An Example of IR Performance Measures

Dell Zhang
28/11/2006

Adapted from Dr CX Zhai's LM-IR tutorial in ACM SIGIR 2006.
http://sifaka.cs.uiuc.edu/lmir/sigir06-tutorial-lmir.pdf

Retrieval Result

1. d_1 √
2. d_2 √
3. d_3 ×
4. d_4 √
5. d_5 ×
6. d_6 ×
7. d_7 ×
8. d_8 ×
9. d_9 ×
10. d_{10} √

The total number of relevant docs = 8.

Evaluate it as a *set*.

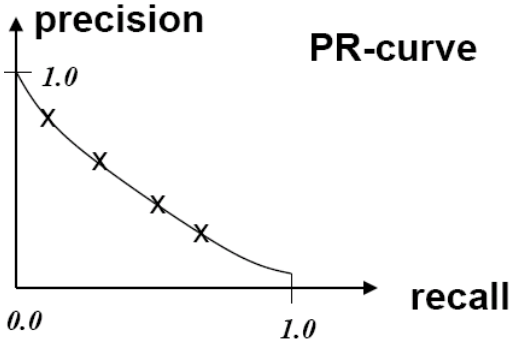
Contingency Table

	Relevant	Not Relevant
Retrieved	$tp = 4$	$fp = 6$
Not Retrieved	$fn = 4$	?

Precision $P = 4 / 10 = 0.4$
Recall $R = 4 / 8 = 0.5$

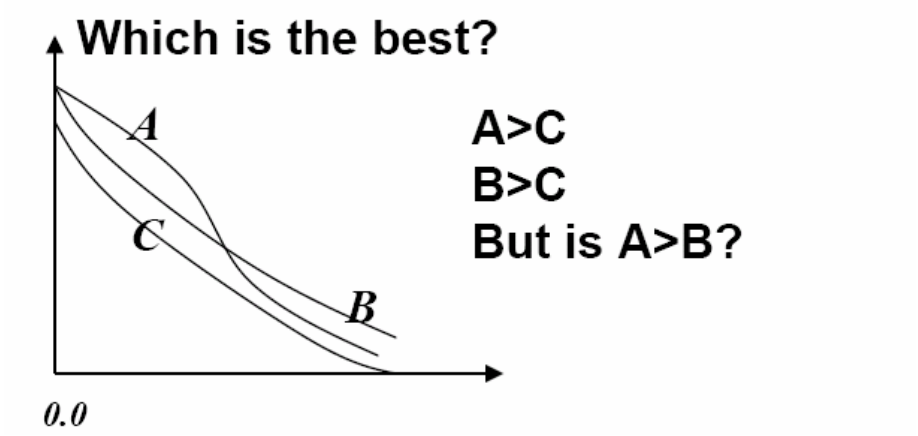
$F_1 = 2PR / (P+R) = 4/9 = 0.44$

Evaluate it as a *ranked list*.



PRBEP = 3/8 = 0.375

How do we compare different rankings?



Summarize a ranking with a single number

$$AvgPrec = \frac{1}{k}\sum_{i=1}^{k} p_i$$

k is the total # of rel docs

p_i = prec at the rank where the i-th rel doc is retrieved

p_i=0 if the i-th rel doc is not retrieved

AvgPrec = (1/1+2/2+3/4+4/10+0+0+0+0)/8=0.394