

**Round 2 results — Run ir\_covid19\_cle\_IB submitted from IR\_COVID19\_CLE**

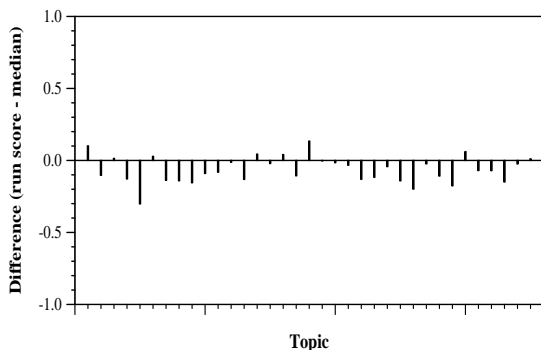
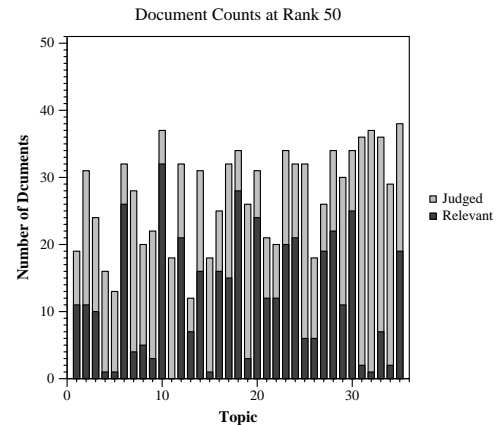
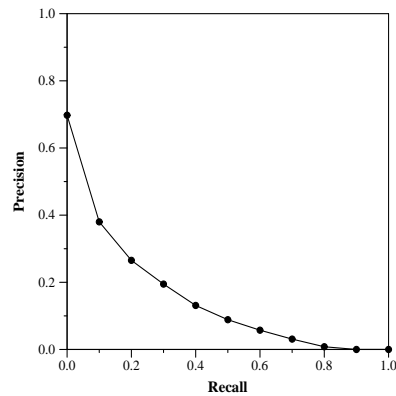
**Run Description**

We have used the data set with all the documents from corpus. Using the metadata available, we added cord\_id. We used around 48000 files for indexing. We used "Paper\_id", "Title Id" and "Abstract" and "body\_text" to index all the documents using Apache Lucene. We have indexed every document for all tokens present with in the document. However, in a collection of documents these tokens can be repeating in multiple documents as well. Here, we use inverted index to store tokens repeating in multiple indexes, so that when searched for a specific token, we can narrow down the search documents specifically all documents that token is present. We have used the query of the topic for querying the index. We parsed the query with English Analyzer and searched on the abstract text field of the index. For each query, We have retrieved the Top 100 documents and the relevance scores using Information based (IB Similarity) which models rely on normalized values of occurrence of a word in documents. Information Models are characterized by three elements that are normalization function, probability distribution and retrieval function. Reference Paper: Clinchant , S., & Gaussier, E. (2010). Information-based models for adhoc IR. In Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval (SIGIR '10).

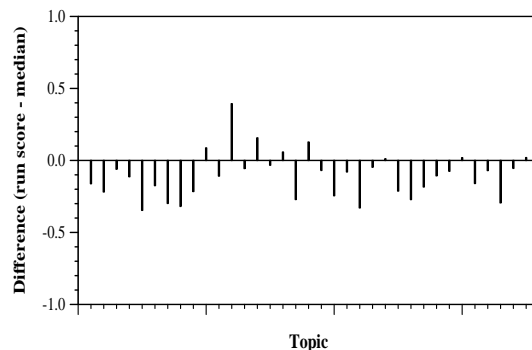
Summary Statistics	
Run ID	ir_covid19_cle_IB
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	35
Total number retrieved	30768
Total relevant	3002
Total relevant retrieved	1482
MAP	0.1423
Mean Bpref	0.3050
Mean NDCG@10	0.3662
Mean RBP(p=0.5)	0.4283 +0.0029

Document Level Averages	
	Precision
At 5 docs	0.4686
At 10 docs	0.4086
At 15 docs	0.3752
At 20 docs	0.3300
At 30 docs	0.2952
R-Precision	
Exact	0.1939



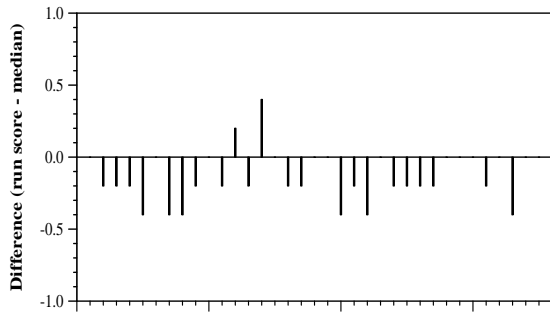
Per-topic difference from median bpref for all Round 2 runs



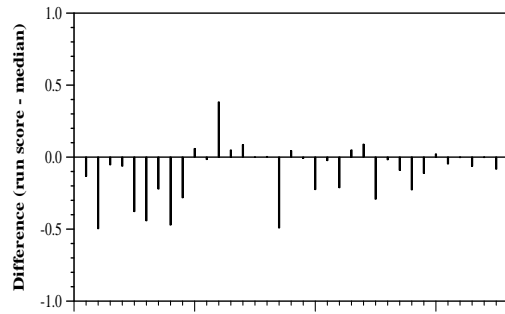
Per-topic difference from median NDCG@10 for all Round 2 runs

Round 2 results — Run ir\_covid19\_cle\_IB submitted from IR\_COVID19\_CLE

---



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