

Round 3 results — Run ruir-q2narr4task4 submitted from RUIR

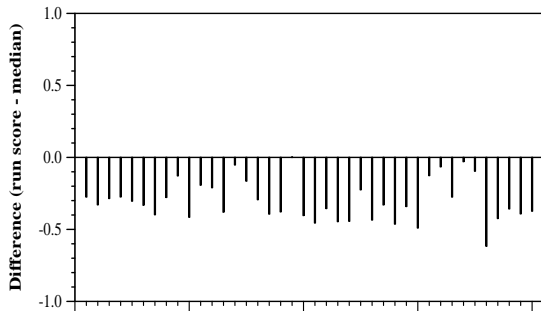
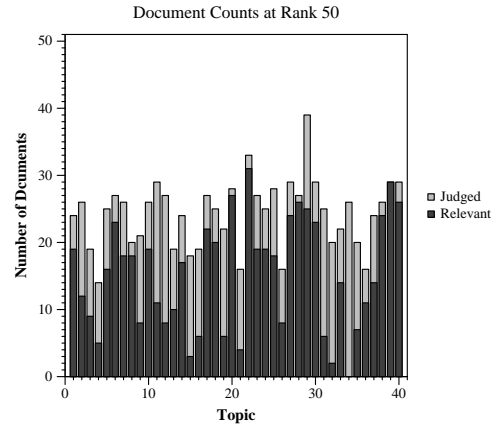
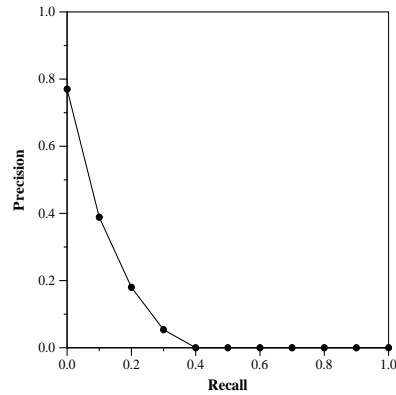
Run Description

manual classification of trec topics into the original 10 kaggle tasks. rm3 query expansion where we weight terms as 0.2 \* query terms + 0.4 \* narrative terms + 0.4 task terms task terms found by tf-idf from kaggle task description against corpus of paper abstracts. parameters found by coarsely tuning towards ndcg after filtering docs with unknown qrels (may be more stable than bpref - sakai 2007)

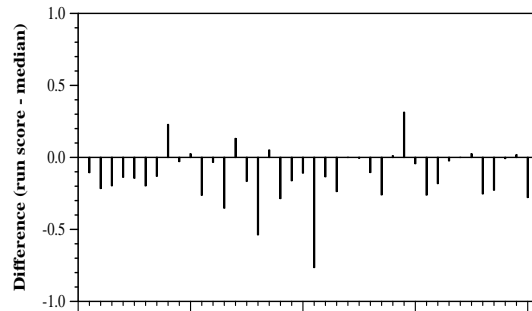
Summary Statistics	
Run ID	ruir-q2narr4task4
Topic type	manual
Contributed to judgment sets?	yes

Overall measures	
Number of topics	40
Total number retrieved	3131
Total relevant	4698
Total relevant retrieved	785
MAP	0.0925
Mean Bpref	0.1603
Mean NDCG@10	0.4672
Mean RBP(p=0.5)	0.5038 +0.0186

Document Level Averages	
	Precision
At 5 docs	0.5850
At 10 docs	0.5525
At 15 docs	0.4550
At 20 docs	0.4275
At 30 docs	0.3600
R-Precision	
Exact	0.1674



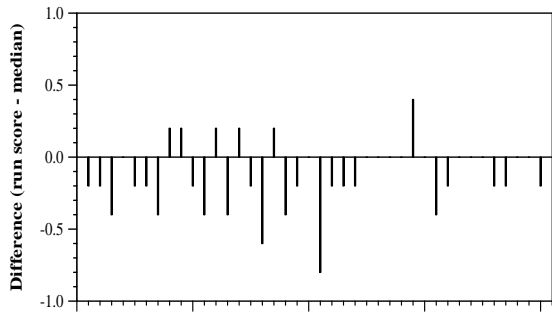
Per-topic difference from median bpref for all Round 3 runs



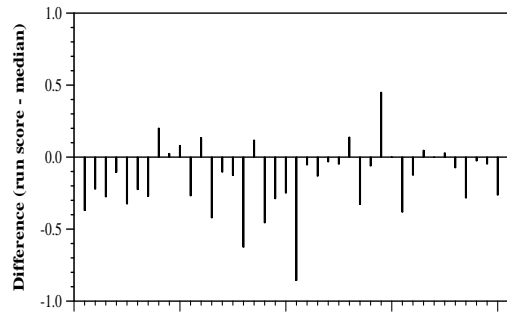
Per-topic difference from median NDCG@10 for all Round 3 runs

Round 3 results — Run ruir-q2narr4task4 submitted from RUIR

---



Per-topic difference from median P@5 for all Round 3 runs



Per-topic difference from median RBP(p=0.5) for all Round 3 runs