Web Information Retrieval

Assignment 2

Lukas Schmelzeisen lukas@uni-koblenz.de
Dr. Chandan Kumar kumar@uni-koblenz.de

Institute of Web Science and Technologies
Department of Computer Science
University of Koblenz-Landau

Submission until: May 16, 2017, 10:00 a.m.
Tutorial on: May 16 and May 19, 2017

This assignment covers *Evaluation* in Information Retrieval. As supplementary material to the lecture, I can recommend reading *Chapter 8* of the book "Introduction to Information Retrieval" by Manning, C. D., Raghavan, P., and Schütze, H. from 2008. The book is available for free under http://informationretrieval.org

1 Calculation of Evaluation Measures (9 Points)

For the evaluation of an information retrieval system, three queries have been performed against the system. The following tables indicate the results of the queries along with relevance judgments provided by human experts:

Query q_1		
Rank	Doc-ID	Rel.?
1	d_{11}	
2	d_{12}	✓
3	d_{13}	✓
4	d_{14}	
5	d_{15}	✓
6	d_{16}	
7	d_{17}	

Total number of documents relevant to q_1: 4

Query q_2		
Rank	Doc-ID	Rel.?
1	d_{21}	✓
2	d_{22}	
3	d_{23}	✓
4	d_{24}	
5	d_{25}	
6	d_{26}	✓
7	d_{27}	
8	d_{28}	✓

Total number of documents relevant to q_2: 8

Query q_3		
Rank	Doc-ID	Rel.?
1	d_{31}	
2	d_{32}	
3	d_{33}	✓
4	d_{34}	

Total number of documents relevant to q_3: 1

1. Calculate *Precision*, *Recall* and F_1*-Measure* for all three result sets, treating them as unranked result sets.
2. Calculate *Precision at* k (P@k) with $k = 1$ and $k = 5$, and *R-Precision* for all three result sets, treating them as ranked result sets.

Note: for calculating *Precision at* k, if a result set contains less than k documents, treat the result set as being filled up with non-relevant documents until is would contain k documents.

2 F-Measure (5 Points)

In Information Retrieval (and many other fields) one typically uses the F_1-score as an evaluation measures that combines both precision and recall.

1. Give the definition of the more-general F_β-score for arbitrary values of β.
2. Explain why the choice of $\beta = 1$ is most common.
3. Explain how the behavior of the measure would change for $\beta = \frac{1}{3}$.

   Think of and describe a scenario (or domain) in Information Retrieval where such a choice would make sense.

3 Programming (26 Points)

3.1 Evaluation Measures (15 Points)

Your task is to implement the computation of the following evaluation measures:

- Precision
- Recall
- F_1-Score
- Precision@k
- R-Precision
- Mean Average Precision

For your solution, implement the respective functions inside the provided `evaluation.py`. You can test your solution using the provided tests in `test_evaluation.py`.

The classes `MyRankedRetrievalSystem` and `MyExpertJury` implement a dummy ranked retrieval system, that ranks documents on relevance to a query, and an artificial human expert jury, that defines a gold standard of relevant documents for a query, respectively.

3.2 Precision/Recall Graph (11 Points)

Your task is to write a programm that plots the *precision/recall graph* and the *11-point interpolated precision graph* together in one figure.

Specifically, your solution should

1. Let the user enter an arbitrary query keyword.
2. Calculate the *precision@k* and *recall@k* measures for all possible values of k and plot them in a *precision/recall graph.*
3. Calculate the *interpolated precision* for the 11 values of recall and plot them to a *11-point interpolated precision graph.*

For your solution, create a new file `precision_recall_graph.py` that creates a figure using matplotlib and saves it to `precision_recall_graph.pny`. Also commit that figure file to the SVN for the query `"once in a blue moon"`. Make sure that your graph's axes are labeled and that you add a legend to your figure that denotes which curve corresponds to which task.

Important Notes

Ask questions and discussion with regard to the lecture, tutorial, or assignments in

- The WebScience newsgroup:
  https://webnews.uni-koblenz.de/newsgroups.php?group=infko.webscience
- Our Facebook group:
  https://facebook.com/groups/InformationRetrievalUniKoblenz

Submission

- Solutions have to be checked into the SVN repository. Use the directory name `solutions/assignment2/` in your group's repository *mandatory*.
- The SVN repository is available via
  https://svn.uni-koblenz.de/westteaching/ir-ss17/<groupname>
- Solution format: all solutions for theoretical taks as *one* PDF document. Programming code has to be submitted as Python code.
- The name of the group and the names of all participating students must be listed on each submission:
  - at the top of each PDF file
  - at the top of each Python file (in comments)

  Only named students will receive credit.
- With the submission of your solution you confirm that you created the solution independently as a group, especially without using other intellectual contributions. That is, you submission should not be plagiarism!

Programming Assignments

The programming assignments require you to have a Python 3.6+ interpreter (older versions may still work coincidentally, but are not officially supported) and the SciPy stack installed. For an installation guide, see http://west.uni-koblenz.de/en/studying/installing-python

The following rules apply for the *implementation* of your solution:

- You can create as many additional code files as you need. You can create as many additional classes or methods as you need (even in the provided code files). However, do not forget to submit *all* `.py` code files to the SVN.
- Check that your code compiles and runs without errors.

- Use `UTF-8` as the file encoding. *Other encodings will not be taken into account!*
- Test your implementation with the provided test cases.

  Passing of all tests is a *necessity* to receive full score, no *sufficiency*. In general, programming against the provided test cases should assist you in finding correct solutions.
- Make sure your code is formatted to be easy to read.
  - Make sure you code has consistent indentation.
  - Make sure you comment and document your code adequately in English.
  - Choose consistent and intuitive names for your identifiers.
- Do *not* use any accents, spaces or special characters in your filenames.