

Round 4 results — Run mpiid5_run1 submitted from mpiid5

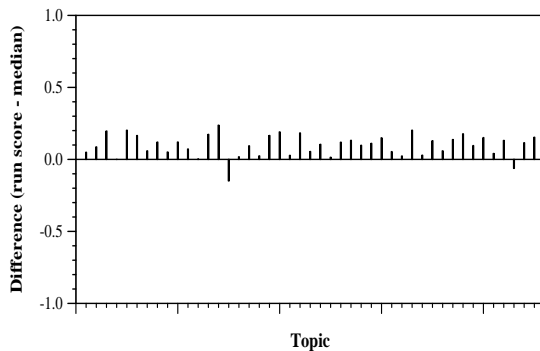
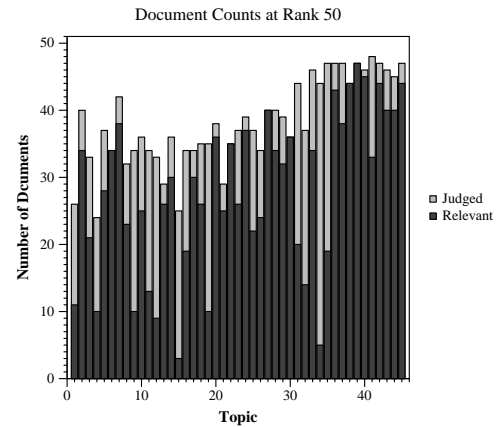
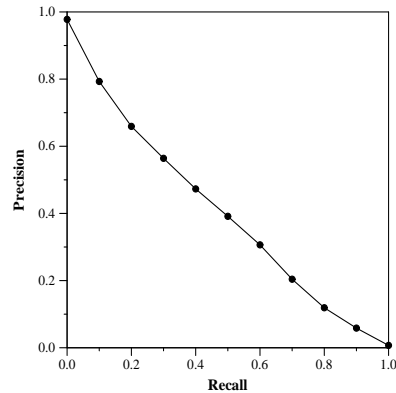
Run Description

We re-rank the top 100 then the top 101-2000 documents from the Anserini fusion2 baseline (8th row). For the re-ranking method, we use the ELECTRA-Base model fine-tuned on the MSMARCO passage dataset. The model is later fine-tuned on the TREC COVID round 1-3 full-text collection. We use the question queries for re-ranking.

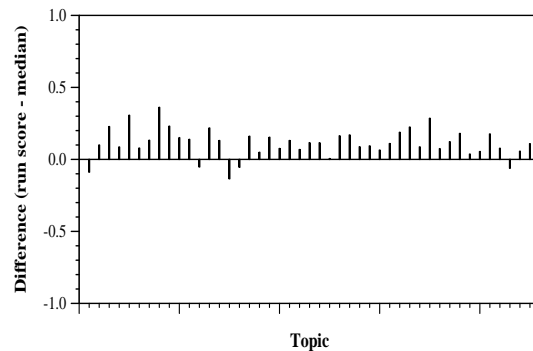
Summary Statistics	
Run ID	mpiid5_run1
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	45
Total number retrieved	44980
Total relevant	5824
Total relevant retrieved	4285
MAP	0.3993
Mean Bpref	0.6132
Mean NDCG@20	0.7391
Mean RBP(p=0.5)	0.8622 +0.0004

Document Level Averages	
	Precision
At 5 docs	0.9067
At 10 docs	0.8422
At 15 docs	0.7926
At 20 docs	0.7589
At 30 docs	0.6585
R-Precision	
Exact	0.4332

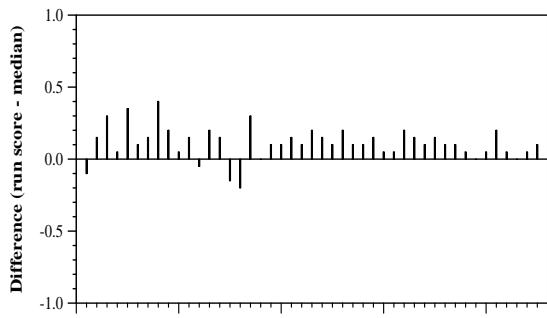


Per-topic difference from median bpref for all Round 4 runs

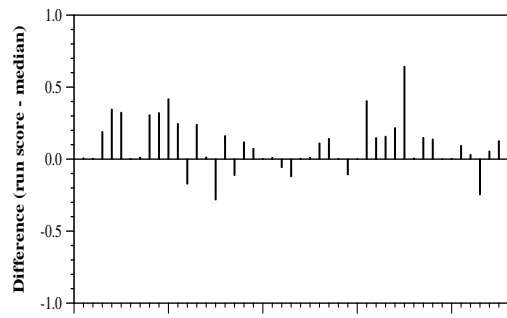


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run mpiid5_run1 submitted from mpiid5



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