

Round 4 results — Run UPrrf38rrf2-r4 submitted from unique_ptr

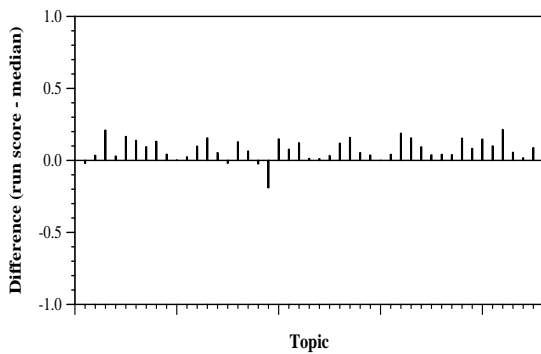
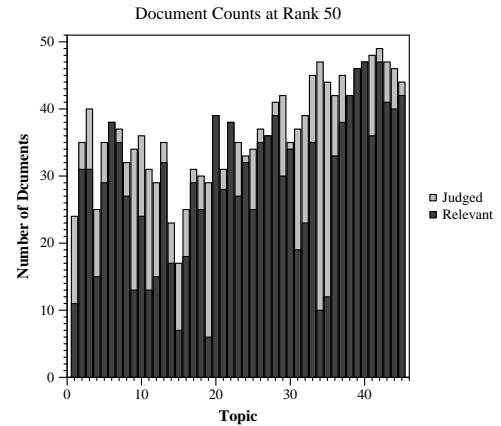
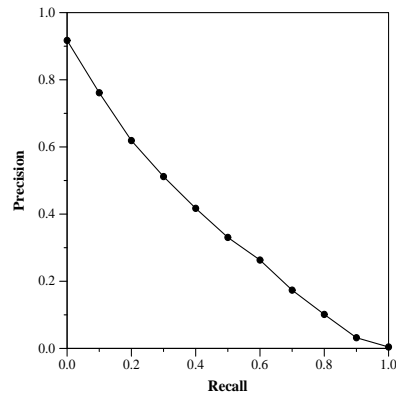
Run Description

This run is produced by Reciprocal Rank Fusion (Cormack et al., 2009) of (a) Terrier & Anserini retrieval runs and neural retrieval runs based on synthetic query generation (<https://arxiv.org/abs/2004.14503>) with various combinations of query, question and narrative (b) TF-Ranking + ELECTRA with softmax loss (<https://arxiv.org/abs/2004.08476>) trained on MS-Marco dataset. This is a fully automatic run, with no extra-tuning based on the existing judgments from TREC-COVID.

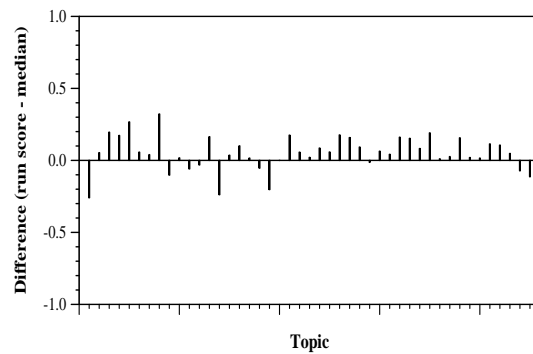
Summary Statistics	
Run ID	UPrrf38rrf2-r4
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	45
Total number retrieved	44982
Total relevant	5824
Total relevant retrieved	4334
MAP	0.3612
Mean Bpref	0.5917
Mean NDCG@20	0.6792
Mean RBP(p=0.5)	0.7602 +0.0067

Document Level Averages	
	Precision
At 5 docs	0.8133
At 10 docs	0.7844
At 15 docs	0.7437
At 20 docs	0.7156
At 30 docs	0.6541
R-Precision	
Exact	0.3959

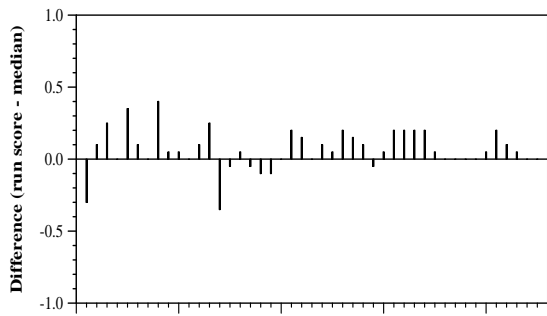


Per-topic difference from median bpref for all Round 4 runs

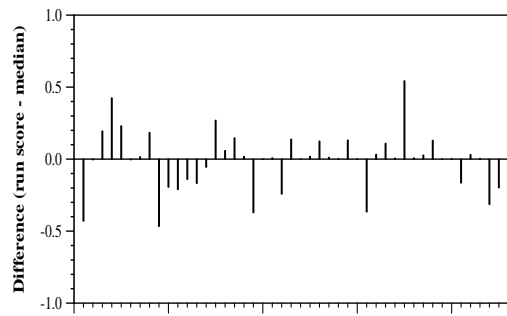


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run UPrrf38rrf2-r4 submitted from unique_ptr



Topic
Per-topic difference from median P@20 for all Round 4 runs



Topic
Per-topic difference from median RBP(p=0.5) for all Round 4 runs