Chair for Algorithms
and Data Structures
Prof. Dr. Hannah Bast
Claudius Korzen

Information Retrieval
WS 2017/2018

`http://ad-wiki.informatik.uni-freiburg.de/teaching`

Exercise Sheet 2

Submit until Tuesday, November 7 at **12:00pm**

Exercise 1 (8 points)

Copy your code from *sheet-01* to a new subfolder *sheet-02*. Extend your code to incorporate BM25 scores, as explained in the lecture. This entails the following:

1. In your method *read_from_file*, add BM25 scores to the inverted lists. Pay attention to the implementation advice given in the lecture, and avoid unnecessary complexity. (4 points)

2. Change your method *intersect* into a method *merge* for merging two lists. Note that this can be done with a relatively minor change. (2 points)

3. In your method *process_query*, sort the results by the aggregated BM25 scores and output the top-3 results. You don't have to implement the sorting algorithm yourself; you can use one of the built-in sorting functions. (2 points)

There is again a TIP file on the Wiki with a suggestion for the structure of your code and test cases for some of the functions. As described in the guidelines on the back of Exercise Sheet 1, you have to implement these test cases (otherwise your submission will not be graded; the content of the test case is important, not the exact syntax). This goes without saying from now on.

Exercise 2 (4 points)

Find a good setting for the BM25 parameters by inspecting the results for a variety of queries of your choice. Optionally (= you don't have to do this to get full points), feel free to improve your ranking in any way you see fit; we discussed various possibilities in the lecture (slide 19).

This development phase should be completed **before** you proceed with Exercise 3. Briefly(!) describe your insights from this phase in your *experiences.txt*.

[please turn over]

Exercise 3 (8 points)

Evaluate your system on the benchmark provided on the Wiki. The file *movies-benchmark.txt* provides 10 queries, and for each query the ids (line numbers in *movies.txt*) of the relevant documents.

1. Write a function *read_benchmark* that reads each query with the associated set of relevant document ids from the file (one query per line). (2 points)

2. Write two functions *precision_at_k* and *average_precision* that compute the measures *P@k* and *AP* for a given list of result ids as it was returned by your inverted index for a single query, and a given set of ids of all documents relevant for the query. (2 points)

3. Write a function *evaluate* that evaluates a given inverted index against a given benchmark and computes the measures *MP@3*, *MP@R*, and *MAP* by aggregating *P@3*, *P@R* and *AP* for each query in the benchmark. (2 points)

4. Write a *main* function that takes the paths to the dataset (*movies.txt*) and the benchmark file (*movies-benchmark.txt*) as command line arguments, constructs an inverted index from the dataset and does the evaluation of the inverted index against the benchmark using the methods above. Optionally, you can allow more arguments for the value of b and k or other parameters that influence the ranking. (2 points)

Report your results in the table on the Wiki, following the examples already given there. In the last column, provide your BM25 parameter settings + a very brief (not complete) description of any additional feature you might have added.

Once you start testing your system on this benchmark you should not go back anymore to Exercise 2 and change your ranking method or tune your parameters. That would be overfitting. In a real competition, your system would be evaluated by the organizers after the submission deadline, on an undisclosed benchmark.

Commit your subfolder *sheet-02* to our SVN. The dataset and the benchmark file should not be commited. Also take care not to commit any by-products (like Python's cache folder or Java's jar file or C++'s object files).

As usual, in your *experiences.txt*, provide a brief account of your experience with this sheet and the corresponding lecture. As a minimum, say how much time you invested and if you had major problems, and if yes, where.