

Round 4 results — Run jlQErnd4 submitted from julielab

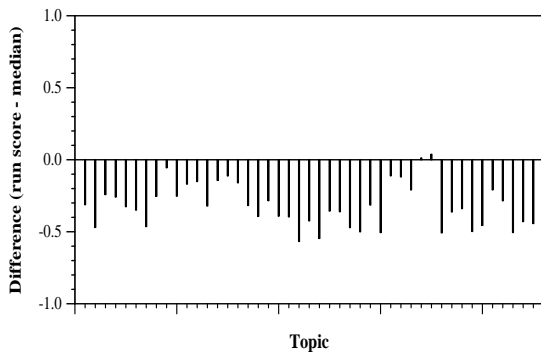
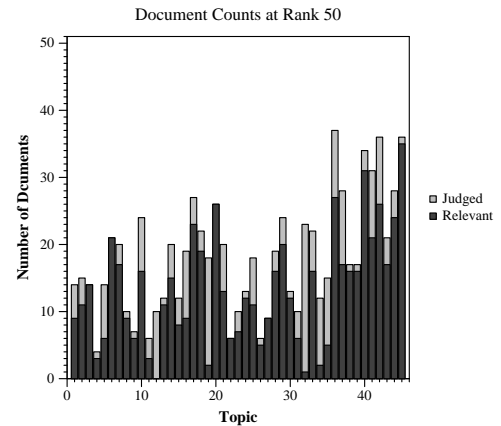
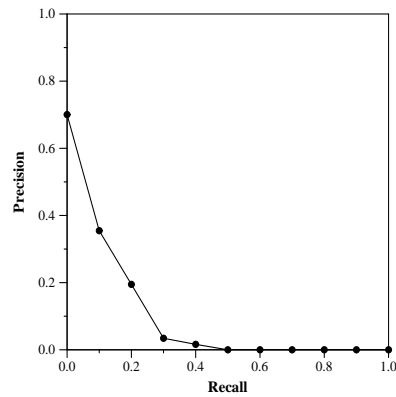
Run Description

In most queries the token "coronavirus" is present. However, coronaviridae are a family of viruses, that are not limited to only SARS-CoV-2. Thus many false positive are likely to be found. This holds also true for other terms, such as "animal model". This term does not occur often, as most of the researchers specify the animal they used as model organism (such as mice or rats). So we created a list of synonyms to specify these general terms. We found out, that mostly nouns are the terms that contained the most information. Thus we used a part-of-speech tagger to isolate nouns and used a manually filled blacklist as well as a general stop word list to filter them. Our final query consists of four different parts: several synonyms and spellings for the illness and the virus respectively, the query, the question and finally a bag-of-words containing the filtered nouns from both the question and the query. We planned on including the filtered nouns of the narrative as optional part of the query, but it turned out during evaluation on the previous part, that it does not have any benefit on the result.

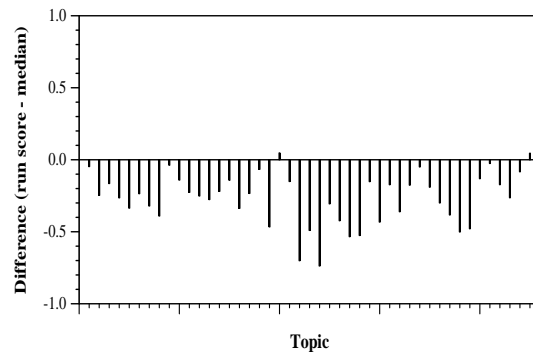
Summary Statistics	
Run ID	jlQErnd4
Topic type	manual
Contributed to judgment sets?	no

Overall measures	
Number of topics	45
Total number retrieved	8789
Total relevant	5824
Total relevant retrieved	1158
MAP	0.0868
Mean Bpref	0.1999
Mean NDCG@20	0.3601
Mean RBP(p=0.5)	0.4518 +0.3549

Document Level Averages	
	Precision
At 5 docs	0.4800
At 10 docs	0.4178
At 15 docs	0.3896
At 20 docs	0.3722
At 30 docs	0.3356
R-Precision	
Exact	0.1741

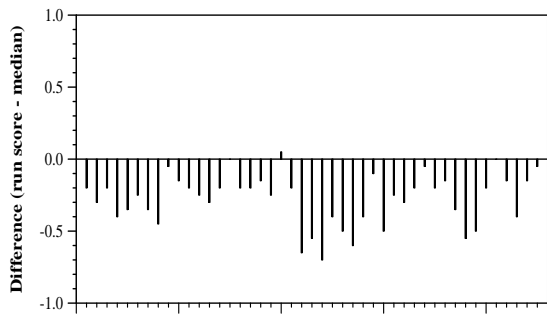


Per-topic difference from median bpref for all Round 4 runs

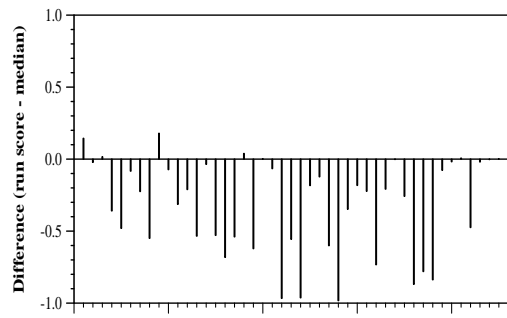


Per-topic difference from median NDCG@20 for all Round 4 runs

Round 4 results — Run jlQErnd4 submitted from julielab



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