~condi~~ considering each word as dimension,

option (b)  $\rightarrow$  (1) & (2).