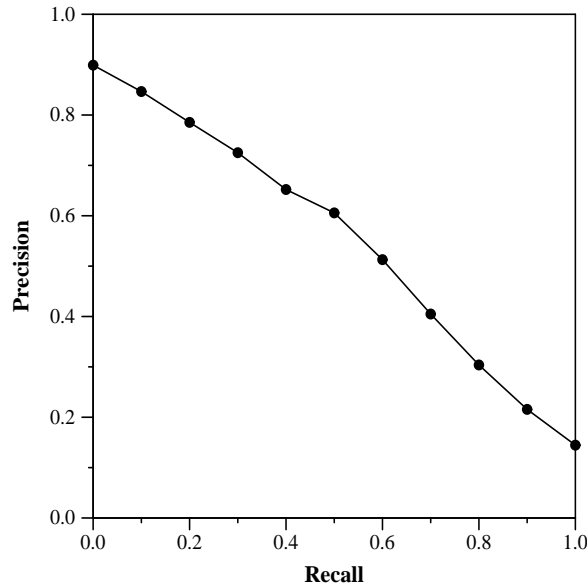
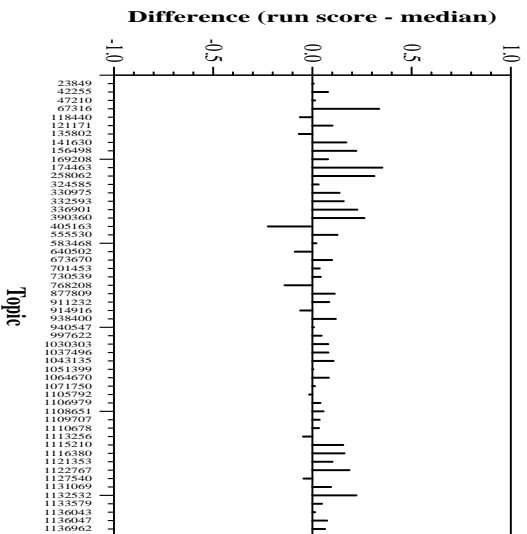


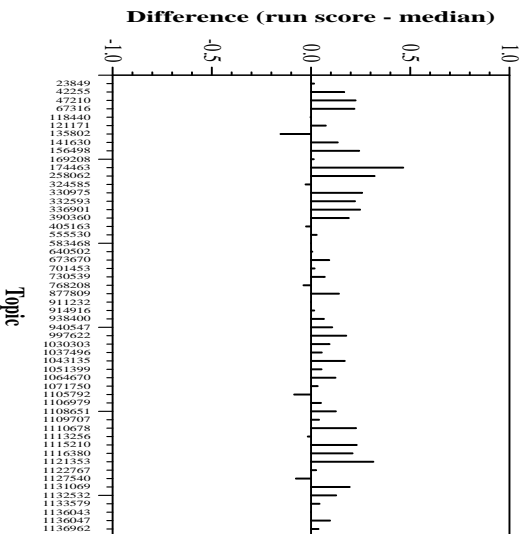
Run Description	
Run ID	CoRT-electra
Task	Passage Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Unsupervised data/model; Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	- pretrained ALBERT model "alber-base-v2" <a href="https://huggingface.co/albert-base-v2">https://huggingface.co/albert-base-v2</a> - pretrained ELECTRA discriminator <a href="https://huggingface.co/google/electra-large-discriminator">https://huggingface.co/google/electra-large-discriminator</a> - 36h of finetuning on a single RTX 2080 Ti (for CoRT) + 92h for the ELECTRA Re-ranker

Overall measures		Document Level Averages	
Number of topics	54	Precision	
Total number retrieved	54000	At 5 docs	0.7148
Total relevant	1666	At 10 docs	0.5889
Total relevant retrieved	1148	At 15 docs	0.5185
MAP	0.5399	At 20 docs	0.4667
Mean NDCG@10	0.7566	At 30 docs	0.3840
Mean NDCG@1000	0.7438	R-Precision	
Mean Reciprocal Rank	0.8703	Exact	0.5353

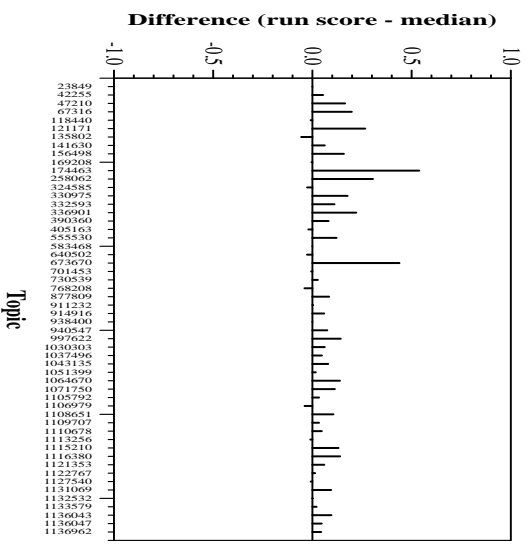




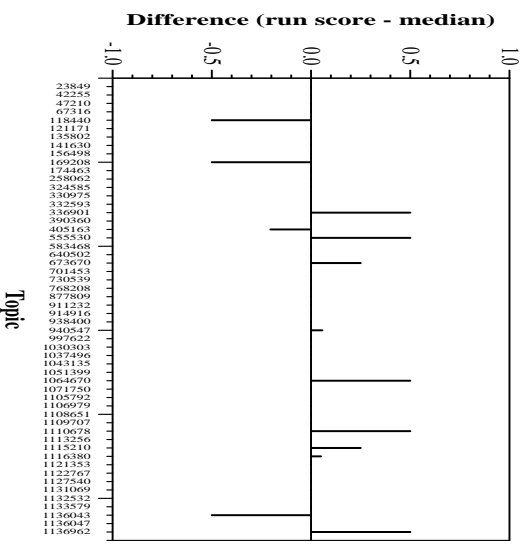
Per-topic difference from median for NDCG@10 for Passage Ranking runs



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



Per-topic difference from median for NDCG@1000 for Passage Ranking runs



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