Run Description

We use the reciprocal rank fusion method to combine the rankings induced by Run2, Run3 with the ranking induced by the BE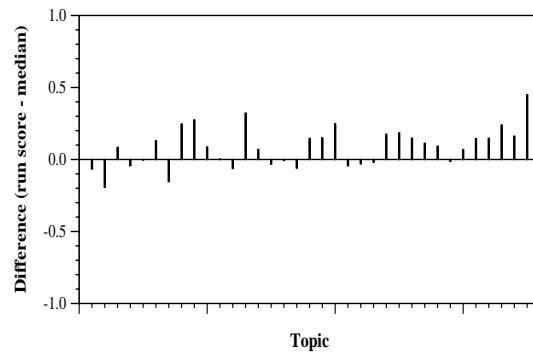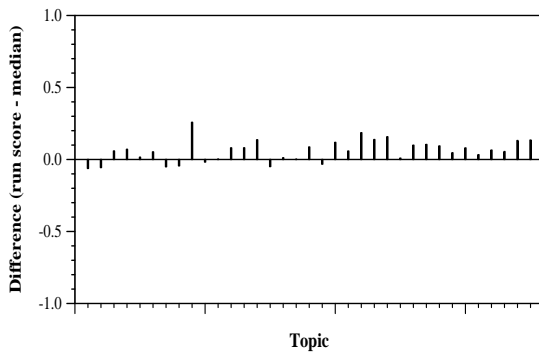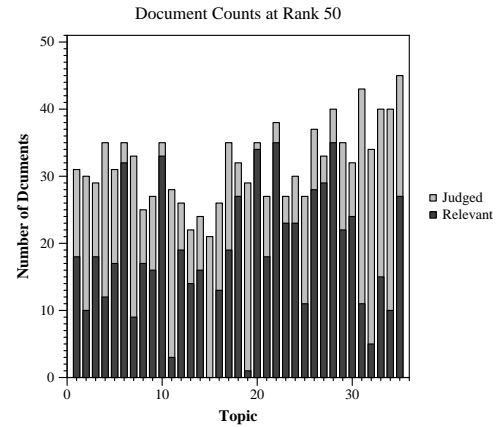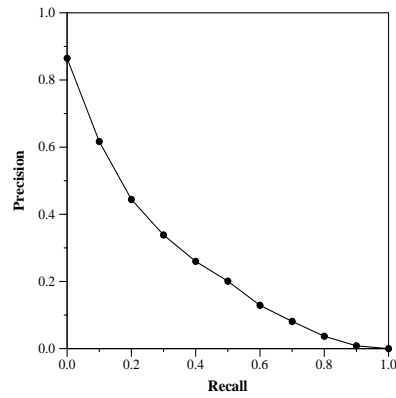RT-base (12 layers, 768 hidden size) model which was fine-tuned on the Ms Marco passage dataset. The model's free parameter was set to 60.

Summary Statistics	
Run ID	Technion-RRF
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	35
Total number retrieved	35000
Total relevant	3002
Total relevant retrieved	1883
MAP	0.2465
Mean Bpref	0.4290
Mean NDCG@10	0.5566
Mean RBP(p=0.5)	0.6236 +0.0016

Document Level Averages	
	Precision
At 5 docs	0.6800
At 10 docs	0.6000
At 15 docs	0.5562
At 20 docs	0.5071
At 30 docs	0.4543
R-Precision	
Exact	0.2995

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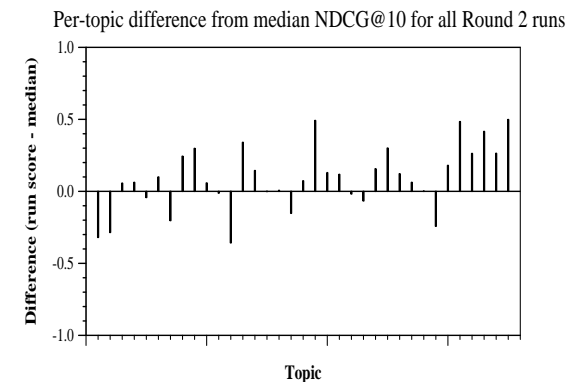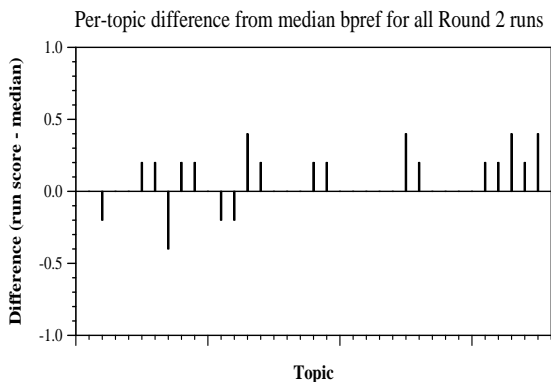
Per-topic difference from median P@5 for all Round 2 runs

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