

Round 4 results — Run run2_Crf_A_SciB_MAP submitted from CIR

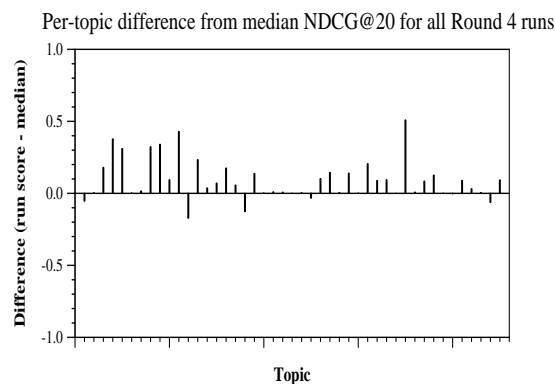
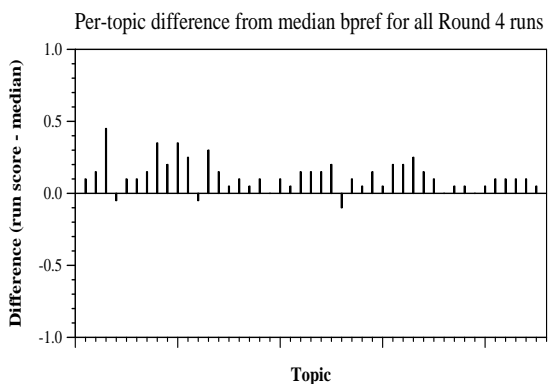
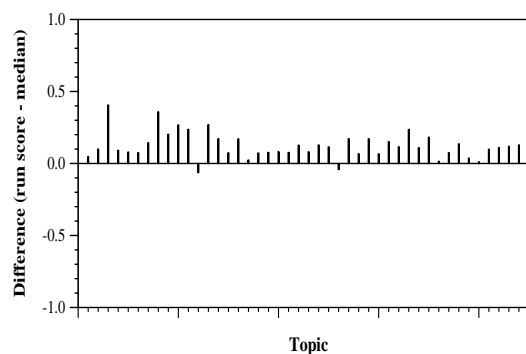
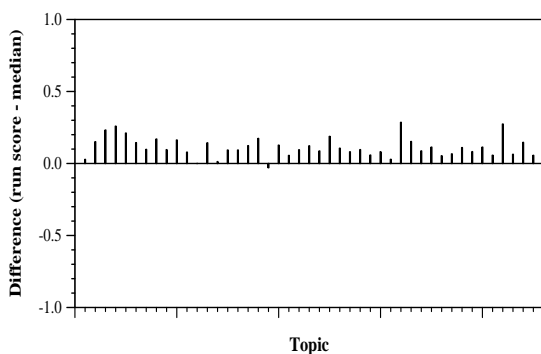
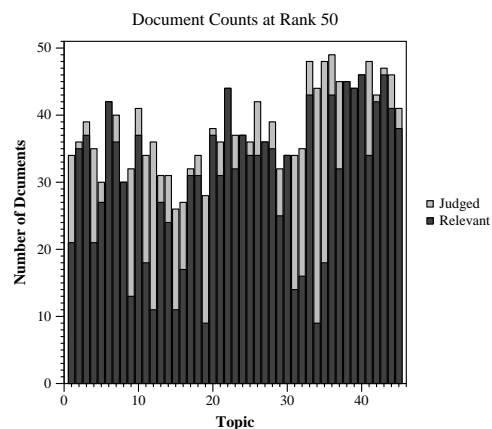
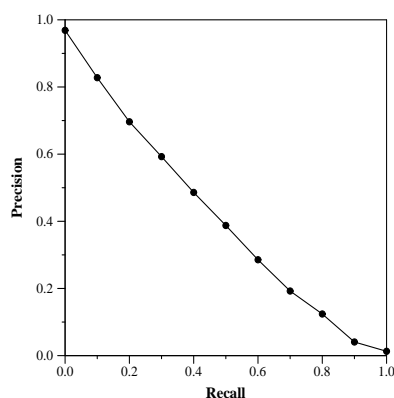
Run Description

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

Summary Statistics	
Run ID	run2_Crf_A_SciB_MAP
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	45
Total number retrieved	44979
Total relevant	5824
Total relevant retrieved	4373
MAP	0.4079
Mean Bpref	0.6292
Mean NDCG@20	0.7470
Mean RBP(p=0.5)	0.8612 +0.0002

Document Level Averages	
	Precision
At 5 docs	0.9022
At 10 docs	0.8511
At 15 docs	0.7985
At 20 docs	0.7700
At 30 docs	0.7111
R-Precision	
Exact	0.4327



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