

Round 2 results — Run covidex.t5 submitted from covidex

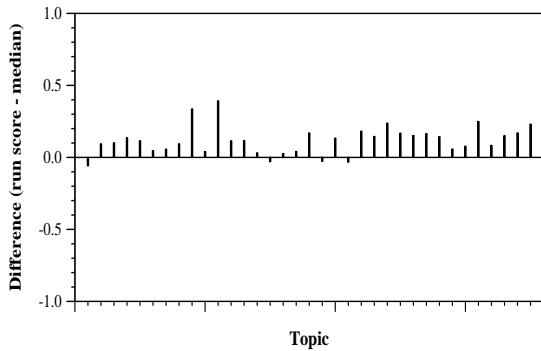
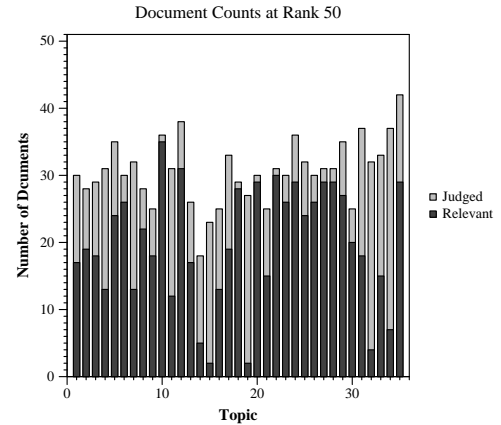
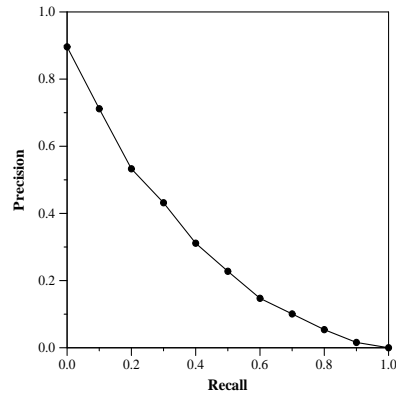
Run Description

Reciprocal rank fusion of two runs: Anserini r2.fusion1, reranked with medT5-3B; Anserini r2.fusion2, reranked with medT5-3B; Anserini fusion baselines for round 2: <https://github.com/castorini/anserini/blob/master/docs/exp/covid.md> medT5-3B: a T5-3B reranker fine-tuned on MS MARCO then fine-tuned (again) on MS MARCO medical subset (from MacAvaney et al., 2020, "SLEDGE")

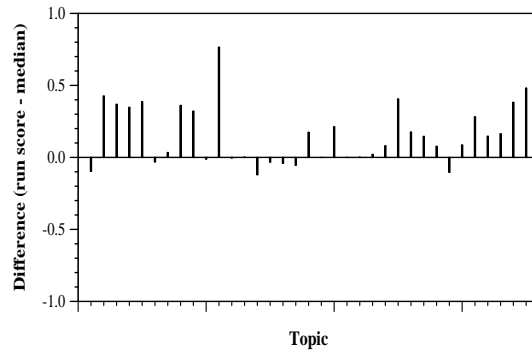
Summary Statistics	
Run ID	covidex.t5
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	35
Total number retrieved	30506
Total relevant	3002
Total relevant retrieved	1963
MAP	0.2880
Mean Bpref	0.4876
Mean NDCG@10	0.6250
Mean RBP(p=0.5)	0.6634 +0.0013

Document Level Averages	
	Precision
At 5 docs	0.7314
At 10 docs	0.6629
At 15 docs	0.6210
At 20 docs	0.5743
At 30 docs	0.4886
R-Precision	
Exact	0.3245

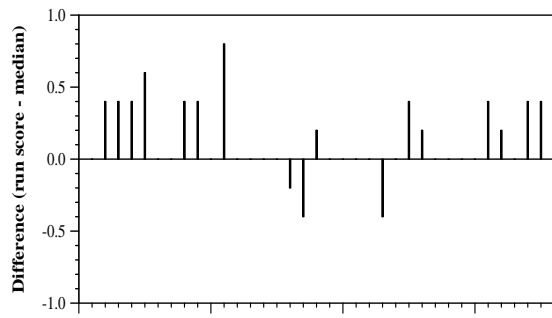


Per-topic difference from median bpref for all Round 2 runs

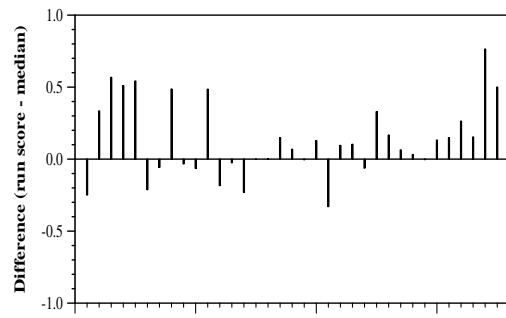


Per-topic difference from median NDCG@10 for all Round 2 runs

Round 2 results — Run covidex.t5 submitted from covidex



Per-topic difference from median P@5 for all Round 2 runs



Per-topic difference from median RBP(p=0.5) for all Round 2 runs