

Round 4 results — Run Marouane_eQQ_EnM-BTN submitted from Marouane

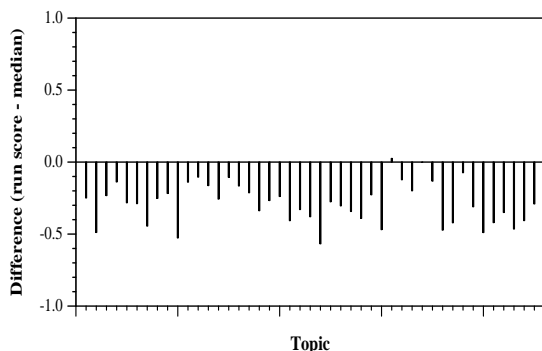
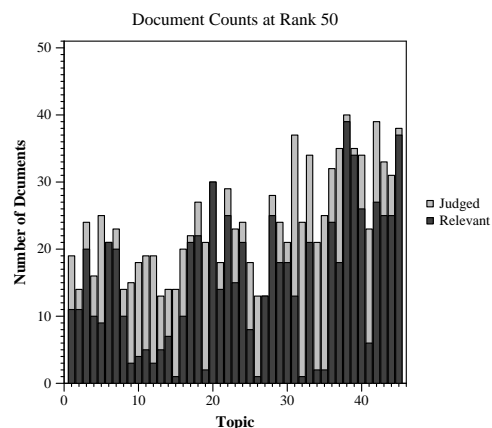
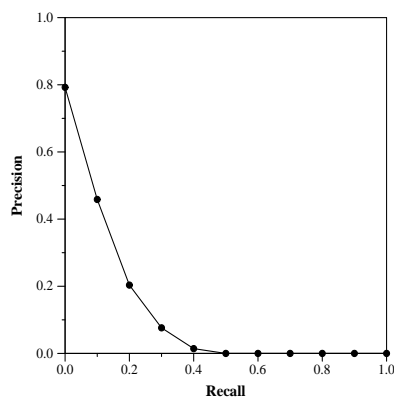
Run Description

In this run we rank the documents using Ensemble Model (EnM), $EnM = 0.38*BM25+0.34*TFIDF+0.28*NMF$. We train/rank on a small subset (#Papers 18674, #Tokens 24909), papers published after 2019-12, containing PMC parsed file Please see the [1](7.2 Coherence) to see why we are using NMF. NMF stands for Non-negative Matrix Factorization. We expand the Query+Question using "Topic models RM" please see [1](Section 3.3). [1] <https://www.kaggle.com/atmarouane/covid-19-search-engine-indexing-by-lda-enm> [2] Indexing by Latent Dirichlet Allocation and Ensemble Model

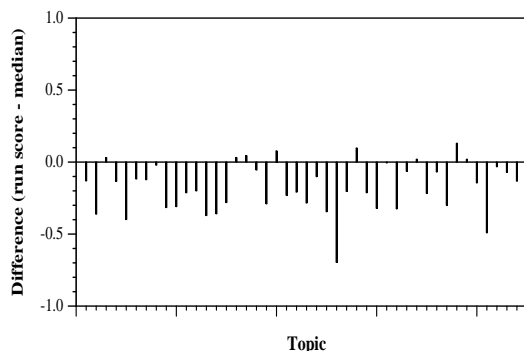
Summary Statistics	
Run ID	Marouane_eQQ_EnM-BTN
Topic type	feedback
Contributed to judgment sets?	yes

Overall measures	
Number of topics	45
Total number retrieved	44996
Total relevant	5824
Total relevant retrieved	1639
MAP	0.1111
Mean Bpref	0.2305
Mean NDCG@20	0.4558
Mean RBP(p=0.5)	0.6079 +0.0015

Document Level Averages	
	Precision
At 5 docs	0.6622
At 10 docs	0.5756
At 15 docs	0.5304
At 20 docs	0.4744
At 30 docs	0.3970
R-Precision	
Exact	0.1699

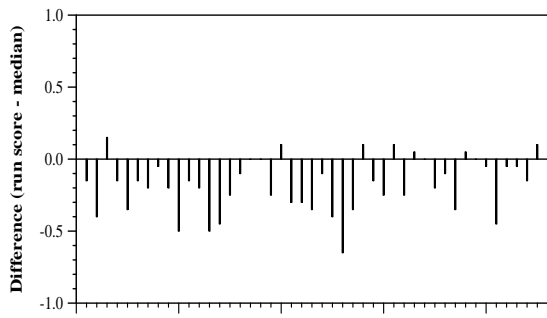


Per-topic difference from median bpref for all Round 4 runs

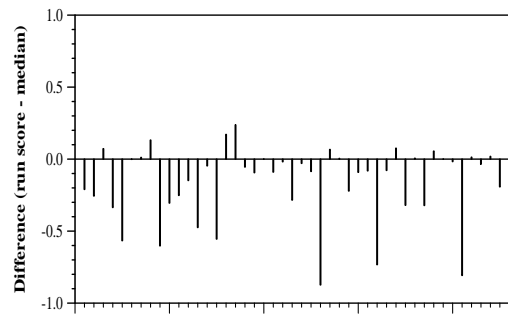


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run Marouane_eQQ_EnM-BTN submitted from Marouane



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