

Paper Summaries

Relational Collaborative Filtering: Modeling Multiple Item Relations for Recommendation <https://arxiv.org/pdf/1904.12796.pdf>

Authors proposed Relational Collaborative Filtering (RCF) that exploits multiple item relations in recommender systems. Both the relation type and the relation value are crucial and a two-level hierarchical attention mechanism is developed to model user preference. The first-level attention extracts important relations, and the second-level attention estimates the contribution of a historical item by considering the specific relation values. The item relations are preserved, and jointly trained with user preference modeling. Experiments are performed on two public datasets: MovieLens3 and KKBox.

Learning to link with Wikipedia

<https://www.cs.waikato.ac.nz/~ihw/papers/08-DNM-IHW-LearningToLinkWithWikipedia.pdf>

The paper proposes a method to automatically cross-reference documents with Wikipedia. It explains how significant terms can be identified within unstructured text and enriched with links to the appropriate Wikipedia articles using Machine Learning. A machine-learning approach to disambiguation is used. It balances the commonness (i.e. prior probability) of a sense with its relatedness to the surrounding context. The disambiguation classifier was evaluated on dataset created by gathering 11,000 anchors from 100 randomly selected articles and disambiguated automatically.

Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs <https://arxiv.org/pdf/1908.00469.pdf>

Authors presented their unsupervised method QUEST that answers complex questions directly from textual sources on-the-fly, by computing similarity joins over partial results from different documents. A noisy quasi Knowledge Graph with node and edge weights is built consisting of

dynamically retrieved entity names and relational phrases. The graph is augmented with types and semantic alignments and the best answers are computed by an algorithm for Group Steiner Trees.

BERT with History Answer Embedding for Conversational Question Answering

Authors presented history answer embeddings for conversational question answering. Conversation history is integrated into a conversational question answering model built on BERT. There are two aspects of the proposed model defined by the general framework. History selection selects a subset of the history turns that are more helpful than others. Second is history modeling to model the selected history turns in an MC model.

Deeper Text Understanding for IR with Contextual Neural Language Modeling

In the paper, BERT is used to provide deeper text understanding for IR and experimental results demonstrate that the BERT contextual text representations are more effective than traditional word embeddings. Query tokens and the document tokens with token [SEP] separating the two segments are input. Because the document text can be long, passage level approach for document retrieval is used where a document is split into overlapping passages. The relevance of each passage is predicted independently by the neural ranker. Document score is the score of the first passage, the best passage, or the sum of all passage scores. BERT is augmented with search knowledge by tuning it on a large search log. Robust04 corpus and ClueWeb09-B corpuses were used for experiments.

CEDR: Contextualized Embeddings for Document Ranking

<https://arxiv.org/pdf/1904.07094.pdf>

Authors investigated the utilization of two pretrained contextualized language models (ELMo and BERT) for ad-hoc document ranking. Experiments were done on Trec benchmarks. A joint approach incorporating BERT's classification vector into existing neural models was proposed and it outperformed state-of-the-art ad-hoc ranking baselines. The joint approach is CEDR(Contextualized Embeddings for Document Ranking).

The Benefits of Magnitude Estimation Relevance Assessments for Information Retrieval Evaluation <https://dl.acm.org/doi/10.1145/2766462.2767760>

The use of magnitude estimation in judging the relevance of documents in the context of information retrieval evaluation is investigated. Magnitude estimation is a scaling technique for the measurement of sensation, and numbers are assigned by observers to stimuli in response to the perceived intensity. A large-scale user study is carried out across 18 TREC topics and more than 50,000 magnitude estimation judgments were collected. The analysis shows that on average, magnitude estimation judgments are rank-aligned with ordinal judgments made by expert relevance assessors.