

Round 5 results — Run UPrrf89-r5 submitted from unique_ptr

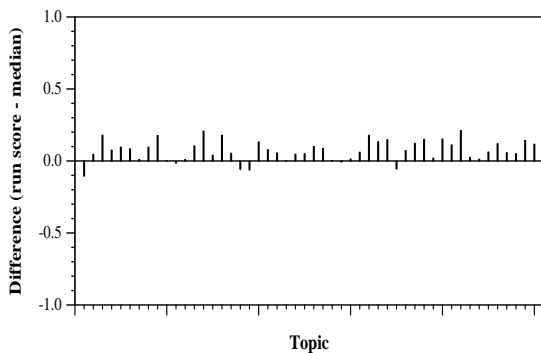
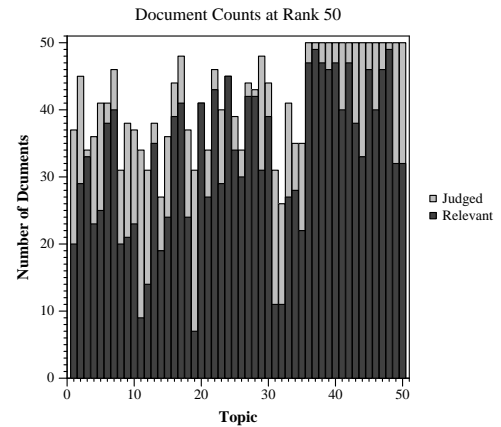
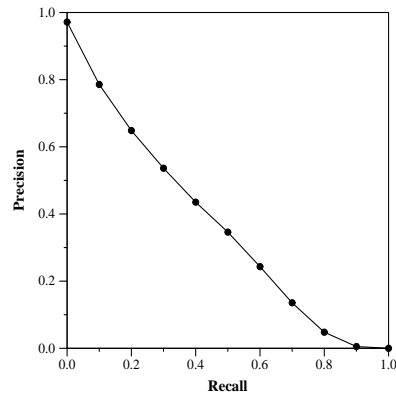
Run Description

A reciprocal rank fusion of (a) Anserini & Terrier runs based on different indexing schemes and topic variants (b) Neural retrieval runs based on synthetic query generation (<https://arxiv.org/abs/2004.14503>) (c) TF-Ranking + multiple BERT variants with softmax loss (<https://arxiv.org/abs/2004.08476>) trained on MS-Marco (d) TF-Ranking + multiple BERT variants with softmax loss trained on BIOASQ data (<https://arxiv.org/abs/1809.01682>). This is a fully automatic run, with no extra tuning based on the existing judgments from TREC-COVID.

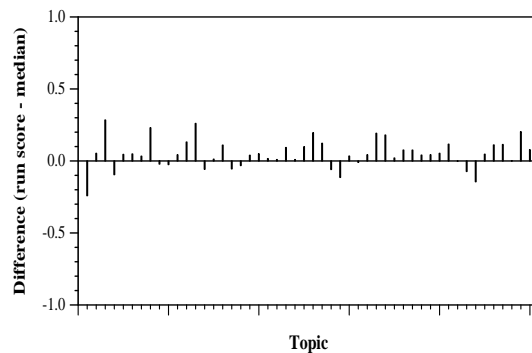
Summary Statistics	
Run ID	UPrrf89-r5
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	50
Total number retrieved	49901
Total relevant	10910
Total relevant retrieved	7508
MAP	0.3612
Mean Bpref	0.5364
Mean NDCG@20	0.7235
Mean RBP(p=0.5)	0.8022 +0.0003

Document Level Averages	
	Precision
At 5 docs	0.8680
At 10 docs	0.8320
At 15 docs	0.7880
At 20 docs	0.7590
At 30 docs	0.7133
R-Precision	
Exact	0.4073

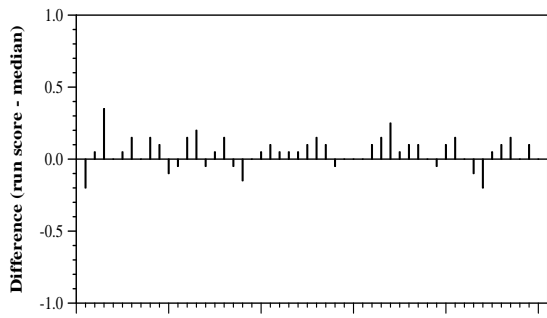


Per-topic difference from median bpref for all Round 4 runs

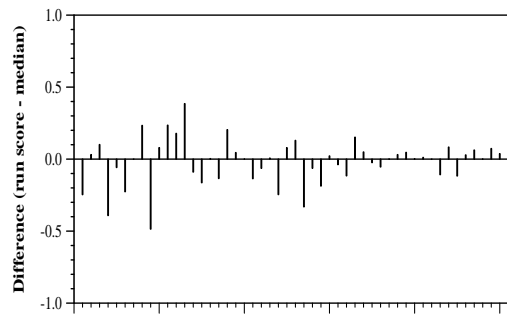


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 5 results — Run UPrrf89-r5 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