Round 1 results — Run irc_entrez submitted from IRC

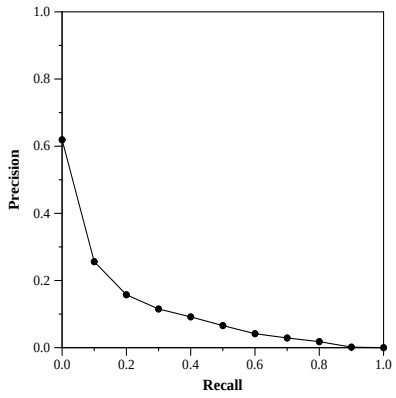
Run Description

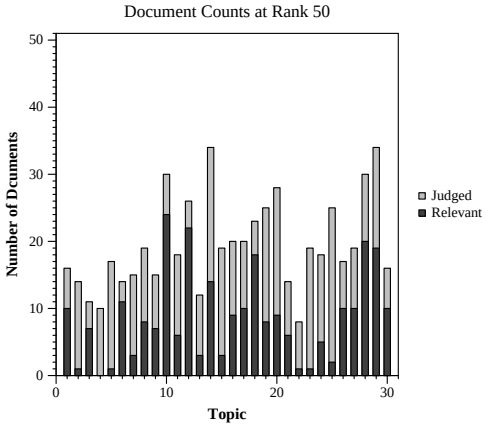
As part of TREC-COVID, we submit automatic runs based on (pseudo) relevance feedback in combination with a reranking approach. The reranker is trained on relevance feedback data that is retrieved from PubMed/PubMed Central (PMC). The training data is retrieved with queries using the contents of the '<query>' tags only. For each topic a new reranker is trained. We consider those documents retrieved by the specific topic query as relevant training data, and the documents of the other 29 topics as non-relevant training data. Given a baseline run, the trained system reranks documents. The baseline run is retrieved with the default ranker of Elasticsearch/Lucene (BM25) and queries using the contents of the '<query>' tags only. For our reranker we use GloVe embeddings in combination with the Deep Relevance Matching Model (DRMM). Our three run submissions differ by the way training data is retrieved from PubMed/PMC. 'irc_entrez': This run is trained on titles and abstracts retrieved from the Entrez Programming Utilities API with "type=relevance".

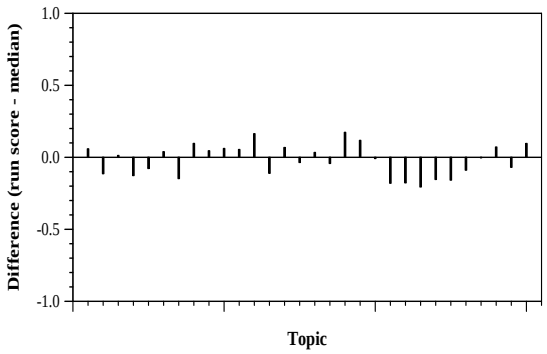
Summary Statistics	
Run ID	irc_entrez
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	1162
MAP	0.0988
Mean Bpref	0.3054
Mean NDCG@10	0.2533

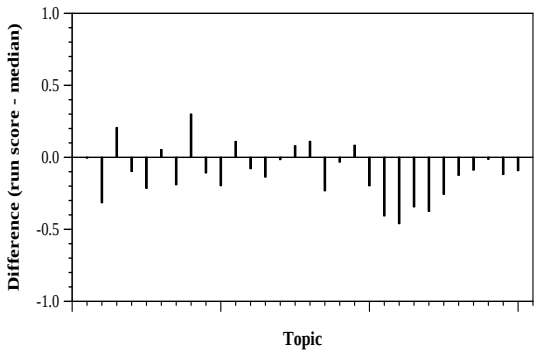
Document Level Averages	
	Precision
At 5 docs	0.3667
At 10 docs	0.3033
At 15 docs	0.2622
At 20 docs	0.2400
At 30 docs	0.2044
R-Precision	
Exact	0.1463





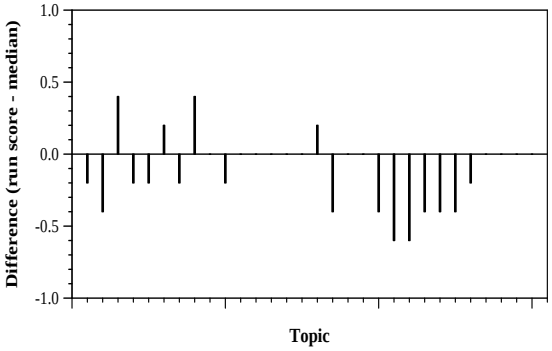


Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs

Round 1 results — Run irc_entrez submitted from IRC



Per-topic difference from median P@5 for all Round 1 runs