

Round 4 results — Run SFDC-fus12-enc34 submitted from SFDC

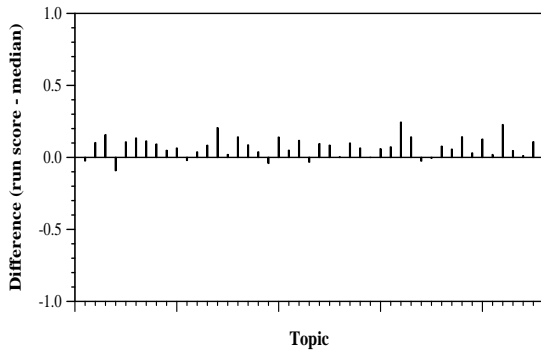
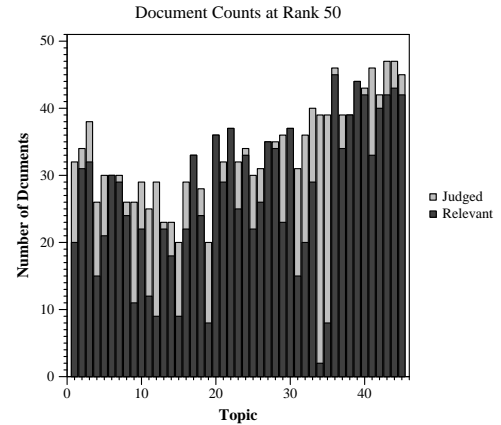
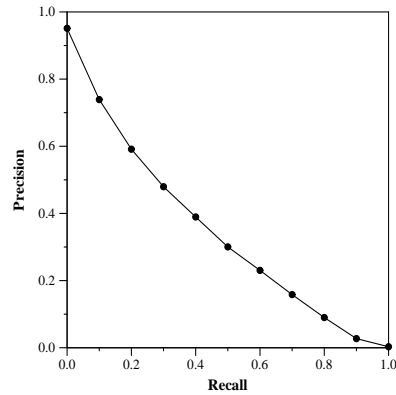
Run Description

Implements: RRF(EncoderRerank(RRF(fusion1, fusion2)), encoder3, encoder4) fusion1 is anserini baseline 135 fusion2 is anserini baseline 246 encoders are hybrid semantic (bipartite-graph-trained SBERT) + keyword-based (TFIDF) encoder3 means TREC-3-data trained encoder, indexing round 4 data encoder4 means TREC-4-data trained encoder, indexing round 4 data

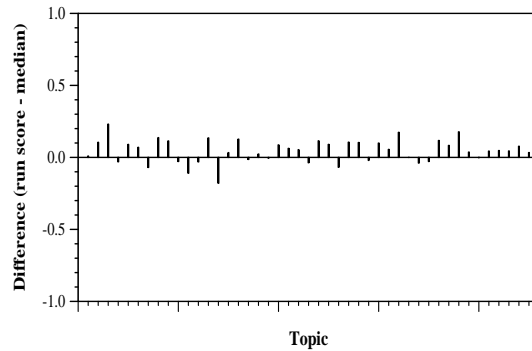
Summary Statistics	
Run ID	SFDC-fus12-enc34
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	45
Total number retrieved	44979
Total relevant	5824
Total relevant retrieved	4175
MAP	0.3436
Mean Bpref	0.5887
Mean NDCG@20	0.6726
Mean RBP(p=0.5)	0.7697 +0.0007

Document Level Averages	
	Precision
At 5 docs	0.8222
At 10 docs	0.7800
At 15 docs	0.7363
At 20 docs	0.7011
At 30 docs	0.6304
R-Precision	
Exact	0.3759

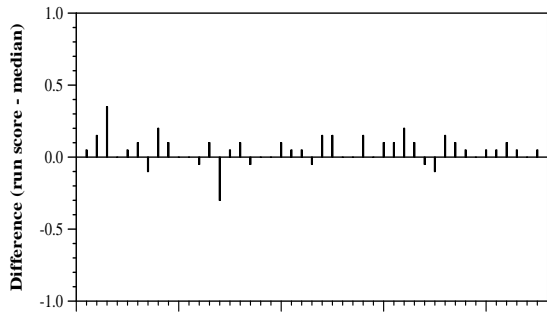


Per-topic difference from median bpref for all Round 4 runs

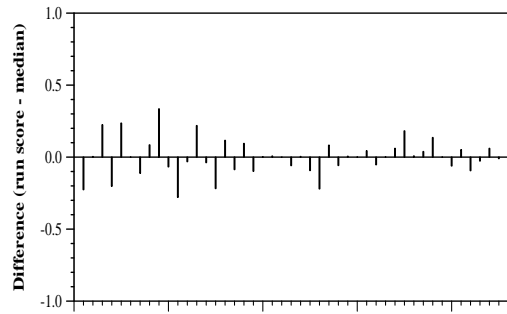


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run SFDC-fus12-enc34 submitted from SFDC



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



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