Round 2 results — Run elhuyar_rRnk_cbert1 submitted from Elhuyar_NLP_team

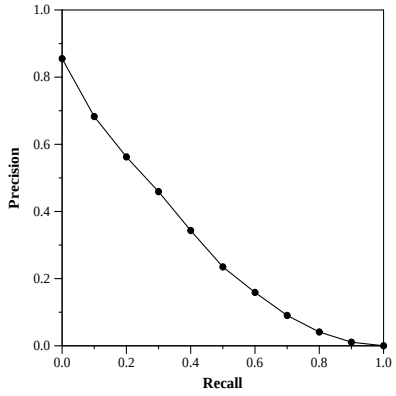
Run Description

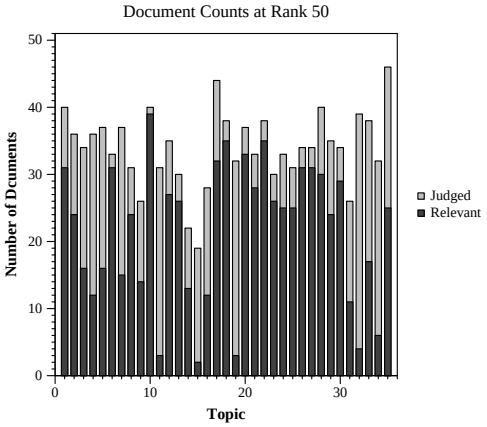
We tackle this document retrieval task in two steps: a) a first ranking and b) re-ranking. In order to obtain the first ranking of relevant documents of the collection corresponding to the queries, we use a language modeling based information retrieval approach (Ponte & Croft, 1998) including pseudo relevance feedback. For that purpose, we used the Indri search engine (Strohman, 2005), which combines Bayesian networks with language models. Then, we make a re-ranking based on BERT following a strategy similar to the one proposed by Nogueira and Cho (2019). As we do not have a collection of query pairs and relevant abstracts for tuning BERT for this passage retrieval task, we simulate a training collection composed of titles and their corresponding abstracts from the COVID-19 Open Research dataset. Through this training collection we tuned the Clinical BERT model (Alsentzer et al., 2019) to the task of identifying relevant queries and abstracts. Indri and Tuned Clinical BERT scores are linearly combined and re-ranking is performed according to that new score. In this run we use weights of 0.9 and 0.1 for Indri and Clinical BERT scores respectively.

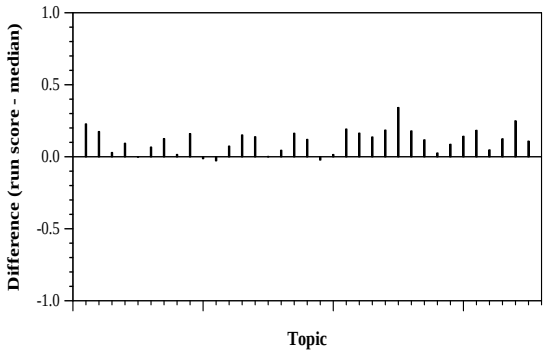
Summary Statistics	
Run ID	elhuyar_rRnk_cbert1
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	35
Total number retrieved	31888
Total relevant	3002
Total relevant retrieved	1962
MAP	0.2941
Mean Bpref	0.4793
Mean NDCG@10	0.5912
Mean RBP(p=0.5)	0.6313 +0.0011

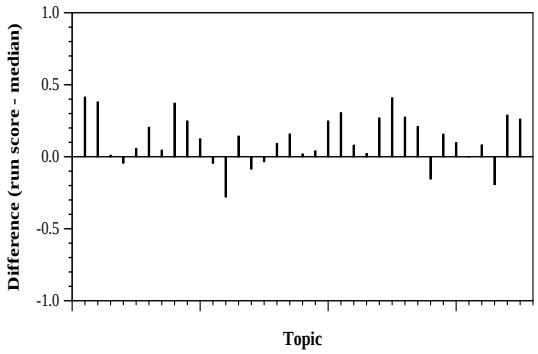
Document Level Averages	
	Precision
At 5 docs	0.7200
At 10 docs	0.6400
At 15 docs	0.6095
At 20 docs	0.5857
At 30 docs	0.5257
R-Precision	
Exact	0.3429





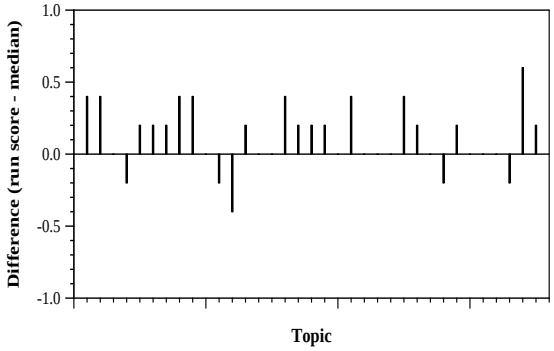


Per-topic difference from median bpref for all Round 2 runs

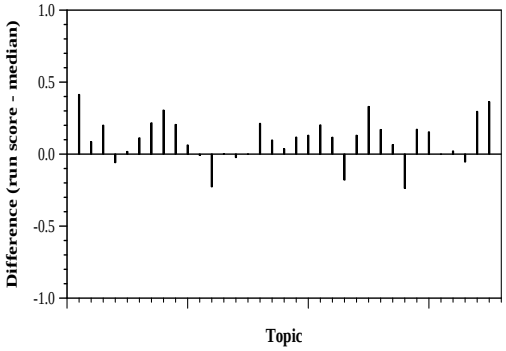


Per-topic difference from median NDCG@10 for all Round 2 runs

Round 2 results — Run elhuyar_rRnk_cbert1 submitted from Elhuyar_NLP_team



Per-topic difference from median P@5 for all Round 2 runs



Per-topic difference from median RBP(p=0.5) for all Round 2 runs