

National University of Singapore  
School of Computing  
CS3245: Information Retrieval  
Tutorial 5

## IR Evaluation, Relevance Feedback and XML Retrieval

**Readings:** IIR Chapters 8 (excluding 8.5), 9 & 10

### 1. Non-ranked Evaluation.

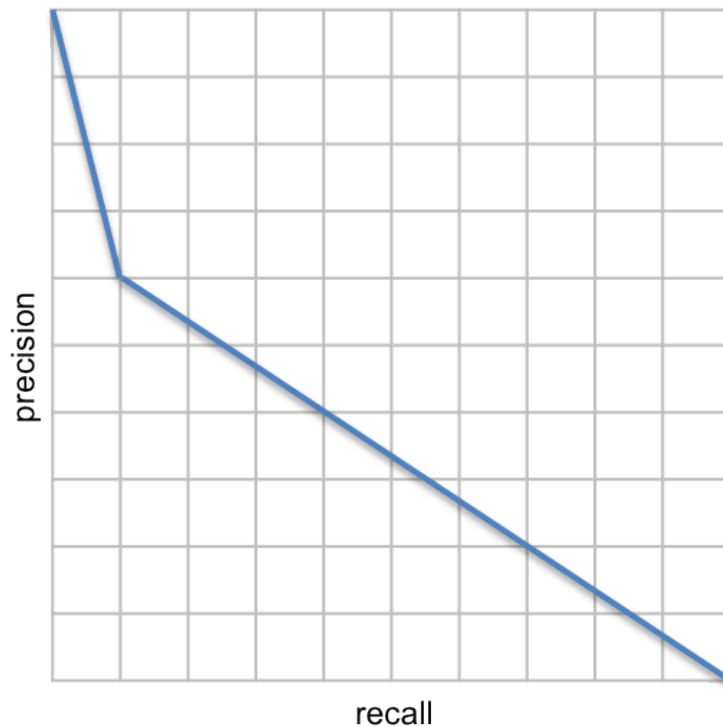
A related problem to information retrieval is *information extraction (IE)*, where automated programs pore over natural language text to try to identify certain key information, and extract it for use downstream.

Let us use an example task of an IE system that identifies people and organization names from news text, such as from the Reuters newswire. In evaluating such a system we have three labels that can be assigned to a word: {(part of) person name, (part of) organization name, neither}.

- (a) Two competing companies, *A* and *B*, are bidding for a contract to create a production system for the above task to showcase for further funding, and have produced two prototypes. *System A*'s claims 95% accuracy in its system's judgment, whereas *B* claims only 90% accuracy. Which is the better system?
- (b) Your friend, Shuhrakh Kan, states that he can build a system with 95% accuracy on the task, within a few minutes. Is he kidding? If not, explain how he might accomplish his task.
- (c) Your other friend, Sylvia Taronel, suggests that the evaluation metric of accuracy isn't appropriate. Is she right? She suggests using  $F_1$  instead. Will that solve the problem?

### 2. Ranked Evaluation.

- (a) The below line plot shows the precision-recall curve for a ranked retrieval system. At what point is  $F_1$  maximized? What about  $F_{0.5}$ ?  $F_2$ ?



We have seen that interpolated precision curves generally look like a concave curve connecting the upper left corner (high precision, low recall) with the bottom right (high recall, low precision), such as the above plot in (a). Let's look into this in a bit more detail.

- (b) Do we always have a point at the upper left hand corner (that is, at the  $y = 1, x = 0$ ) on every precision and recall curve?
- (c) Do we always have a point at the lower right (that is, at the  $y = 0, x = 1$  point) on every precision and recall curve?
- (d) Is precision always a non-increasing function with respect to larger recall values on an interpolated precision-recall curve? What about on a non-interpolated curve?

**(The following questions will be covered in Tutorial 6 instead.)**

### 3. Relevance Feedback in Vector Space.

The below diagram show documents as 'X's in a two-dimensional vector space. A query given by the user is shown as a filled circle.

- (a) Show the ranking for the documents under the VSM for the original query, by numbering any relevance documents with their rank ("1" for 1st ranked, "2" for 2nd ranked, etc.) If two or more documents are tied in rank, use the same number; and if a document is not relevant, mark it with a "0".



how we could modify the context resemblance algorithm to cater for these types of queries.