Novel methods for query-performance prediction

Anna Shtok

Supervisor: Prof. Oren Kurland

The query-performance prediction (QPP) task is estimating retrieval effectiveness when no relevance judgments are provided. We present several novel state-of-the-art approaches to query-performance prediction. We first devise a query-performance predictor that is based on analyzing the retrieval-scores distribution of documents in the result list. The predictor is based on estimating the potential amount of query drift in the list, that is, the presence (and dominance) of aspects or topics not related to the query in the top-retrieved documents. We then present a novel general framework to query-performance prediction. The framework is based on using statistical decision theory for estimating the utility that a document ranking provides with respect to an information need expressed by a query. To generalize further we present a unified framework for post-retrieval prediction. The framework is based on using a pseudo effective and/or ineffective ranking as reference comparisons to the ranking at hand, the quality of which we want to predict. We will demonstrate the concepts and the effectiveness of the above approaches using several benchmarks provided by TREC.