

# Effective and Practical Neural Ranking

Sean MacAvaney\*  
IR Lab, Georgetown University  
*sean@ir.cs.georgetown.edu*

## Abstract

Supervised machine learning methods that use neural networks (“deep learning”) have yielded substantial improvements to a multitude of Natural Language Processing (NLP) tasks in the past decade. Improvements to Information Retrieval (IR) tasks, such as ad-hoc search, lagged behind those in similar NLP tasks, despite considerable community efforts. Although there are several contributing factors, I argue in this dissertation that early attempts were not more successful because they did not properly consider the unique characteristics of IR tasks when designing and training ranking models. I first demonstrate this by showing how large-scale datasets containing weak relevance labels can successfully replace training on in-domain collections. This technique improves the variety of queries encountered when training and helps mitigate concerns of overfitting particular test collections. I then show that dataset statistics available in specific IR tasks can be easily incorporated into neural ranking models alongside the textual features, resulting in more effective ranking models. I also demonstrate that contextualized representations, particularly those from transformer-based language models, considerably improve neural ad-hoc ranking performance. I find that this approach is neither limited to the task of ad-hoc ranking (as demonstrated by ranking clinical reports) nor English content (as shown by training effective cross-lingual neural rankers). These efforts demonstrate that neural approaches can be *effective* for ranking tasks. However, I observe that these techniques are impractical due to their high query-time computational costs. To overcome this, I study approaches for offloading computational cost to index-time, substantially reducing query-time latency. These techniques make neural methods *practical* for ranking tasks. Finally, I take a deep dive into better understanding the linguistic biases of the methods I propose compared to contemporary and traditional approaches. The findings from this analysis highlight potential pitfalls of recent methods and provide a way to measure progress in this area going forward.

**Awarded by:** Georgetown University, Washington, DC, USA **on** 24 March 2021.

**Supervised by:** Nazli Goharian and Ophir Frieder.

**Available at:** <http://ir.cs.georgetown.edu/downloads/diss-sean-macavaney.pdf>.

---

\* Currently at the University of Glasgow.

---

## Selected Publications

- Sean MacAvaney, Andrew Yates, Arman Cohan, and Nazli Goharian. CEDR: Contextualized embeddings for document ranking. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1101–1104, 2019a. doi: 10.1145/3331184.3331317. URL <https://dl.acm.org/doi/10.1145/3331184.3331317>.
- Sean MacAvaney, Andrew Yates, Kai Hui, and Ophir Frieder. Content-based weak supervision for ad-hoc re-ranking. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 993–996, 2019b. doi: 10.1145/3331184.3331316. URL <https://dl.acm.org/doi/10.1145/3331184.3331316>.
- Sean MacAvaney, Arman Cohan, and Nazli Goharian. SLEDGE: A simple yet effective baseline for covid-19 scientific knowledge search. *arXiv*, abs/2005.02365, 2020a. URL <https://arxiv.org/abs/2005.02365>. See also: extended version published in EMNLP 2020 (<https://www.aclweb.org/anthology/2020.emnlp-main.341/>).
- Sean MacAvaney, Franco Maria Nardini, Raffaele Perego, Nicola Tonellotto, Nazli Goharian, and Ophir Frieder. Efficient document re-ranking for transformers by precomputing term representations. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 49–58, 2020b. doi: 10.1145/3397271.3401093. URL <https://dl.acm.org/doi/10.1145/3397271.3401093>.
- Sean MacAvaney, Franco Maria Nardini, Raffaele Perego, Nicola Tonellotto, Nazli Goharian, and Ophir Frieder. Expansion via prediction of importance with contextualization. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1573–1576, 2020c. doi: 10.1145/3397271.3401262. URL <https://dl.acm.org/doi/10.1145/3397271.3401262>.