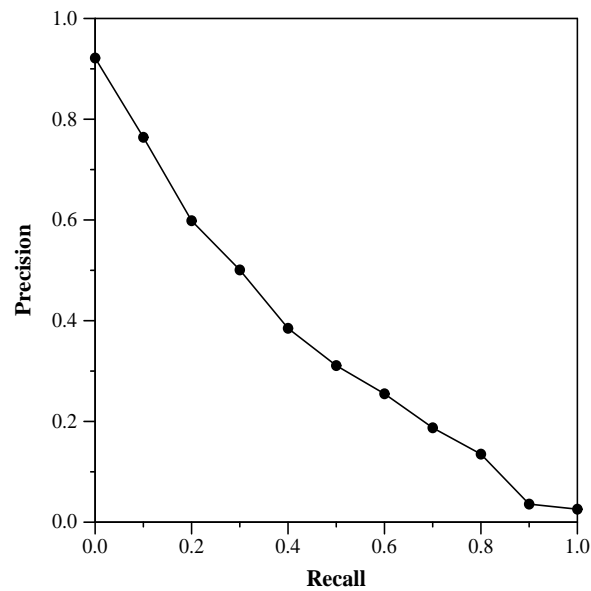
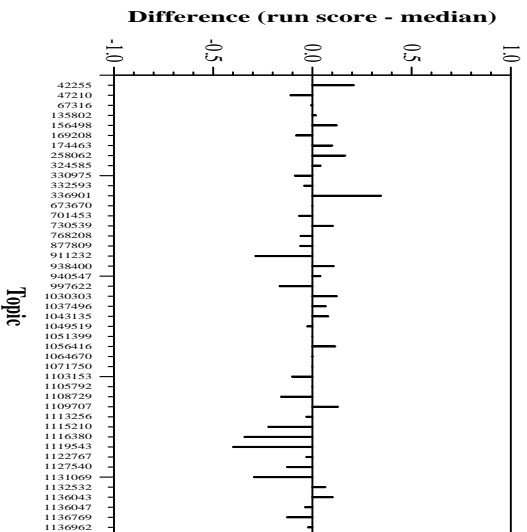


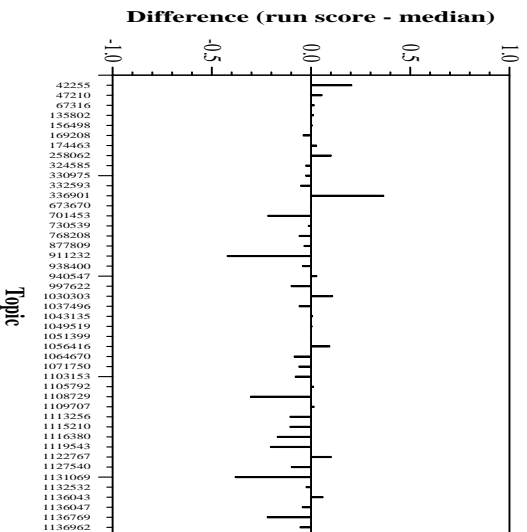
Run Description	
Run ID	longformer_1
Task	Document Ranking
Subtask	Reranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used document task training labels
ORCAS data used for something other than training?	No
Training cost?	4 days on 4 Pascal GPUs for training

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.6711
Total relevant	1767	At 10 docs	0.5511
Total relevant retrieved	814	At 15 docs	0.4904
MAP	0.3503	At 20 docs	0.4244
Mean NDCG@10	0.5520	At 30 docs	0.3578
Mean NDCG@100	0.5675	R-Precision	
Mean Reciprocal Rank	0.8889	Exact	0.3781

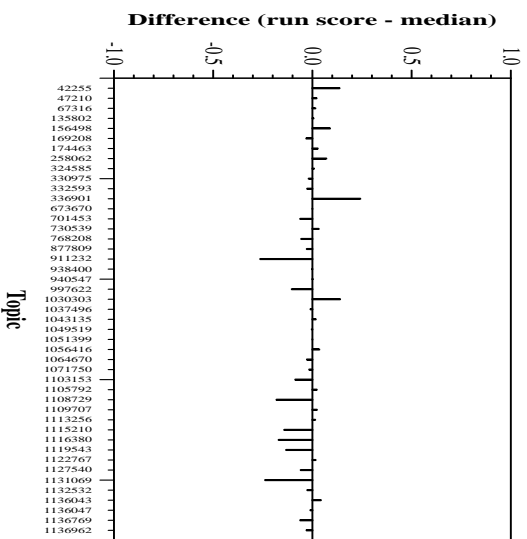




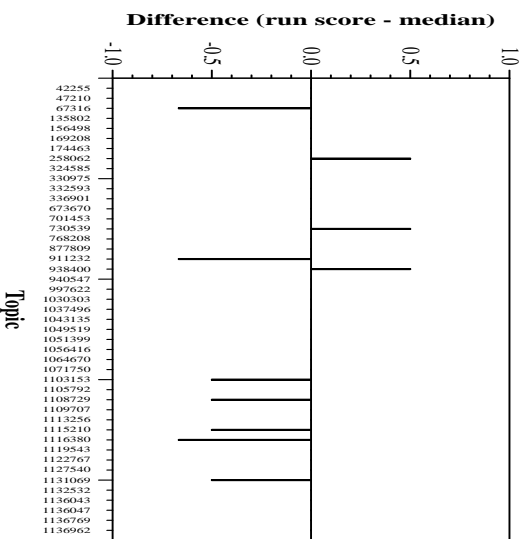
Per-topic difference from median for NDCG@10 for Document Ranking runs



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



Per-topic difference from median for NDCG@100 for Document Ranking runs



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