

Round 2 results — Run PS-r2-termebd submitted from PITT

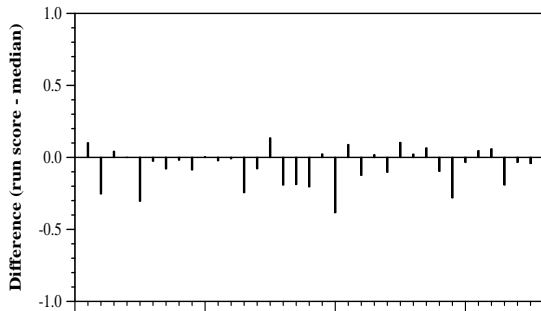
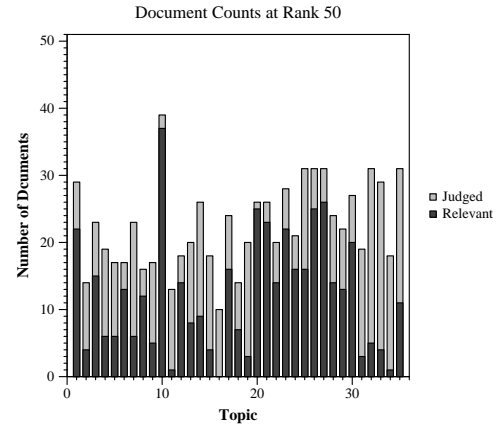
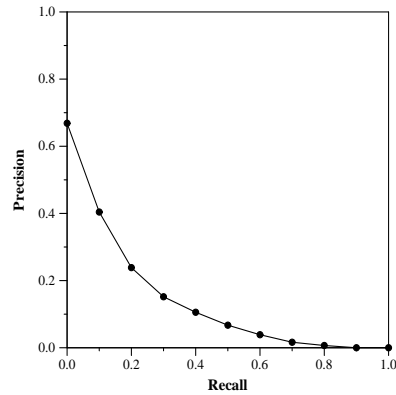
Run Description

The query was generated manually through human by reading the topic, question and narrative section of the topics file and find keywords from it. The query is then expand through the Metamap + Term level Word embedding. The expanded query is then sent to Elasticsearch engine which use tf-idf model to index and use "BM25" retrieval model to rank the result. Currently, all the words weigh the same inside the query.

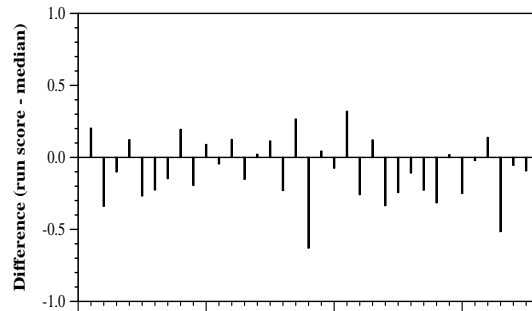
Summary Statistics	
Run ID	PS-r2-termebd
Topic type	manual
Contributed to judgment sets?	no

Overall measures	
Number of topics	35
Total number retrieved	27695
Total relevant	3002
Total relevant retrieved	1303
MAP	0.1287
Mean Bpref	0.3049
Mean NDCG@10	0.3850
Mean RBP(p=0.5)	0.4503 +0.0075

Document Level Averages	
	Precision
At 5 docs	0.5029
At 10 docs	0.4314
At 15 docs	0.3829
At 20 docs	0.3514
At 30 docs	0.3067
R-Precision	
Exact	0.1841

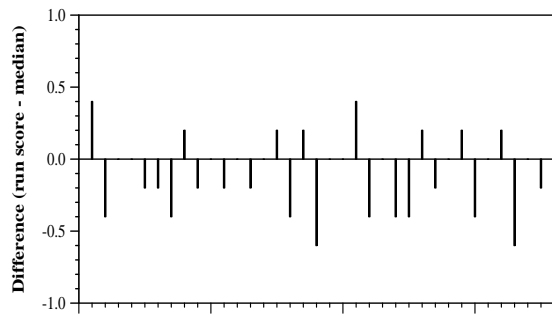


Per-topic difference from median bpref for all Round 2 runs

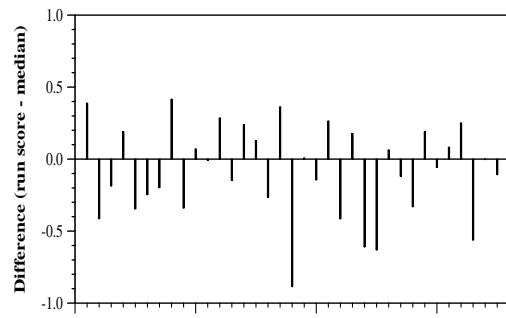


Per-topic difference from median NDCG@10 for all Round 2 runs

Round 2 results — Run PS-r2-termebd submitted from PITT



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