

Round 2 results — Run bm25_safir_rrf submitted from ims_unipd

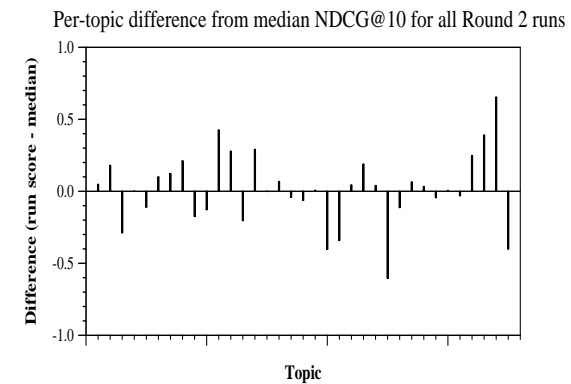
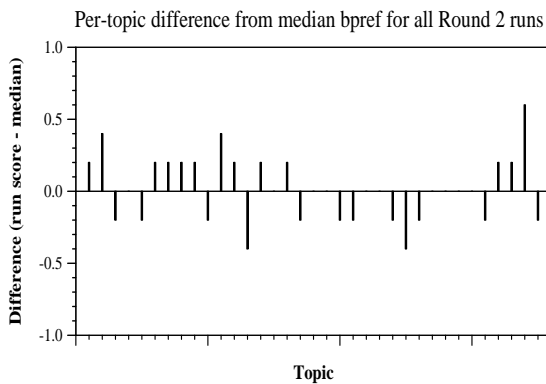
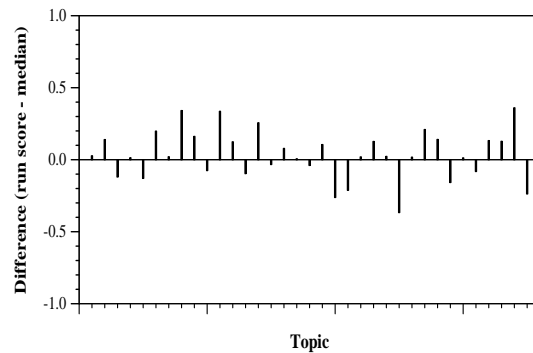
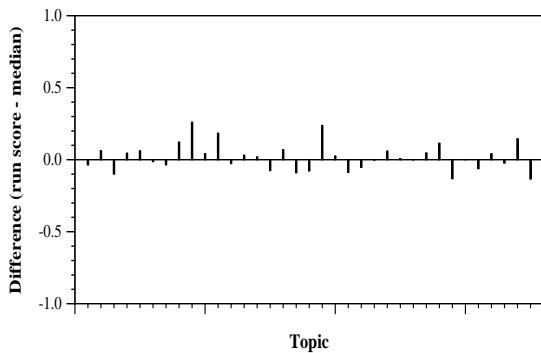
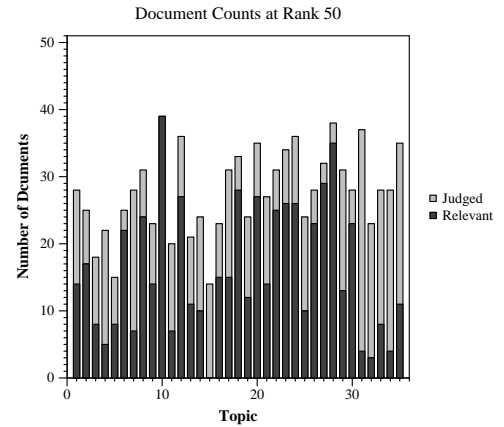
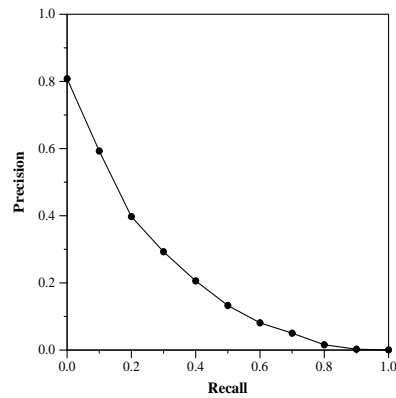
Run Description

We adopt SAFIR (three models) + RM3 and BM25 + RM3. Then we performed RRF to combine the four runs.

Summary Statistics	
Run ID	bm25_safir_rrf
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	35
Total number retrieved	32032
Total relevant	3002
Total relevant retrieved	1612
MAP	0.2090
Mean Bpref	0.3883
Mean NDCG@10	0.5053
Mean RBP(p=0.5)	0.5457 +0.0057

Document Level Averages	
	Precision
At 5 docs	0.6171
At 10 docs	0.5714
At 15 docs	0.4952
At 20 docs	0.4686
At 30 docs	0.3971
R-Precision	
Exact	0.2732



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