

Round 4 results — Run run1_C_A_SciB submitted from CIR

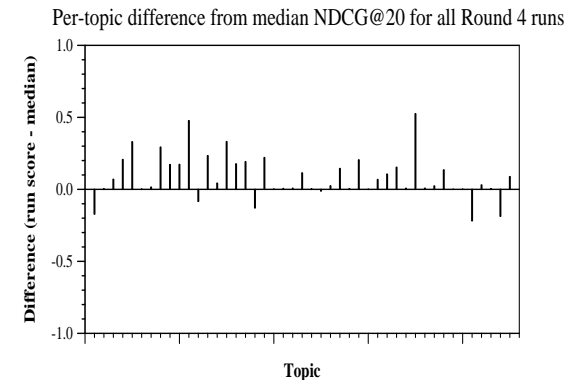
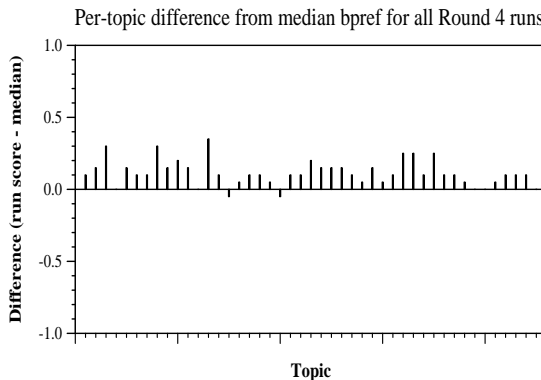
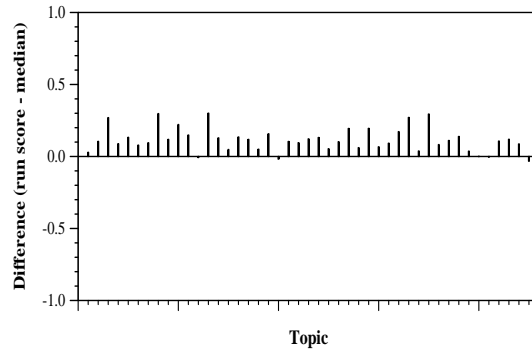
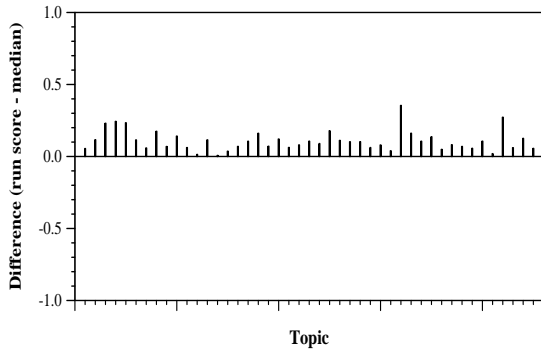
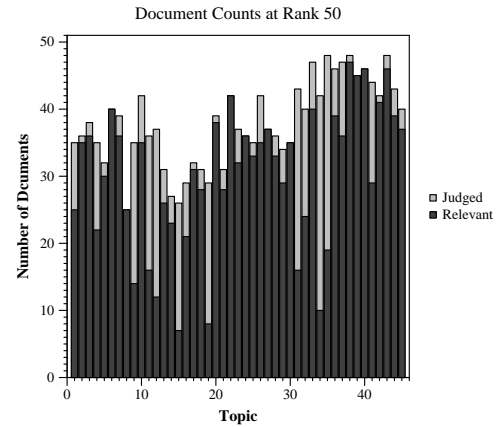
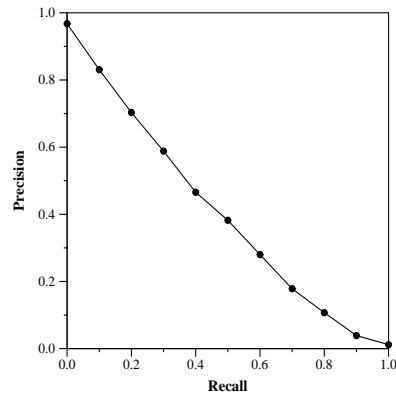
Run Description

Fusion of: (1) Anserini BM25 baseline(r4.rf), (2) Med-MARCO ANN dense embedding paragraph retrieval, (3) MedMARCP SciBERT paragraph reranker(top 1000). Fusion weights tuned. Language model finetuned on Round-4 documents using MLM task.

Summary Statistics	
Run ID	run1_C_A_SciB
Topic type	feedback
Contributed to judgment sets?	no

Overall measures	
Number of topics	45
Total number retrieved	44978
Total relevant	5824
Total relevant retrieved	4313
MAP	0.3992
Mean Bpref	0.6256
Mean NDCG@20	0.7420
Mean RBP(p=0.5)	0.8551 +0.0053

Document Level Averages	
	Precision
At 5 docs	0.8622
At 10 docs	0.8400
At 15 docs	0.8074
At 20 docs	0.7633
At 30 docs	0.7067
R-Precision	
Exact	0.4320



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