

Pseudo-Relevance Feedback in the Era of Dense Retrieval

Nicola Tonellotto

University of Pisa, Italy

Abstract

Pseudo-relevance feedback mechanisms such as relevance models have shown the impact of expanding and re-weighting the users' initial queries exploiting information derived from an initial set of retrieved documents, known as the pseudo-relevant documents. Recently, dense retrieval – through the use of pre-trained language models such as BERT to compute the relevance scores of queries and documents from their contents – has produced significant improvements in the effectiveness of several information retrieval tasks. Up so far, two different dense retrieval families have emerged: the use of single embedded representations for each passage and query, or via multiple representations, for each token in each passage and query. In this talk, we will discuss the first study into the potential for multiple representation dense retrieval to be enhanced using pseudo-relevance feedback, jointly carried out by the University of Pisa and the University of Glasgow.

Speaker Bio

Dr. Nicola Tonellotto (male) is assistant professor at the Information Engineering Department of the University of Pisa since 2019. From 2002 to 2019 he was researcher at the Information Science and Technologies Institute "A. Faedo" of the National Research Council of Italy. His main research interests include Cloud Computing, Web Search, Information Retrieval and Deep Learning. He co-authored more than 80 papers on these topics in peer reviewed international journal and conferences. He is co-recipient of the ACM SIGIR 2015 Best Paper Award.

IIR2022: 12th Italian Information Retrieval Workshop, June 29 - July 1, 2022, Milan, Italy

 nicola.tonellotto@unipi.it (N. Tonellotto)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)