

Round 4 results — Run `poznan_p4_run3` submitted from POZNAN

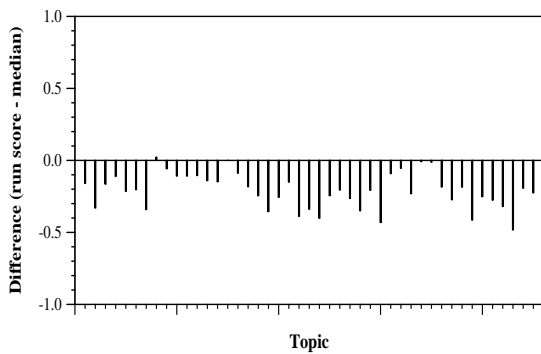
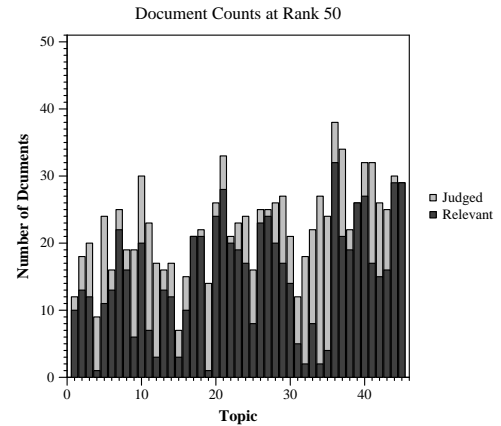
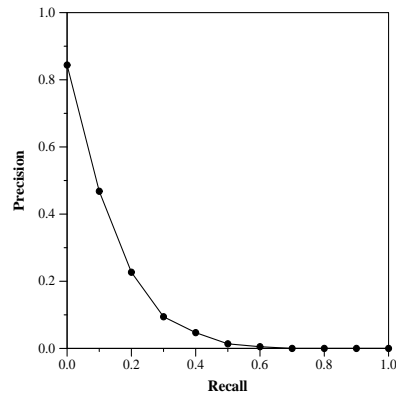
Run Description

Aggressive reranking of BM25 baseline. Abstract is tokenized into sentences. Each word in a short query gets compared with each word in the document abstract. Five sentences with the highest sum of top five most similar words to the query are picked. Sum of the top five word similarities in top five sentences is added to the score. Word similarity is defined as a cosine similarity between word vectors (embedding calculated with `word2vec`).

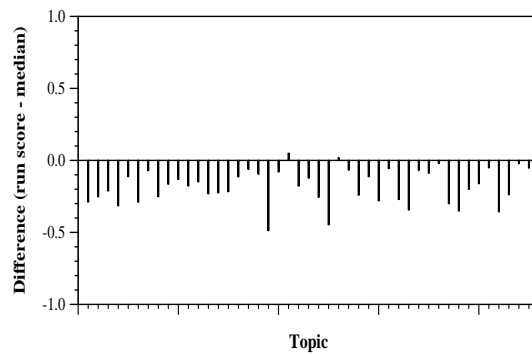
Summary Statistics	
Run ID	<code>poznan_p4_run3</code>
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	45
Total number retrieved	31946
Total relevant	5824
Total relevant retrieved	2152
MAP	0.1187
Mean Bpref	0.3066
Mean NDCG@20	0.4468
Mean RBP( $p=0.5$ )	0.5843 +0.1332

Document Level Averages	
	Precision
At 5 docs	0.6356
At 10 docs	0.5822
At 15 docs	0.5126
At 20 docs	0.4589
At 30 docs	0.3881
R-Precision	
Exact	0.1815



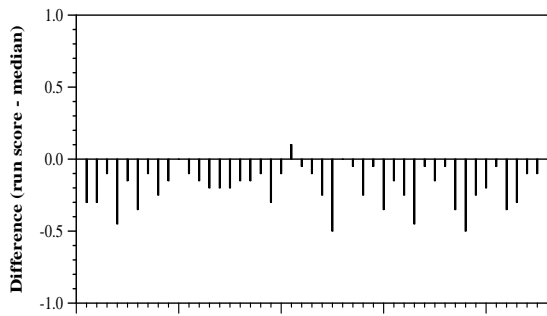
Per-topic difference from median bpref for all Round 4 runs



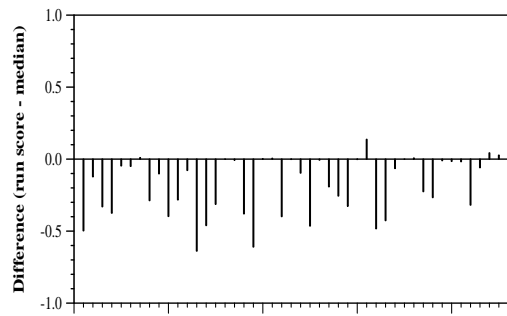
Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run poznan\_p4\_run3 submitted from POZNAN

---



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