

## Round 5 results — Run UPrrf102-r5 submitted from unique\_ptr

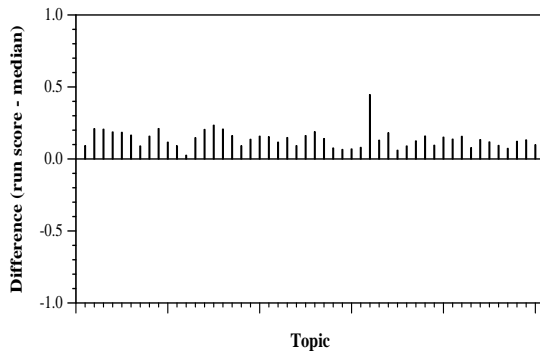
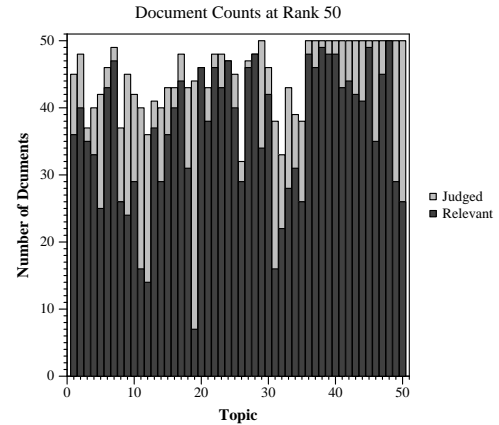
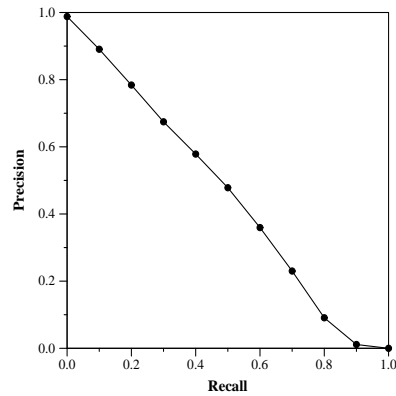
### Run Description

A reciprocal rank fusion of (a) Anserini & Terrier runs based on different indexing schemes and topic variants, including relevance feedback (b) Neural retrieval runs based on synthetic query generation (<https://arxiv.org/abs/2004.14503>) (c) TF-Ranking + multiple BERT variants with softmax loss (<https://arxiv.org/abs/2004.08476>) trained on MS-Marco (d) TF-Ranking + multiple BERT variants and topic variants fine-tuned with softmax loss on the relevance judgments from Rounds 1 - 4.

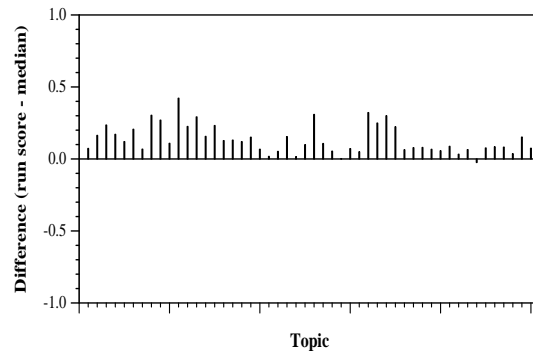
Summary Statistics	
Run ID	UPrrf102-r5
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	50
Total number retrieved	49889
Total relevant	10910
Total relevant retrieved	8153
MAP	0.4569
Mean Bpref	0.6043
Mean NDCG@20	0.8092
Mean RBP(p=0.5)	0.9056 +0.0000

Document Level Averages	
	Precision
At 5 docs	0.9600
At 10 docs	0.9140
At 15 docs	0.8773
At 20 docs	0.8530
At 30 docs	0.8073
R-Precision	
Exact	0.4805



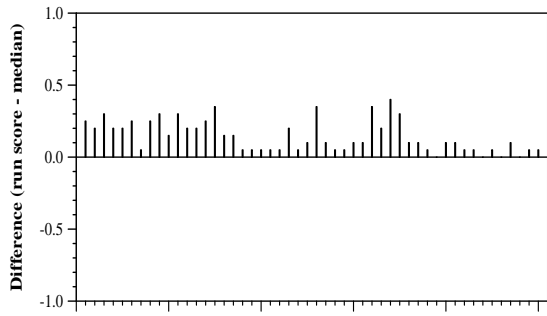
Per-topic difference from median bpref for all Round 4 runs



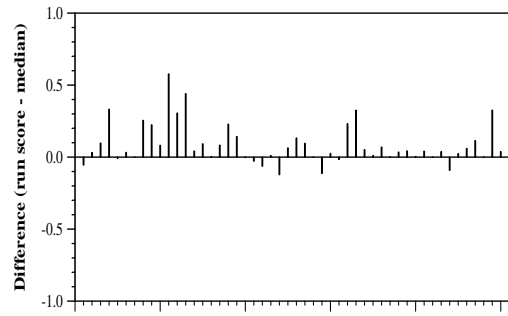
Per-topic difference from median NDCG@20 for all Round 4 runs

Round 5 results — Run UPrrf102-r5 submitted from unique\_ptr

---



Topic  
Per-topic difference from median P@20 for all Round 4 runs



Topic  
Per-topic difference from median RBP(p=0.5) for all Round 4 runs