

Short Problem Set (Module 5)

1. [10%] Briefly describe the three key assumptions of the Cranfield paradigm for information retrieval evaluation.

2. [10%] What is pooling and why is it used in large-scale text retrieval evaluations?

3. [50%] Consider a query with 10 relevant documents whose docids are: D3, D27, D30, D39, D51, D54, D69, D72, D81, and D96. Assume that all other documents are not relevant. On this query two retrieval systems *FastSearch* and *Telescope* produce the following ranked lists. (Note: D17 is the 1st ranked doc by *FastSearch*; D4 is its 2nd ranked doc, etc...)

FastSearch: D17, D4, D69, D54, D37, D41, D89, D85, D3, D5, D91, D39

Telescope: D3, D1, D94, D27, D50, D54, D16, D7, D72, D39, D95, D62

(a) How many relevant documents are found by each system?

(b) For both systems what is $P@10$ (precision at 10 documents) for this query?

(c) For *FastSearch* what is the uninterpolated precision at 30% Recall?

(d) Assuming that *FastSearch* returns no other documents other than this top-12 ranked list, what is *FastSearch*'s Recall for this query?

(e) For both systems what is average precision on this query?

4. [15%] Given two retrieval systems (called A and B), is it possible for System A to be better than System B in average precision, but for System B to have higher $P@10$ than System A? Briefly justify your response.

5. [15%] Consider the contingency tables below for the word pairs (bicycle, helmet) and (bicycle, repairs). Suppose we are looking to expand a query containing the word *bicycle* by adding some potentially useful search terms. Using pointwise mutual information (PMI) to score candidate terms, calculate scores for both *helmet* and *repairs*, and indicate which of the two would be the better expansion term. $N = 15,000$ documents. Use base 2 logs.

$$PMI(x, y) = \log \left(\frac{P(x, y)}{P(x) P(y)} \right) = \log \left(\frac{N \times a}{(a+b)(a+c)} \right)$$

A: docs with both terms together	B: docs with first term, but not second
C: docs with second term, but not first	D: docs that contain neither term

	has helmet	missing helmet
has bicycle	22	54
missing bicycle	87	14837

	has repairs	missing repairs
has bicycle	31	45
missing bicycle	164	14760