

Round 5 results — Run covidex.r5.2s submitted from covidex

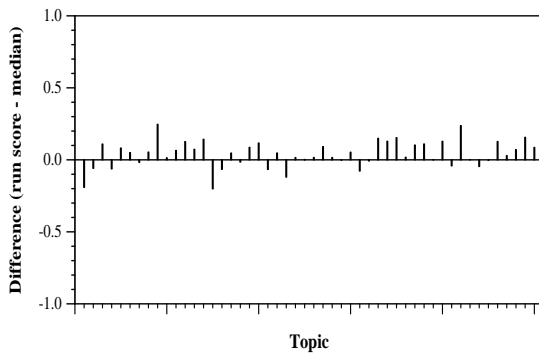
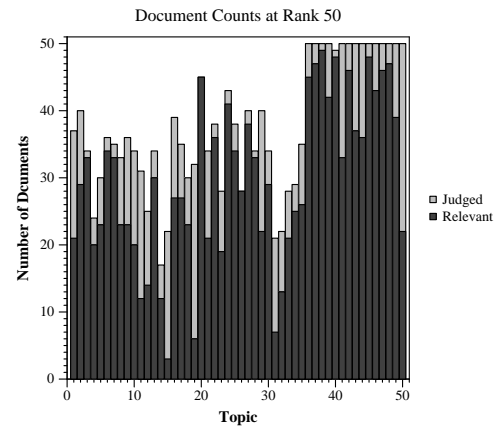
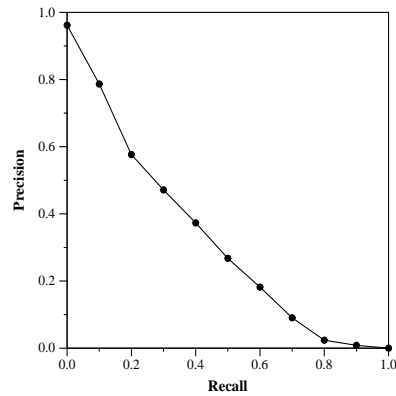
Run Description

A reciprocal rank fusion of two baseline runs that were independently reranked by a pointwise reranker (monoT5) then by a pairwise reranker (duoT5). Both rerankers were trained on MedMARCO (MacAvaney et al., SLEDGE, 2020). Baseline runs are rows 7 and 8 from <https://github.com/castorini/anserini/blob/master/docs/experiments-covid.md#round-5>

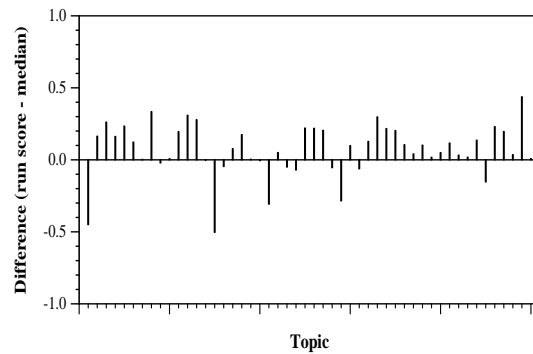
Summary Statistics	
Run ID	covidex.r5.2s
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	50
Total number retrieved	49900
Total relevant	10910
Total relevant retrieved	6722
MAP	0.3212
Mean Bpref	0.5052
Mean NDCG@20	0.7457
Mean RBP(p=0.5)	0.8447 +0.0006

Document Level Averages	
	Precision
At 5 docs	0.8960
At 10 docs	0.8540
At 15 docs	0.8053
At 20 docs	0.7610
At 30 docs	0.6793
R-Precision	
Exact	0.3721

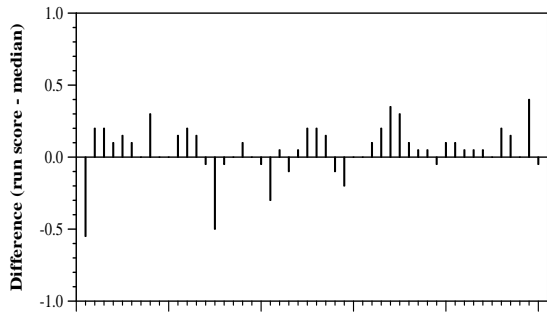


Per-topic difference from median bpref for all Round 4 runs

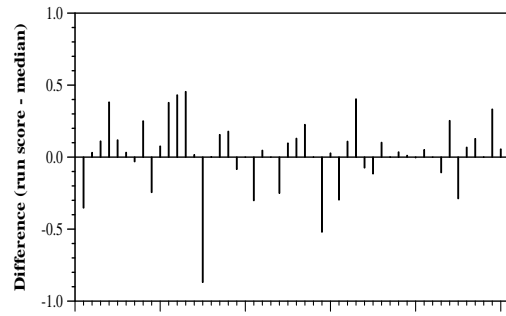


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 5 results — Run covidex.r5.2s submitted from covidex



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



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