Run Description

Fusion of SciBERT over abstracts, SPECTER feedback (top 5 SciBERT documents), and BM25 (paragraph). SciBERT model trained on MS-MARCO. Tuning conducted on sample of COVID-QA queries.
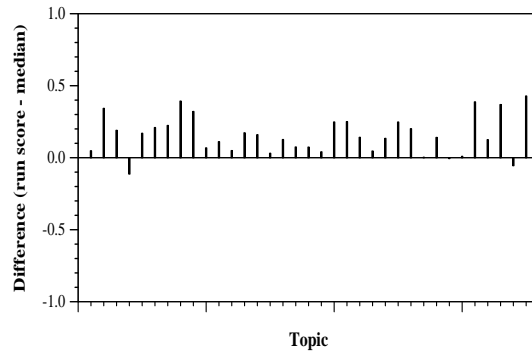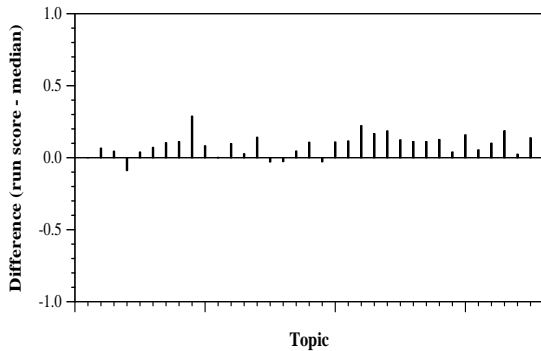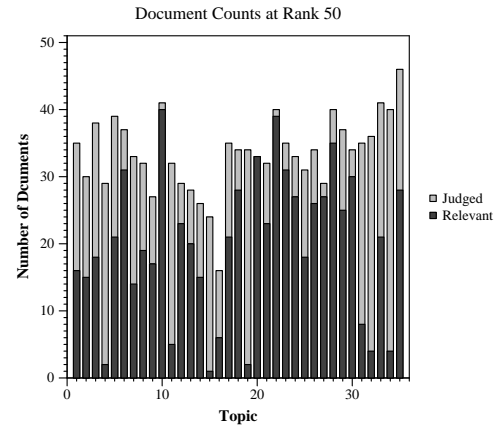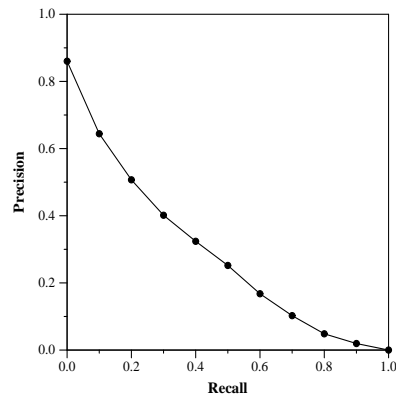
Summary Statistics	
Run ID	GUIR_S2_run1
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	35
Total number retrieved	30727
Total relevant	3002
Total relevant retrieved	1981
MAP	0.2842
Mean Bpref	0.4569
Mean NDCG@10	0.6251
Mean RBP(p=0.5)	0.6787 +0.0041

Document Level Averages	
	Precision
At 5 docs	0.7486
At 10 docs	0.6571
At 15 docs	0.5943
At 20 docs	0.5500
At 30 docs	0.4857
R-Precision	
Exact	0.3253

Document Counts at Rank 50

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

Per-topic difference from median NDCG@10 for all Round 2 runs
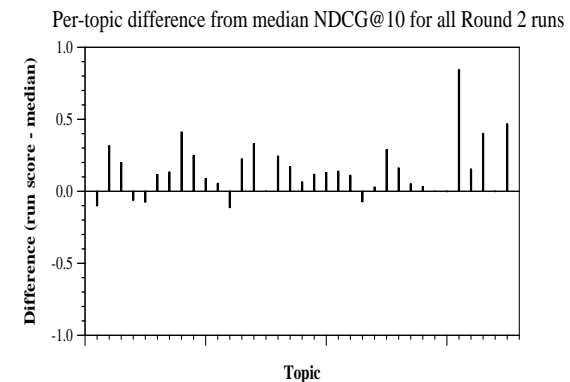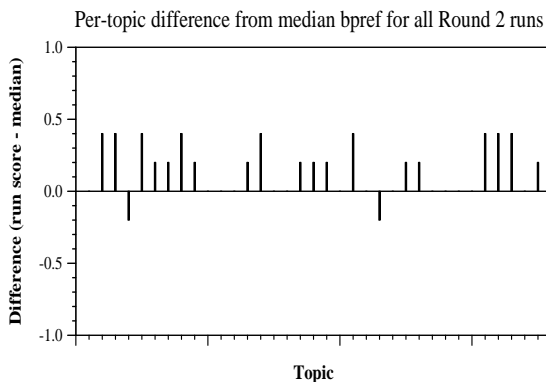
Per-topic difference from median P@5 for all Round 2 runs

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