Neural Models for Information Retrieval without Labeled Data

Hamed Zamani
Center for Intelligent Information Retrieval
University of Massachusetts Amherst
Amherst, MA 01003
zamani@cs.umass.edu

Abstract

Recent developments of machine learning models, and in particular deep neural networks, have yielded significant improvements on several computer vision, natural language processing, and speech recognition tasks. Progress with information retrieval (IR) tasks has been slower, however, due to the lack of large-scale training data as well as neural network models specifically designed for effective information retrieval [9]. In this dissertation, we address these two issues by introducing task-specific neural network architectures for a set of IR tasks and proposing novel unsupervised or *weakly supervised* solutions for training the models. The proposed learning solutions do not require labeled training data. Instead, in our weak supervision approach, neural models are trained on a large set of noisy and biased training data obtained from external resources, existing models, or heuristics.

We first introduce relevance-based embedding models [3] that learn distributed representations for words and queries. We show that the learned representations can be effectively employed for a set of IR tasks, including query expansion, pseudo-relevance feedback, and query classification [1, 2].

We further propose a standalone learning to rank model based on deep neural networks [5, 8]. Our model learns a sparse representation for queries and documents. This enables us to perform efficient retrieval by constructing an inverted index in the learned semantic space. Our model outperforms state-of-the-art retrieval models, while performing as efficiently as term matching retrieval models.

We additionally propose a neural network framework for predicting the performance of a retrieval model for a given query [7]. Inspired by existing query performance prediction models, our framework integrates several information sources, such as retrieval score distribution and term distribution in the top retrieved documents. This leads to state-of-the-art results for the performance prediction task on various standard collections.

We finally bridge the gap between retrieval and recommendation models, as the two key components in most information systems. Search and recommendation often share the same goal: helping people get the information they need at the right time. Therefore, joint modeling and optimization of search engines and recommender systems could potentially benefit both systems [4]. In more detail, we introduce a retrieval model that is trained

using user-item interaction (e.g., recommendation data), with no need to query-document relevance information for training [6].

Our solutions and findings in this dissertation smooth the path towards learning efficient and effective models for various information retrieval and related tasks, especially when large-scale training data is not available.

Doctoral Supervisor: W. Bruce Croft

The dissertation is available at: `http://bit.ly/2CGVOPn`

References

[1] H. Zamani and W. B. Croft. Embedding-based Query Language Models. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, ICTIR '16, pages 147–156, Newark, Delaware, USA, 2016. ACM.

[2] H. Zamani and W. B. Croft. Estimating Embedding Vectors for Queries. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, ICTIR '16, pages 123–132, Newark, Delaware, USA, 2016. ACM.

[3] H. Zamani and W. B. Croft. Relevance-based Word Embedding. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, pages 505–514, Shinjuku, Tokyo, Japan, 2017. ACM.

[4] H. Zamani and W. B. Croft. Joint Modeling and Optimization of Search and Recommendation. In *Proceedings of the First International Conference on Design of Experimental Search and Information Retrieval Systems*, DESIRES '18, pages 36–41, Bertinoro, Italy, 2018. CEUR.

[5] H. Zamani and W. B. Croft. On the Theory of Weak Supervision for Information Retrieval. In *Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieval*, ICTIR '18, pages 147–154, Tianjin, China, 2018. ACM.

[6] H. Zamani and W. B. Croft. Learning a Joint Search and Recommendation Model from User-Item Interactions. In *Proceedings of the 13th ACM International Conference on Web Search and Data Mining*, WSDM '20, Houston, TX, USA, 2020. ACM. (to appear).

[7] H. Zamani, W. B. Croft, and J. S. Culpepper. Neural Query Performance Prediction Using Weak Supervision from Multiple Signals. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '18, pages 105–114, Ann Arbor, MI, USA, 2018. ACM.

[8] H. Zamani, M. Dehghani, W. B. Croft, E. Learned-Miller, and J. Kamps. From Neural Re-Ranking to Neural Ranking: Learning a Sparse Representation for Inverted Indexing. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, pages 497–506, Torino, Italy, 2018. ACM.

[9] H. Zamani, M. Dehghani, F. Diaz, H. Li, and N. Craswell. SIGIR 2018 Workshop on Learning from Limited or Noisy Data for Information Retrieval. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '18, pages 1439–1440, Ann Arbor, MI, USA, 2018. ACM.