Round 1 results — Run BioinfoUA-emb submitted from BioinformaticsUA

Run Description

This run corresponds to the results of a system that was tunned for the BioASQ challenge (a more broad biomedical Adhoc retrieval challenge). So this submission tries to explore the possible similarity between the data domains in order to train a neural ranking model. The system uses a standard BM25 + Neural ranking model. In the retrieval were considered only documents that have title+abstract to be more similar to the BioASQ data. The neural ranking is built upon the DeepRank model and a more complete description can be found here [1]. The word embeddings were computed on CORD+Pubmed corpus using word2vec. For each topic, the field "question" was used to express the information need. REFs: [1] T. Almeida and S. Matos, 'Calling Attention to Passages for Biomedical Question Answering,' in Advances in Information Retrieval, 2020, pp. 69–77.

Summary Statistics	
Run ID	BioinfoUA-emb
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	1227
MAP	0.1903
Mean Bpref	0.3530
Mean NDCG@10	0.5298

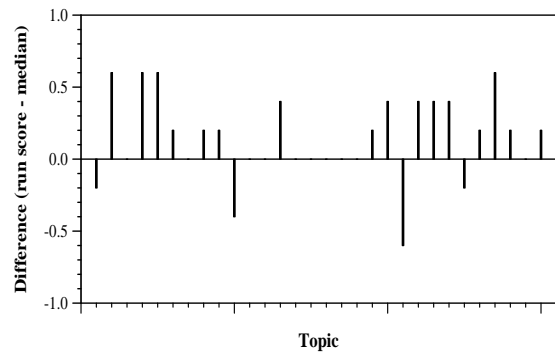| Document Level Averages	
	Precision
At 5 docs	0.6333
At 10 docs	0.5533
At 15 docs	0.4756
At 20 docs	0.4283
At 30 docs	0.3700
R-Precision	
Exact	0.2367

Document Counts at Rank 50
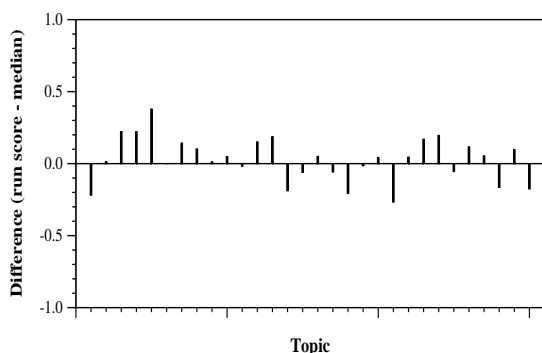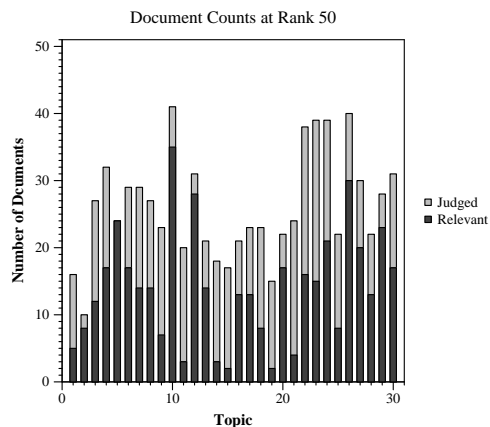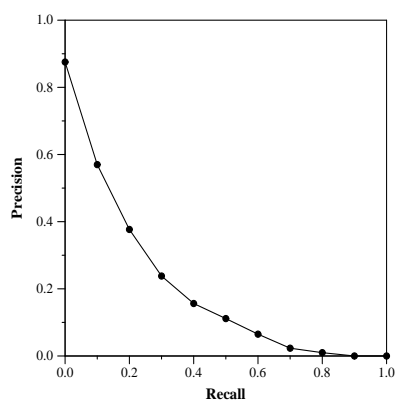
Per-topic difference from median bpref for all Round 1 runs
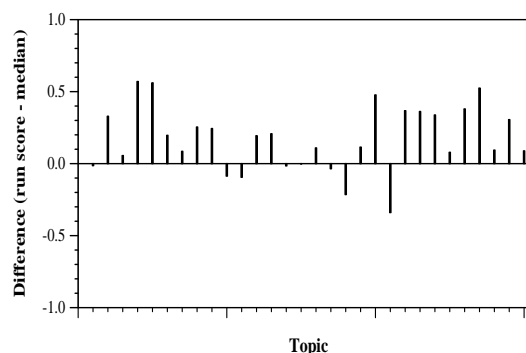
Per-topic difference from median NDCG@10 for all Round 1 runs

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