

Round 5 results — Run udel_fang_ltr_split submitted from udel_fang

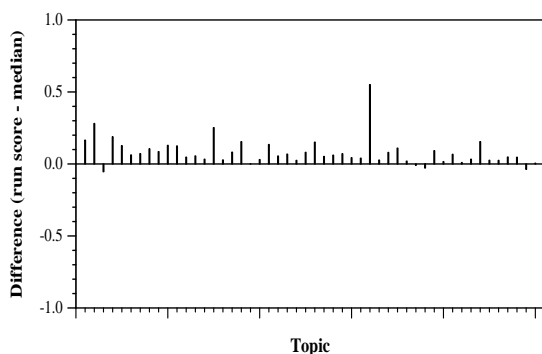
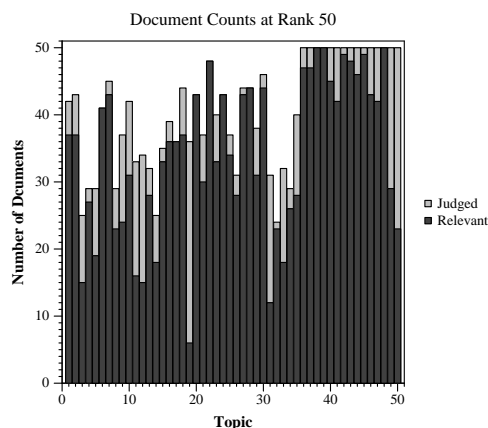
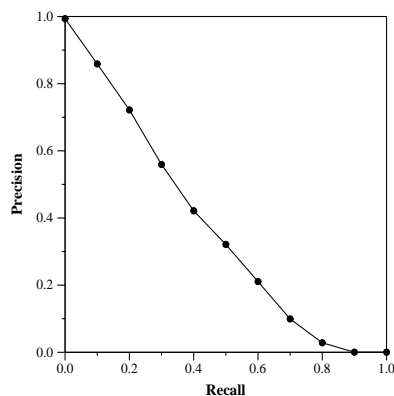
Run Description

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 45 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. We tune the hyper-parameters of relevance feedback and pseudo relevance feedback based on different validation methods.

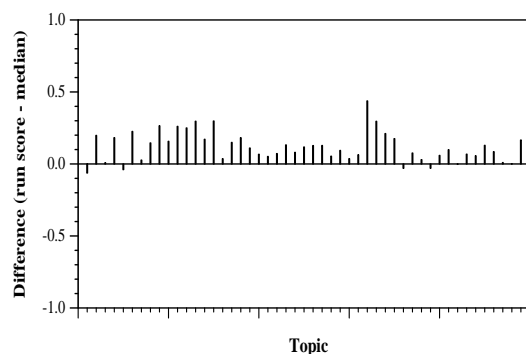
Summary Statistics	
Run ID	udel_fang_ltr_split
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	50
Total number retrieved	49894
Total relevant	10910
Total relevant retrieved	7041
MAP	0.3682
Mean Bpref	0.5451
Mean NDCG@20	0.7929
Mean RBP(p=0.5)	0.8988 +0.0005

Document Level Averages	
	Precision
At 5 docs	0.9640
At 10 docs	0.9060
At 15 docs	0.8693
At 20 docs	0.8270
At 30 docs	0.7853
R-Precision	
Exact	0.4007

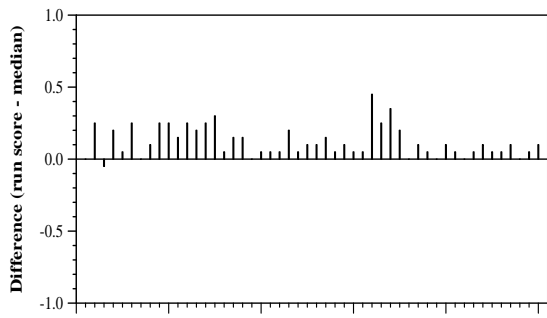


Per-topic difference from median bpref for all Round 4 runs

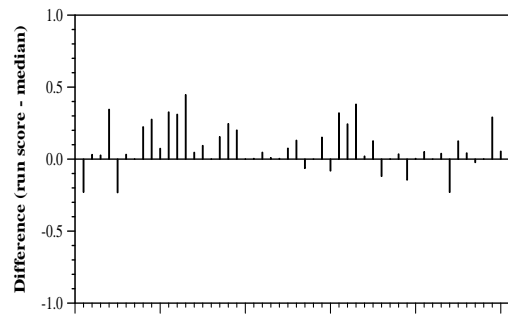


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 5 results — Run udel_fang_itr_split 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