Preface

QPP++ 2023: Query Performance Prediction and Its Evaluation in New Tasks is the first edition of a workshop that aims to foster a discussion within the community on how Query Performance Prediction (QPP) can be applied to new techniques in Information Retrieval (IR) and how such techniques can be exploited to define new QPP models. This first edition was hosted by the European Conference on Information Retrieval (ECIR) 2023 in Dublin (Ireland).

QPP++ 2023 received nine scientific submissions, of which seven papers (four long and three short) were accepted. Two to three program committee members reviewed each submission, and the program chairs oversaw the reviewing. The accepted papers included authors from 8 countries and 14 institutions, as some publications resulted from international collaborations. Researchers addressed the following challenges: QPP for conversational search, known-item search and passage retrieval, QPP in the learning-to-rank and neural information retrieval domains, issues with using correlation metrics to evaluate QPP, QPP evaluation using pointwise approaches, continuous evaluation, and using information theory for QPP.

Motivation

The advent of large language models and the rise of new tasks, such as conversational search, semantic search and question answering, enabled by the availability of new powerful technological tools, have led to a previously unseen rapid growth in the variety and quality of Information Retrieval (IR) systems. Several ancillary research fields have also flourished due to the scientific uptake of new Natural Language Processing (NLP) methodologies, facilitating advancement in new IR tasks. The Query Performance Prediction and Its Evaluation in New Tasks (QPP++ 2023) workshop [3] aimed to further fuel such growth in the renowned and important area of Query Performance Prediction (QPP).

The QPP task is defined as estimating search effectiveness in the absence of human relevance judgments [1]. Since its introduction at the beginning of the 21st century, QPP has established itself as an essential tool in numerous tasks, including model selection [1, 14], query suggestion [1, 14], and rank fusion [11]. The QPP++ 2023 workshop was a collaborative effort of researchers to master the new tools made available by the NLP community and learn how to effectively use them for the QPP task. The workshop focused on applying QPP in traditional scenarios, such as ad-hoc retrieval, and in new domains, including conversational and semantic search, passage retrieval, and question answering. QPP++ 2023 also allowed the community to reexamine past weaknesses and challenges linked to the QPP task, such as its evaluation, while establishing a roadmap to organize and guide the community's future efforts to advance the QPP research field.

QPP and Novel Search Paradigms Given the recent developments in IR, the prediction quality of existing QPP approaches may be significantly affected in new domains and scenarios for the following three reasons. First, some of the traditional predictors exploit statistics derived from the collection [6], while new IR models often use indexes of embeddings or apply machine learning to re-rank documents [8]. Second, the vast majority of the recently developed retrieval models in IR utilize semantic information that, with a few notable exceptions [9, 13], is rarely exploited by QPP models. This, in turn, impairs the performance of traditional QPP models applied on IR systems based on new paradigms [4]. Finally, QPP can be used for new processes such as selective query processing [2].

The QPP++ 2023 workshop aimed to provide a platform for the community to jointly discuss ways to address these challenges and create a better alignment between the latest technologies, retrieval models, and QPP approaches. Along with the challenges mentioned above, the recent advances in NLP present great opportunities for enhancing the state of the art in QPP. The workshop also sought to encourage collaboration between researchers to exploit these opportunities.

QPP and its Evaluation on New Tasks The quality of QPP methods is typically evaluated by computing the correlation between the scores assigned to queries by a QPP method and the true performance values, e.g., Average Precision (AP), attained for these queries using relevance judgements. Previous research demonstrated the unreliability of this approach when multiple experimental factors (i.e., IR models, corpora, and predictors) are considered [7, 12, 5]. In addition, researchers demonstrated that high correlation does not necessarily translate to improved retrieval effectiveness [10, 7]. These issues are further exacerbated in new domains, such as question answering or conversational search, where the evaluation of the retrieval models is often more challenging. The QPP++ 2023 workshop aimed at fostering discussion in the community regarding these challenges.

The workshop provided a forum for researchers and practitioners to discuss the following key research challenges emerging following the recent advances in IR:

– Can existing QPP techniques be exploited, or which new QPP theories and models need to be devised, for new tasks, such as image retrieval, passage-retrieval, question answering, and conversational search?
– How can new technologies, such as contextualized embeddings, large language models, and neural networks be exploited to improve QPP?
– How should QPP techniques be evaluated, including best practices, datasets, and resources?
– Should QPP be evaluated in the same manner for different IR tasks?
– What changes should we make to the QPP evaluation paradigm to accommodate new domains and IR techniques?

