Neural Generative Models and Representation Learning for Information Retrieval

Qingyao Ai
University of Massachusetts Amherst
aiqy@cs.umass.edu

Abstract

Information Retrieval (IR) concerns about the structure, analysis, organization, storage, and retrieval of information. Among different retrieval models proposed in the past decades, generative retrieval models, especially those under the statistical probabilistic framework, are one of the most popular techniques that have been widely applied to Information Retrieval problems. While they are famous for their well-grounded theory and good empirical performance in text retrieval, their applications in IR are often limited by their complexity and low extendability in the modeling of high-dimensional information. Recently, advances in deep learning techniques provide new opportunities for representation learning and generative models for information retrieval. In contrast to statistical models, neural models have much more flexibility because they model information and data correlation in latent spaces without explicitly relying on any prior knowledge. Previous studies on pattern recognition and natural language processing have shown that semantically meaningful representations of text, images, and many types of information can be acquired with neural models through supervised or unsupervised training. Nonetheless, the effectiveness of neural models for information retrieval is mostly unexplored. In this thesis, we study how to develop new generative models and representation learning frameworks with neural models for information retrieval. Specifically, our contributions include three main components: (1) *Theoretical Analysis*: We present the first theoretical analysis and adaptation of existing neural embedding models for ad-hoc retrieval tasks; (2) *Design Practice*: Based on our experience and knowledge, we show how to design an embedding-based neural generative model for practical information retrieval tasks such as personalized product search; And (3) *Generic Framework*: We further generalize our proposed neural generative framework for complicated heterogeneous information retrieval scenarios that concern text, images, knowledge entities, and their relationships. Empirical results show that the proposed neural generative framework can effectively learn information representations and construct retrieval models that outperform the state-of-the-art systems in a variety of IR tasks.

Advisor: W. Bruce Croft
Committee Members: James Allan, Erik G. Learned-Miller, Rajesh Bhatt, Jiafeng Guo
The dissertation is available at `https://scholarworks.umass.edu/dissertations_2/1676/`