



School of Computing, Engineering, and the Built Environment Edinburgh Napier University

PHD STUDENT PROJECT

Funding and application details

Funding status: Fully funded project (worldwide)

Application instructions:

Detailed instructions are available at <https://blogs.napier.ac.uk/scebe-research/available-phd-student-projects/>

Prospective candidates are encouraged to contact the Director of Studies (see details below) to discuss the project and their suitability for it.

Project details

Supervisory Team:

- DIRECTOR OF STUDY: Md Zia Ullah (Email: M.Ullah@napier.ac.uk)
- 2ND SUPERVISOR:

Subject Group: Computer science

Research Areas: Artificial Intelligence, Data Science, Machine Learning

Project Title: Query Performance Prediction for Neural Information Retrieval: Application to Conversational Search

Project description:

In information retrieval (IR), query performance prediction (QPP) aims to predict the search effectiveness for a given query without resorting to relevance judgments. QPP may be advantageous in many ways, such as signaling an IR system whether a search query would be effective or underperforming. Based on that information, the system can either apply a query reformulation [6] or an adaptive retrieval configuration [1,4] or engage in an interactive session with the user (i.e., conversational search [7]) to understand the search intent and provide a better search experience.

Predicting query performance is a challenging problem due to many characteristics of queries, collections, and search systems. Existing QPPs are extracted from traditional retrieval models (e.g., BM25 or Divergence from randomness) using the pre-retrieval features based on the collection statistics or the post-retrieval features based on the top-retrieved documents [2, 3]. With the advent of language models (e.g., BERT [6]), neural IR models have been proposed and shown to have better retrieval effectiveness [8]. However, QPP on the neural retrieval model has not yet been explored [9].

This Ph.D. project aims to develop neural query performance predictors from neural IR models and combine them with existing QPPs based on traditional IR models. Experiments could be conducted on standard TREC collections (e.g., MS MARCO and TREC Deep learning tracks) to demonstrate the effectiveness of the QPPs and compare them with the state-of-the-art approaches. Another goal of this project would be to design appropriate metrics to evaluate the QPPs and apply them to Conversational search [7].

Prospective applicants are encouraged to contact the Supervisor before submitting their applications. Please consult with the Supervisor if you want to work on the broad areas of IR and NLP.

References:

- [1] Mothe J and Ullah MZ, Defining an Optimal Configuration Set for Selective Search Strategy – A Risk-Sensitive Approach, Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM 2021), 2021.
- [2] Déjean S, Ionescu RT, Mothe J, and Ullah MZ, Forward and backward feature selection for query performance prediction, The 35th ACM/SIGAPP Symposium on Applied Computing (SAC), 2020.
- [3] Chifu AG, Laporte L, Mothe J, and Ullah MZ, Query Performance Prediction Focused on Summarized Letor Features, The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2018.
- [4] Deveau R, Mothe J, Ullah MZ, Nie JY, Learning to Adaptively Rank Document Retrieval System Configurations, ACM Transactions of Information Systems (ACM TOIS), 41 pages, pp.3:1-3:41, Volume 37, Issue 1, 2019.
- [5] Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805, 2018.
- [6] Azad, Hiteshwar Kumar and Deepak, Akshay, Query expansion techniques for information retrieval: a survey, Information Processing and Management, 2019
- [7] Faggioli et. Al., A Geometric Framework for Query Performance Prediction in Conversational Search, ACM SIGIR, 2023
- [8] Zhuang et al., Rankt5: Fine-tuning T5 for text ranking with ranking losses, ACM SIGIR, 2023
- [9] Faggioli et al., Query Performance Prediction for Neural IR: Are We There Yet? Advances in Information Retrieval, ECIR 2023

Candidate characteristics

Education:

A first-class honours degree, or a distinction at master level, or equivalent achievements in Computer Science or Data Science or Software Engineering

Subject knowledge:

- Information retrieval (IR),
- Natural language processing (NLP), and
- Machine learning/Deep learning.

Essential attributes:

- Experimentation experience of Information retrieval, Natural language processing, Language models, and Machine learning/Deep learning projects.
- Knowledge of IR and Deep Learning.
- Competent in R, Python, PyTorch, Java, and Shell scripting
- Good written and oral communication skills
- Strong motivation, with evidence of independent research skills relevant to the project
- Good time management

Desirable attributes:

- Experience of IR Tools, such as Terrier IR, Lemur Indri, Lucene.
- Experience of PyTorch or Keras