

Re-Thinking Re-Ranking

Sean MacAvaney

University of Glasgow, United Kingdom

Abstract

Re-ranking systems take a “cascading” approach, wherein an initial candidate pool of documents are ranked and filtered to produce a final result list. This approach exhibits a fundamental relevance misalignment problem: the most relevant documents may be filtered out by a prior stage as insufficiently relevant, ultimately reducing recall and limiting the potential effectiveness. In this talk, I challenge the cascading paradigm by proposing methods that efficiently pull in additional potentially-relevant documents during the re-ranking process, using the long-standing Cluster Hypothesis. I demonstrate that these methods can improve the efficiency and effectiveness of both bi-encoder and cross-encoder retrieval models at various operational points. Cascading is dead, long live re-ranking!

Speaker Bio

Sean is a Lecturer in Machine Learning at the University of Glasgow and a member of the Terrier Team. His research primarily focuses on effective and efficient neural retrieval. He completed his PhD at Georgetown University in 2021, where he was a member of the IR Lab and an ARCS Endowed Scholar. He was a co-recipient of the ECIR 2023 Best Short Paper award.

IIR2023: 13th Italian Information Retrieval Workshop, 8th - 9th June 2023, Pisa, Italy

✉ sean.macavaney@glasgow.ac.uk (S. MacAvaney)

🌐 <https://macavaney.us/> (S. MacAvaney)

🆔 0000-0002-8914-2659 (S. MacAvaney)



© 2023 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)