

Round 1 results — Run run3 submitted from GUIR\_S2

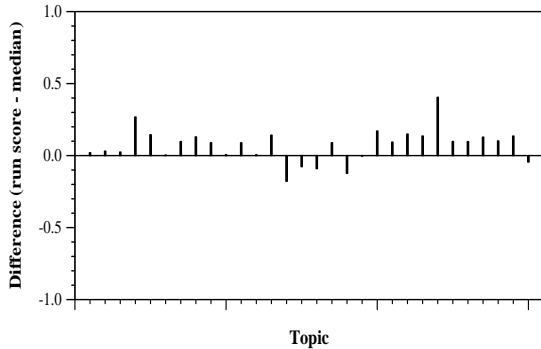
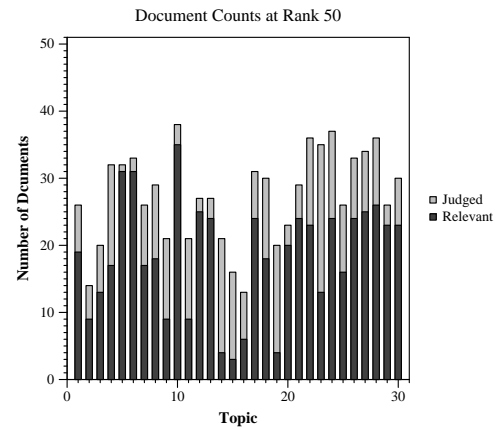
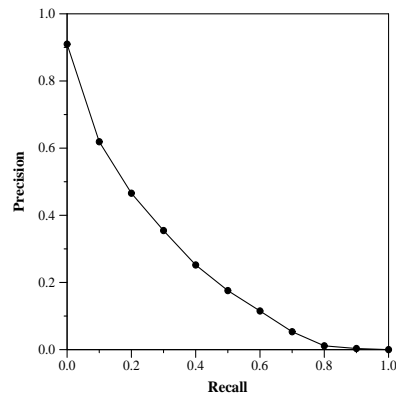
Run Description

Initial ranking with BM25 and question over title and abstract re-ranking using SciBERT (trained on MS-MARCO) using title and abstract with the question. Filtered to 2020 articles only.

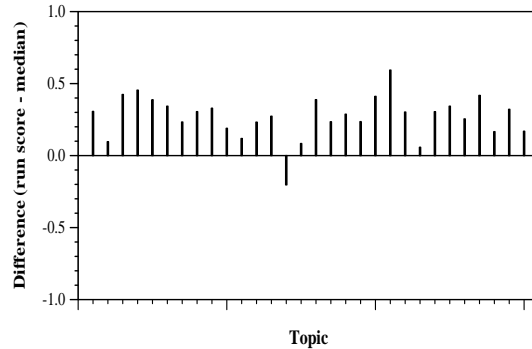
Summary Statistics	
Run ID	run3
Topic type	manual
Contributed to judgment sets?	no

Overall measures	
Number of topics	30
Total number retrieved	14996
Total relevant	2352
Total relevant retrieved	1153
MAP	0.2426
Mean Bpref	0.3969
Mean NDCG@10	0.6273

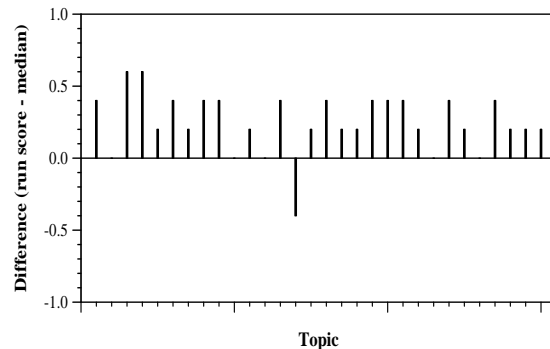
Document Level Averages	
	Precision
At 5 docs	0.7333
At 10 docs	0.6467
At 15 docs	0.5867
At 20 docs	0.5467
At 30 docs	0.4756
R-Precision	
Exact	0.3148



Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs



Per-topic difference from median P@5 for all Round 1 runs