

Round 1 results — Run BERT submitted from THUMSR

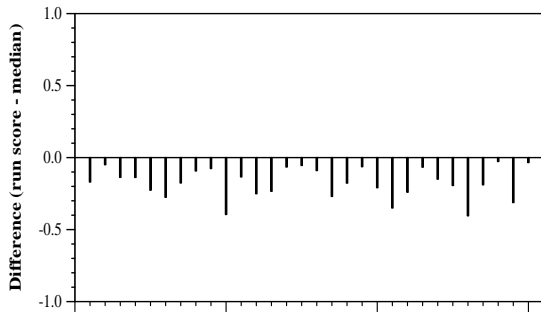
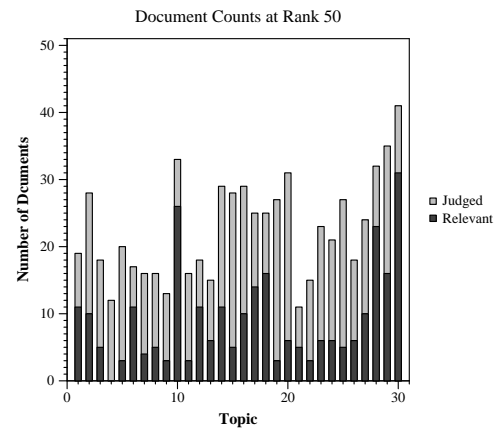
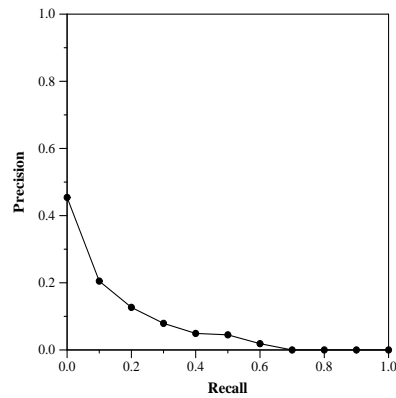
Run Description

Combination of BM25 score and BERT score with Reinforce training.

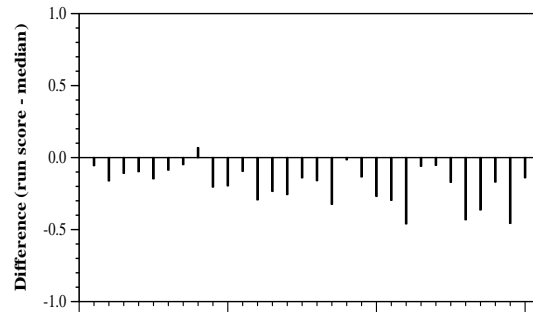
Summary Statistics	
Run ID	BERT
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	30
Total number retrieved	2355
Total relevant	2352
Total relevant retrieved	359
MAP	0.0653
Mean Bpref	0.1506
Mean NDCG@10	0.1731

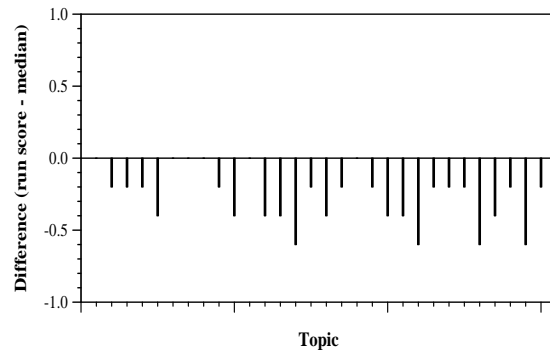
Document Level Averages	
	Precision
At 5 docs	0.2200
At 10 docs	0.2333
At 15 docs	0.2244
At 20 docs	0.2117
At 30 docs	0.2111
R-Precision	
Exact	0.1525



Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs



Per-topic difference from median P@5 for all Round 1 runs