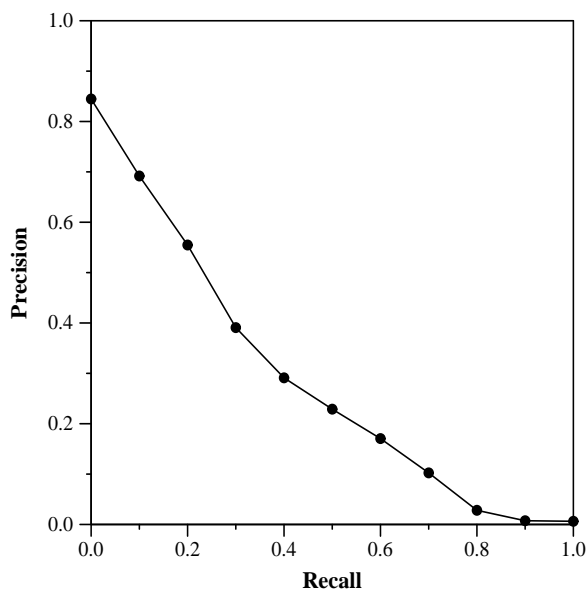
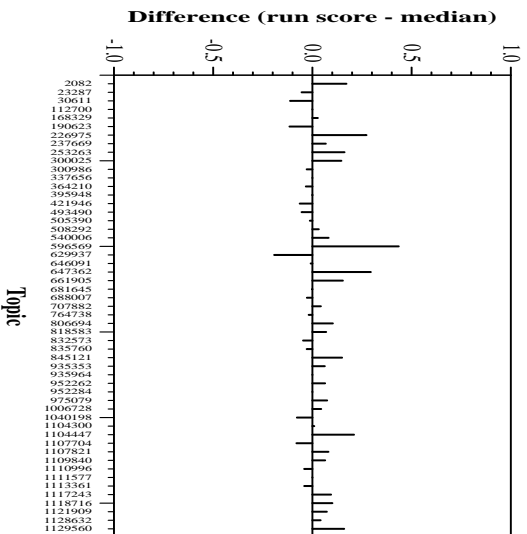


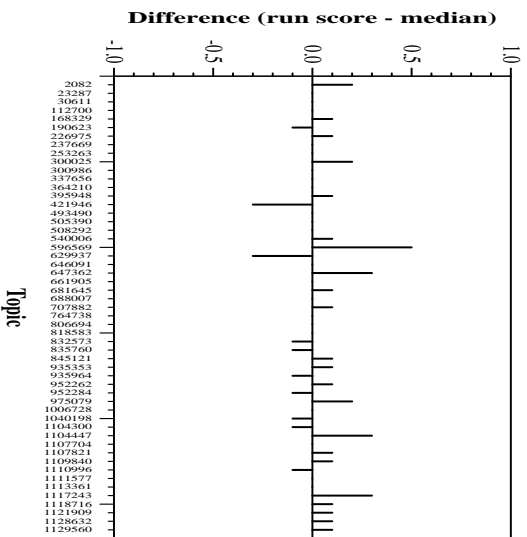
Run Description	
Run ID:	ielab-uniCOIL
Task:	Passage Ranking
Subtask:	Full ranking from the collection
Topic type:	Automatic
Single-stage retrieval?	Yes
Dense retrieval?	No
Used deep nn model?	Pre-trained model
Type of training:	Previous year's MS MARCO training data
Pre-processing/indexing cost:	requires expanding the original passage collection (append additional tokens to each passage); this process took about 1 hour 40 mins per 2 million passages on a single Tesla v100 16G Gpu, we use 6 gpus for this part, thus it took around 20 hours in total. After expanding the whole collection, uniCOIL required encoding the expanded collection which costs around 40 mins per 2 million passages, we again use 6 gpus to do this. Finally, we used Anserini to indexing the encoded collection which took us around 3.5 hours on a 2018 Mac Mini machine with 3.2 GHz 6-Core Intel Core i7 CPU.
Query processing cost:	requires a single BERT inference to process each query, which takes around 40ms per query on GPU. And took 179ms to search the Anserini index per query.

Overall measures		Document Level Averages	
Number of topics	53		Precision
Total number retrieved	5300	At 5 docs	0.5925
Total relevant	3427	At 10 docs	0.5453
Total relevant retrieved	1353	At 15 docs	0.5119
MAP	0.2745	At 20 docs	0.4858
Mean NDCG@10	0.6420	At 30 docs	0.4365
Mean Reciprocal Rank	0.7975	R-Precision	
		Exact	0.3354

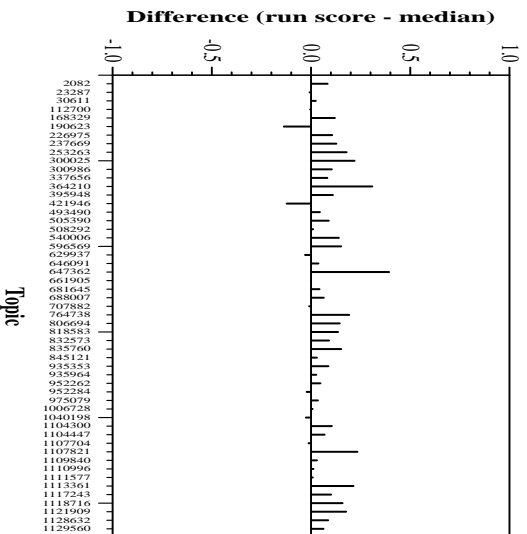




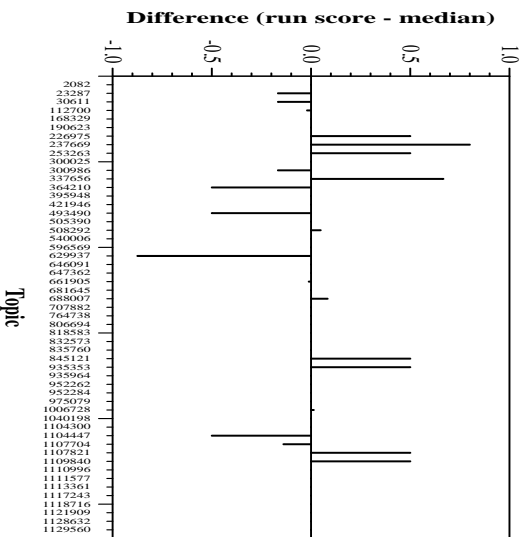
Per-topic difference from median for NDCG@10 for Passage Ranking runs



Per-topic difference from median for Prec@10 for Passage Ranking runs



Per-topic difference from median for Average Precision for Passage Ranking runs



Per-topic difference from median for Reciprocal Rank for Passage Ranking runs