SIGIR'12 tutorial: Predicting Query Performance for IR

SHORT SUMMARY

The goal of this tutorial is to expose participants to current research on query performance prediction (also known as query difficulty estimation). Participants will become familiar with state-of-the-art performance prediction methods, with common evaluation methodologies for prediction quality, and with potential applications that can utilize such performance predictors. In addition, some open issues and challenges in the field will be discussed.

BIO:

David Carmel: is a Research Staff Member at the Information Retrieval group at IBM Haifa Research Lab. David's research is focused on search in the enterprise, query performance prediction, social search, and text mining. David has published more than 80 papers in IR and Web journals and conferences, organized a number of workshops, and taught several tutorials at SIGIR, and WWW. David is co-author of the book "Estimating the Query Difficulty for Information Retrieval".

Oren Kurland: is a Senior Lecturer in the Faculty of Industrial Engineering and Management at the Technion, Israel Institute of Technology. The information retrieval research group that Oren leads at the Technion focuses on developing formal models for information retrieval. Oren published more than 30 peer-reviewed papers in IR conferences and journals. He served as a senior program committee member (area chair) for the SIGIR and CIKM conferences. Oren also serves on the editorial board of the Journal of Information Retrieval.

ABSTRACT:

Many information retrieval (IR) systems suffer from a radical variance in performance when responding to users’ queries. Even for systems that succeed very well on average, the quality of results returned for some of the queries is poor. Thus, it is desirable that IR systems will be able to identify `difficult" queries in order to handle them properly. Understanding why some queries are inherently more difficult than others is essential for IR, and a good answer to this important question will help search engines to reduce the variance in performance, hence serving better their customer needs.
The high variability in query performance has driven a new research direction in the IR field on estimating the expected quality of the search results, i.e., the query difficulty, when no relevance feedback is given. Such estimation is beneficial for many reasons:

1) As feedback to the users: The IR system can provide the users with an estimate of the expected quality of the results retrieved for their queries. The users can then rephrase queries that were found to be "difficult", or alternatively, resubmit the "difficult" query to alternative search resources.

2) As feedback to the search engine: The search engine can invoke alternative retrieval strategies for different queries according to their estimated difficulty. For example, heavy query-analysis procedures that are not feasible for all queries due to time response restrictions, may be invoked selectively only for difficult queries.

3) As feedback to the system administrator: The administrator can identify queries related to a specific subject that are "difficult" for the search engine, and to expand the collection of documents to better answer poorly covered subjects. Identifying missing content queries is especially important for commercial search engines which should better identify, as soon as possible, popular emerging user needs that cannot be answered appropriately due to missing relevant content.

4) For IR applications. For example, difficulty estimation can be used by a distributed search application as a method for merging the results retrieved from different datasets by weighing the results according to their estimated quality.

Estimating the query difficulty is a significant challenge due to the numerous factors that impact retrieval performance. Many prediction methods have been proposed throughout the last years. However, as many researchers observed, the prediction quality of state-of-the-art predictors is still too low to be widely used by IR applications. The low prediction quality is due to query ambiguity, missing content, vocabulary mismatch, and many other factors. This complexity burdens the estimation task and calls for new prediction methods that will be able to cope with this complex prediction challenge.

In this tutorial we will first discuss the reasons that cause search engines to fail for some of the queries. Then, we will overview several state-of-the-art approaches for estimating query difficulty.

We will also describe common methodologies for evaluating the prediction quality of query-performance prediction methods, and present some experimental results for the prediction quality of some predictors, as measured over several TREC's benchmarks. We will cover a few potential applications that can utilize query difficulty estimators. Finally, we will summarize with a discussion on open issues and challenges in the field.