

Round 4 results — Run ILPS_UvA_zeroshot_BE submitted from ILPS_UvA

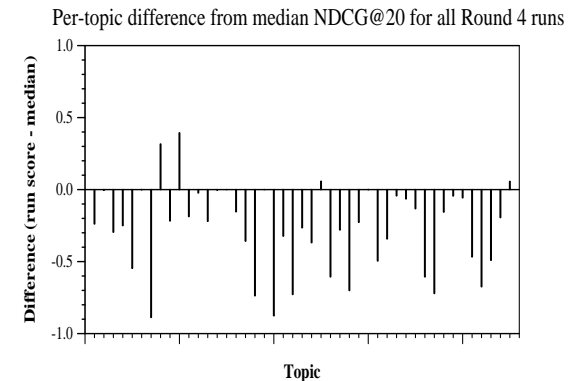
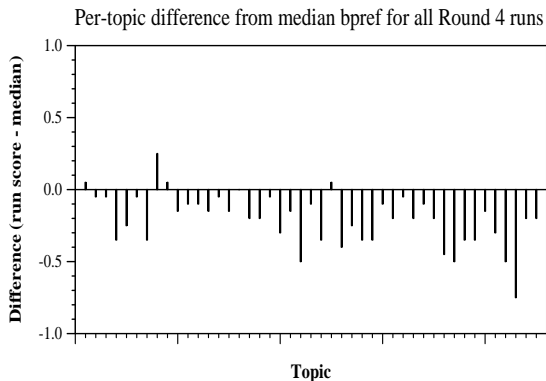
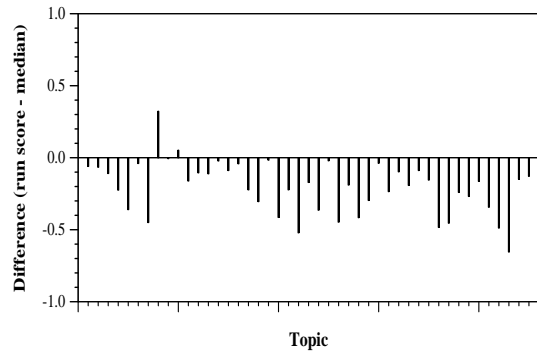
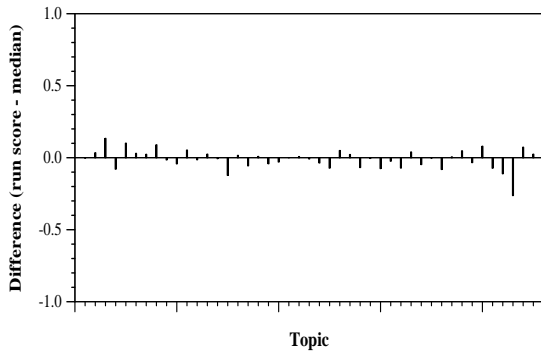
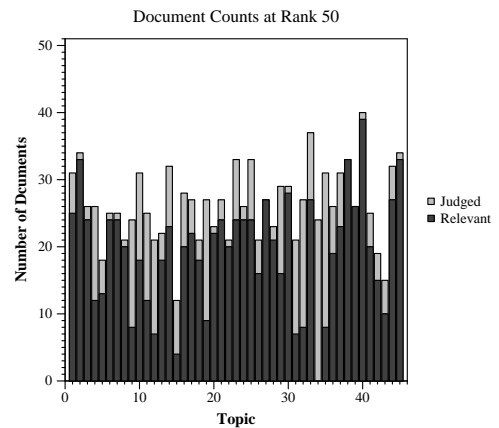
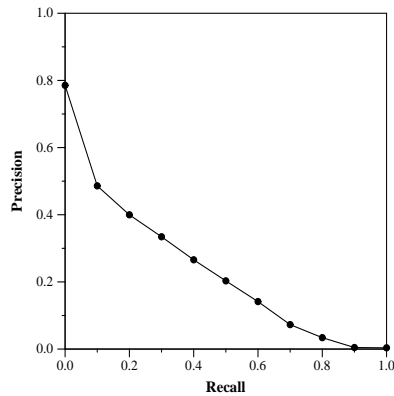
Run Description

Labels: zeroshot weak (with sample size = 5, negcutoff=1000, GM) + BERT re-ranker

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

Overall measures	
Number of topics	45
Total number retrieved	36320
Total relevant	5824
Total relevant retrieved	3628
MAP	0.2206
Mean Bpref	0.5059
Mean NDCG@20	0.4222
Mean RBP(p=0.5)	0.5001 +0.3450

Document Level Averages	
	Precision
At 5 docs	0.5422
At 10 docs	0.4844
At 15 docs	0.4681
At 20 docs	0.4456
At 30 docs	0.4244
R-Precision	
Exact	0.3013



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