Call for Master Thesis

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Making good use of bad data – Experimental negative relevance feedback simulation

Are you interested in working with cutting edge technology in your thesis? Work with us on an innovative approach for relevance feedback.

Interactive Relevance Feedback (RF) has been a longstanding promising retrieval technique that unfortunately failed to deliver good, consistent results for a long time. Now, recent technological advances in the area of Deep Learning have opened up new possibilities. In some areas, e.g. image classification, the improvements to-date have been unprecedented and dramatic. This work focusses on how the new techniques can be applied to negative relevance feedback in patent search. Negative RF is particularly interesting because it can work in the absence of positive results, which are sometimes difficult to find. However, as is the case in most of professional search today, there is an abundance of not or not quite relevant results.

Building on publicly available, well-curated test collections, the thesis encompasses:

- Setup of a baseline system with standard relevance ranking algorithms, e.g. BM25.
- Training of word embedding models, possibly knowledge graph embeddings, depending on the student's expertise.
- Implementation of a re-ranking process which augments or replaces exact matching with semantic matching, e.g. with BERT[1].
- Rigorous evaluation of the above.

The test collections contain relevance judgments which will be used to give the system feedback. Based on that feedback, a re-ranking process is carried out, utilizing deep learning methods and the trained models.

The aim of the thesis is to arrive at an understanding of in how far the examined deep learning methods may contribute to the retrieval success via negative relevance feedback.

This thesis will be supervised by **Prof. Dr. Harald Sack, Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe**.

[1] https://bit.ly/2OqmY5D

Which prerequisites should you have?

- Good programming skills in Python or Java
- Interest in Experiments and Evaluation
- Interest in Natural Language Processing, Semantic Web technologies, or Deep Learning technologies

Contact person:

René Hackl-Sommer
rene.hackl-sommer@fiz-karlsruhe.de

aifb

Institute of Applied Informatics and Formal Description Methods
http://www.aifb.kit.edu/

Karlsruhe Institute of Technology