Trier University   Trier, 02.01.2023
Department IV – Computer Sciences
Prof. Dr.-Ing. Ralf Schenkel (H 539, schenkel@uni-trier.de)

Assignment 8 in the course Digital Libraries and Foundations of Information Retrieval

Winter Semester 2022

Deadline 12:15h on Monday, 09.01.2023

Task 1:   Relevance Assessment   2+4+2+7 Points

In this task we want to evaluate the search engine of Wikipedia (`https://en.wikipedia.org/wiki/Special:Search`). We will use the following topic description (see slide 7-5):

```
<top>
<num> 1
<title> trier university
<desc> Structure and history of Trier University
<narr> Relevant documents describe the structure, organization or
history of the university. Articles on individual programs, departments
or institutions may be relevant if they are of significance for
the university. Articles on alumni are only relevant if they are
related to the history of the university. Articles about other
colleges in Trier or universities in other locations can
be relevant if there is a connection to Trier University.
</top>
```

(a) Determine the top 20 results for this topic with the search engine of Wikipedia and show them in your solution in a reasonable way. Assign a unique number to each result.

(b) Assess each result with respect to its relevance for the topic. Use the relevance degree "highly relevant", if the result contains (almost) only relevant content; "relevant", if a part of it is relevant, but it also contains non-relevant content; "not relevant", if the result is not relevant at all for the topic.

(c) Determine the precision of the result list.

(d) Following the above scheme, develop a new topic by giving a `title` (i.e., a search engine query), a `description` and a `narrative` (the extensive description of the content of relevant and non-relevant documents). Determine the top 20 results for this topic with the Wikipedia search engine, assess the relevance of the results for your topic, and determine the precision of the result list.

Task 2: Relevance Assessment 1+2+3+3+3+3 Points

(a) Give at least one reason why it is practically impossible to assess all documents of a collection with respect to their relevance for a topic.

(b) Explain how one solves this problem with pooling, i.e., how the documents to assess are determined. Why can one ignore the influence of unassessed documents?

(c) Assume now that a new search engine (or a new search algorithm, or a new parameter setting of an existing algorithm) is evaluated with the existing collection. What problem could appear here? How could one solve this problem with some additional effort?

(d) Assume that you should design a benchmark for text search, for which you can perform assessments of up to 10,000 results. You want to compare ten systems. Is it better to (1) consider 10 topics and assess 1,000 results for each, or (2) consider 100 topics and assess 100 results for each? In both scenarios, the documents to assess are determined using pooling. Could the retrieval task under consideration have an influence? Could the selection of systems have an impact? Give a reason for your decision.

(e) In a search engine, you observe that a user clicks the second result of a result list, but not the first. Explain why this can be a signal that the second result is "more relevant" for the user than the first result. Explain also why the first result can still be relevant, even though the user did not click it.

(f) In the same situation as in part (e), you now observe that the user does not click the third result. Explain why this can be a sighnal that the third result is not relevant. Then explain why the third result can still be relevant even though the user did not click it.

Task 3: MAP 6+6+3 Points

In this task we consider a collection with documents A through T.

(a) We consider a query q_1 for which documents A, B, C, and D are relevant. Assume that a search engine returns the (sorted) result list

A — V — R — C — S — B — J — L — E — D

(10 documents). Compute the (non-interpolated) precision at all rank positions and AP.

(b) We now consider a query q_2 for which the documents A, R, S, and T are relevant. Assume that a search engine returns the (sorted) result list

A — V — R — C — S — B — J — L — E — D

Compute the (non-interpolated) precision at all rank positions and AP. If a document does not occur in the result, use a precision of 0 for that document.

(c) Compute MAP (slide 7-26) and GMAP (slide 7-28) for the two result lists from tasks (a) and (b).

These tasks will be discussed in the meeting on January 16.

General remarks:

- The tutorial group takes place on Mondays in the regular meeting in F55 at 12:15.

- The first meeting of the tutorial group will be on November 14, 2022.

- To be admitted to the final oral exam, you need to acquire at least 50% of the points in the assignments. In addition, you need to present at least one solution of a task in a convincing way during the tutorial.

- It is preferred to submit in groups of size up to two (but not larger); in that case, only one submission is sufficient for the whole group. Write the names of all group members on your solutions.

- Solutions must be handed in before the deadline

  - in Moodle (`https://moodle.uni-trier.de/`, course `DL-IR-2022`) as as a PDF or, if submitting multiple files, as an archive (.zip or comparable).

  Submissions that arrive after the deadline will not be considered. The name of at least one group member should occur in the file name of your submission. If you want to modify a previously uploaded solution, just re-upload your solution.

- Graded versions of your submissions will be returned in Moodle until the following tutorial.

- Announcements regarding the lecture **and** the tutorial group will be done in the StudIP course for the lecture.