Round 1 results — Run PL2c1.0 submitted from UB_BW

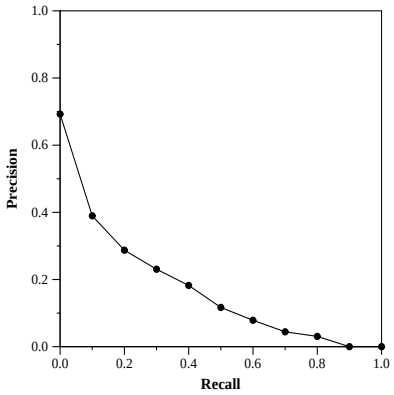
Run Description

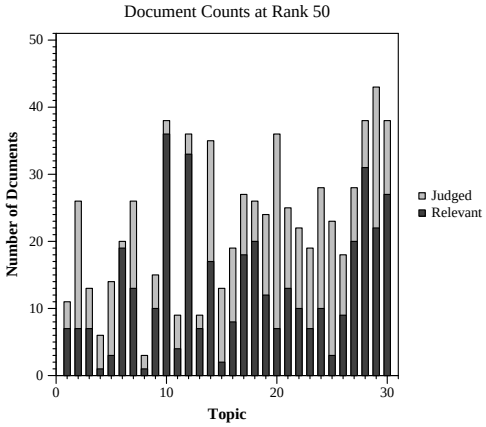
For this run, we used Terrier-v5.2, an open source Information Retrieval (IR) platform. All the documents (title and abstracts) used in this study were first pre-processed before indexing and this involved tokenising the text and stemming each token using the full Porter stemming algorithm. Block indexing was enabled so that we could deploy proximity search. Stopword removal was enabled and we used Terrier-v5.2 stopword list. We used PL2 Divergence from Randomness term weighting model in Terrier-v5.2 IR platform to score and rank the documents. The hyper-parameter for PL2 was set to its default value of b = 1.0. During retrieval we used only the query in the topic. As improvement, we used Markov Random Fields for Term Dependencies. We used the full dependence variant of the model, which models dependencies between adjacent query terms. In this work, we explore a window size of 4, to see what impact it has of the retrieval effectiveness.

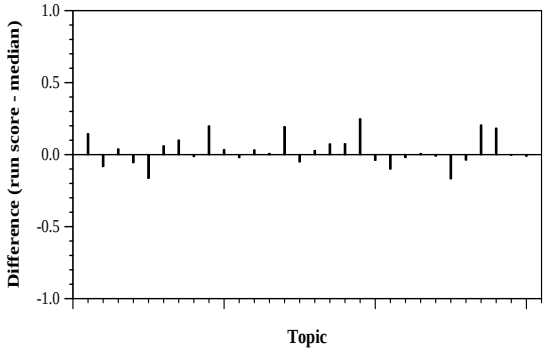
Summary Statistics	
Run ID	PL2c1.0
Topic type	automatic
Contributed to judgment sets?	no

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	1109
MAP	0.1617
Mean Bpref	0.3544
Mean NDCG@10	0.3536

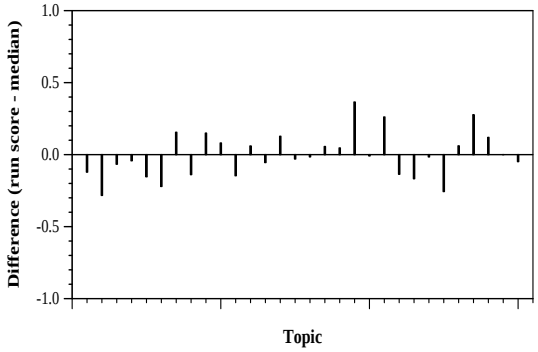
Document Level Averages	
	Precision
At 5 docs	0.4200
At 10 docs	0.4100
At 15 docs	0.3733
At 20 docs	0.3533
At 30 docs	0.3100
R-Precision	
Exact	0.2212





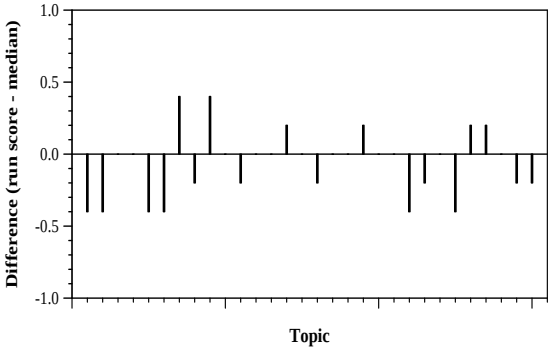


Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs

Round 1 results — Run PL2c1.0 submitted from UB_BW



Per-topic difference from median P@5 for all Round 1 runs