

Round 4 results — Run udel\_fang\_lambdarank submitted from udel\_fang

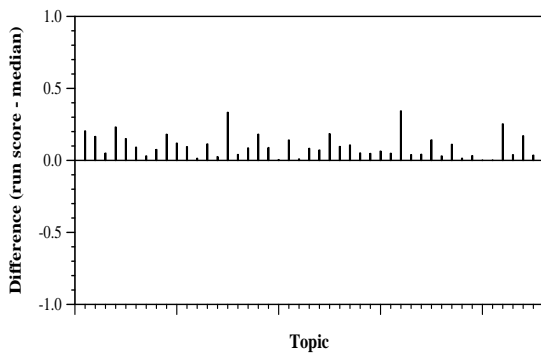
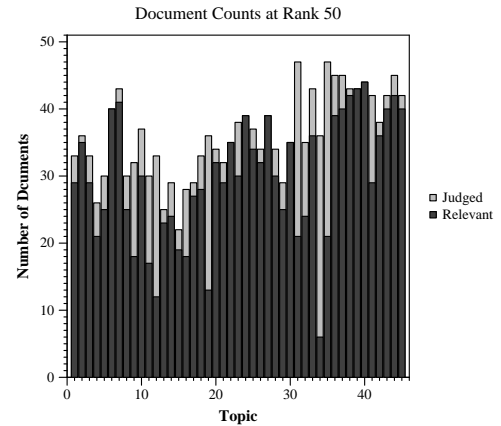
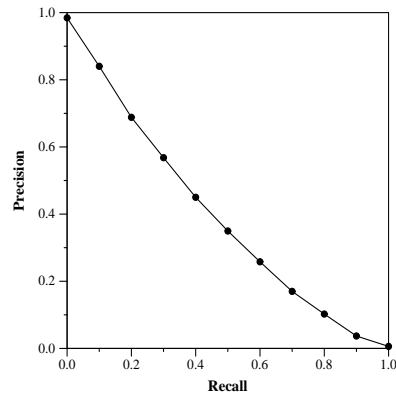
Run Description

We build an index with title and abstract from the metadata file. Non-stopwords in the query and entities tagged by SciSpacy in question and narrative fields are assigned the weight ratio of 2:3:1 to form the query. We generate a run using relevance feedback on the first 40 queries and pseudo relevance feedback on the last 5 queries. LambdaMART is used to re-rank the first 200 results. The features we use include BM25, SciBERT, recency, and so on.

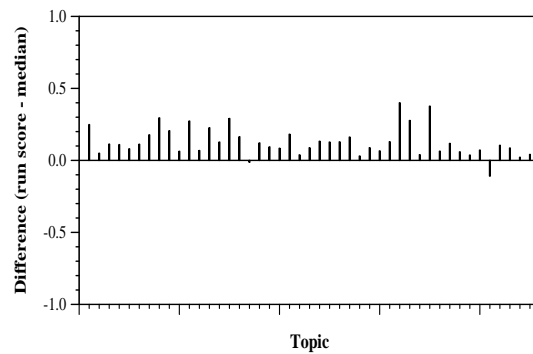
Summary Statistics	
Run ID	udel_fang_lambdarank
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	45
Total number retrieved	44979
Total relevant	5824
Total relevant retrieved	4198
MAP	0.3907
Mean Bpref	0.6161
Mean NDCG@20	0.7534
Mean RBP(p=0.5)	0.8660 +0.0004

Document Level Averages	
	Precision
At 5 docs	0.9378
At 10 docs	0.8911
At 15 docs	0.8281
At 20 docs	0.7844
At 30 docs	0.7133
R-Precision	
Exact	0.4167



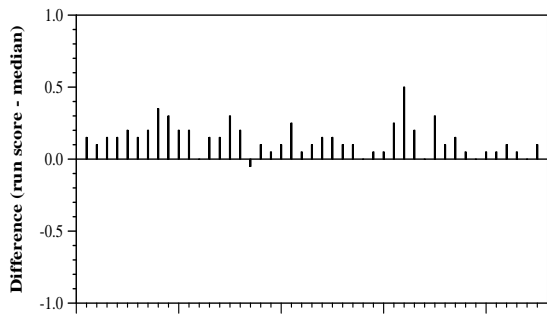
Per-topic difference from median bpref for all Round 4 runs



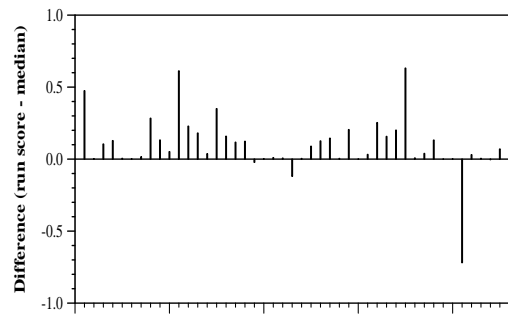
Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run udel\_fang\_lambdarank submitted from udel\_fang

---



Topic  
Per-topic difference from median P@20 for all Round 4 runs



Topic  
Per-topic difference from median RBP(p=0.5) for all Round 4 runs