Round 3 results — Run cord19.vespa.ai-gb-2 submitted from cord19.vespa.ai
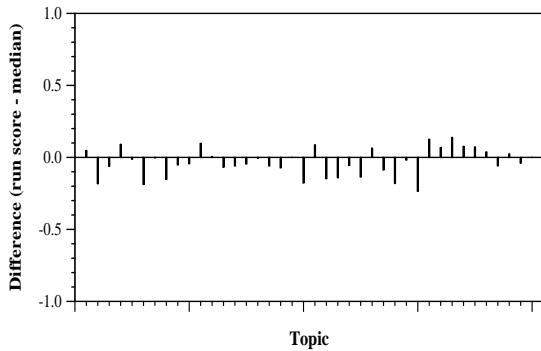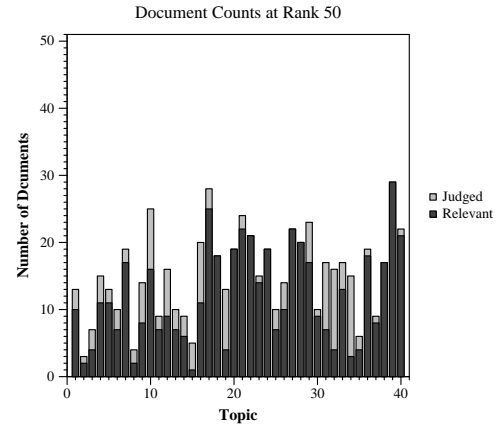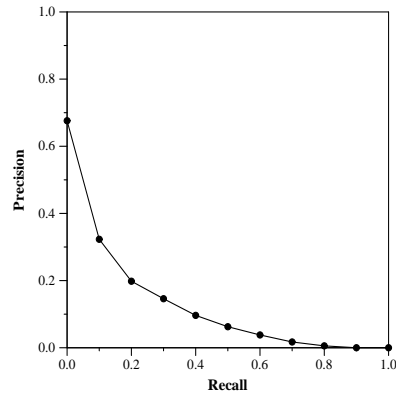
Run Description

Retrieval using entities from query, question and narrative. First-phase ranking using bm25(title) + bm25(abstract) Top 1000 hits from first phase re-ranked using a GBDT model (LightGBM with lambdarank objective). Model trained on all available judgements from round 2 and round 3. https://github.com/vespa-engine/cord-19/tree/master/trec-covid
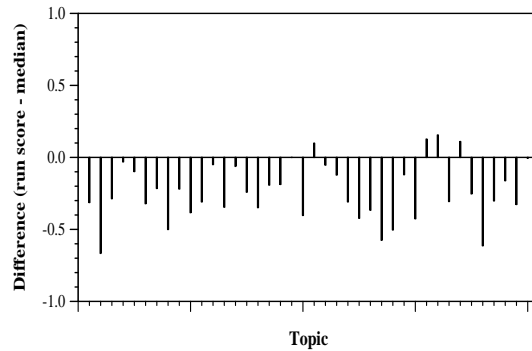
Summary Statistics	
Run ID	cord19.vespa.ai-gb-2
Topic type	feedback
Contributed to judgment sets?	no

Overall measures	
Number of topics	40
Total number retrieved	39941
Total relevant	4698
Total relevant retrieved	2384
MAP	0.1117
Mean Bpref	0.4321
Mean NDCG@10	0.3543
Mean RBP(p=0.5)	0.3810 +0.4741

Document Level Averages	
	Precision
At 5 docs	0.4100
At 10 docs	0.3875
At 15 docs	0.3450
At 20 docs	0.3138
At 30 docs	0.2867
R-Precision	
Exact	0.1806

Document Counts at Rank 50
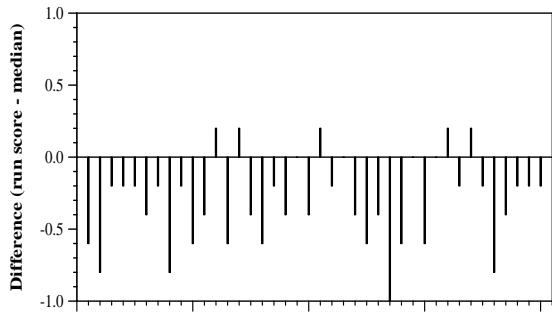

Per-topic difference from median bpref for all Round 3 runs

Per-topic difference from median NDCG@10 for all Round 3 runs

Per-topic difference from median P@5 for all Round 3 runs

Per-topic difference from median RBP(p=0.5) for all Round 3 runs