The workshop is expected to have two main outcomes:

- We intend to compile the workshop proceedings from the submitted papers. The proceedings will be published in the CEUR-WS.org proceedings series.
- We intend to draft a position paper describing the roadmap identified during the discussions and submit it to the SIGIR forum.

Organization

Workshop Organizers

Guglielmo Faggioli, *University of Padova, Italy*
Nicola Ferro, *University of Padova, Italy*
Josiane Mothe, *IRIT UMR5505 CNRS, INSPE, Univ. de Toulouse, France*
Fiana Raiber, *Yahoo Research, Israel*

Program Committee Members

PC members have been carefully chosen to represent the community, including representatives from the industry and academia.
Negar Arabzadeh, *University of Waterloo, Canada*
Fabio Giachelle, *University of Padova, Italy*
Claudia Hauff, *Spotify, Netherlands*
Ornella Irrera, *University of Padova, Italy*
Stefano Marchesin, *University of Padova, Italy*
Jian-Yun Nie, *University de Montreal, Canada*
Haggai Roitman, *eBay Research, Israel*
Laure Soulier, *Sorbonne Université-ISIR, France*
Ellen Voorhees, *NIST, US*

Acknowledgments

The QPP++ workshop would not have been possible without the tremendous effort of several people including the program committee members and the participants who contributed their time to making the workshop a success.

Thank you all very much!

April, 2023

Guglielmo Faggioli,
Nicola Ferro,
Josiane Mothe,
Fiana Raiber

References

1. Carmel, D., Yom-Tov, E.: Estimating the query difficulty for information retrieval. Synthesis Lectures on Information Concepts, Retrieval, and Services **2**(1), 1–89 (2010)
2. Deveaud, R., Mothe, J., Ullah, M.Z., Nie, J.Y.: Learning to adaptively rank document retrieval system configurations. ACM Transactions on Information Systems (TOIS) **37**(1), 1–41 (2018)
3. Faggioli, G., Ferro, N., Mothe, J., Raiber, F.: QPP++ 2023: Query-performance prediction and its evaluation in new tasks. In: Kamps, J., Goeuriot, L., Crestani, F., Maistro, M., Joho, H., Davis, B., Gurrin, C., Kruschwitz, U., Caputo, A. (eds.) Advances in Information Retrieval - 45th European Conference on Information Retrieval, ECIR 2023, Dublin, Ireland, April 2-6, 2023, Proceedings, Part III. Lecture Notes in Computer Science, vol. 13982, pp. 388–391. Springer (2023). https://doi.org/10.1007/978-3-031-28241-6_ 42, https://doi.org/10.1007/978-3-031-28241-6_42
4. Faggioli, G., Formal, T., Marchesin, S., Clinchant, S., Ferro, N., Benjamin, P.: Query Performance Prediction for Neural IR: Are We There Yet? In: Proc. ECIR (2023)
5. Faggioli, G., Zendel, O., Culpepper, J.S., Ferro, N., Scholer, F.: An Enhanced Evaluation Framework for Query Performance Prediction. In: Proc. ECIR. pp. 115–129 (2021)
6. Hauff, C.: Predicting the effectiveness of queries and retrieval systems. In: Ph.D. Dissertation. University of Twente. pp. 1–179 (2010)
7. Hauff, C., Azzopardi, L., Hiemstra, D.: The Combination and Evaluation of Query Performance Prediction Methods. In: Proc. ECIR. pp. 301–312 (2009)
8. Mitra, B., Craswell, N.: An Introduction to Neural Information Retrieval. Foundations and Trends in Information Retrieval **13**(1), 1–126 (2018)
9. Mothe, J., Tanguy, L.: Linguistic Features to Predict Query Difficulty. In: Proc. SIGIR. pp. 7–10 (2005)
10. Raiber, F., Kurland, O.: Query-Performance Prediction: Setting the Expectations Straight. In: Proc. SIGIR. pp. 13–22 (2014)
11. Roitman, H.: Enhanced performance prediction of fusion-based retrieval. pp. 195––198. ICTIR '18 (2018)
12. Scholer, F., Garcia, S.: A Case for Improved Evaluation of Query Difficulty Prediction. In: Proc. SIGIR. pp. 640–641 (2009)
13. Shtok, A., Kurland, O., Carmel, D.: Using Statistical Decision Theory and Relevance Models for Query-Performance Prediction. In: Proc. SIGIR. pp. 259–266 (2010)
14. Thomas, P., Scholer, F., Bailey, P., Moffat, A.: Tasks, queries, and rankers in preretrieval performance prediction. pp. 1–4. ADCS 2017 (2017)