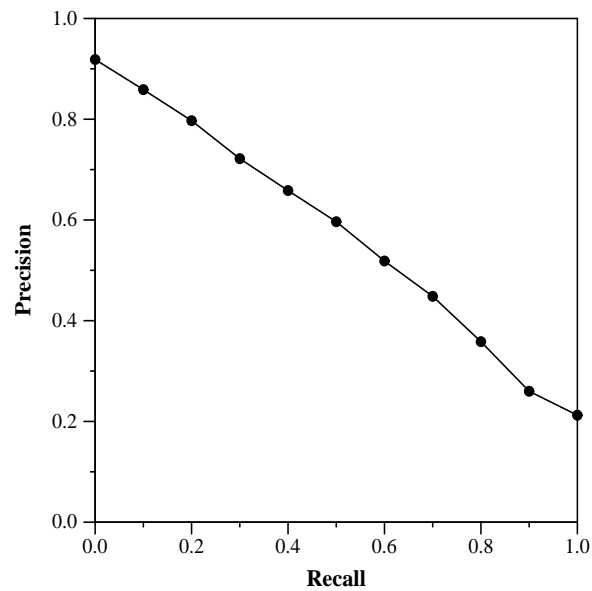
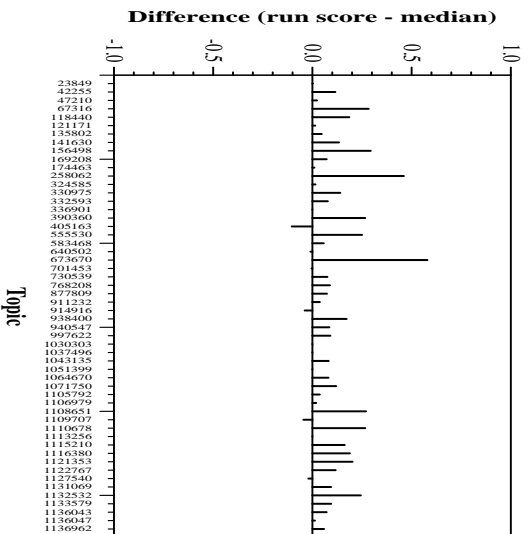


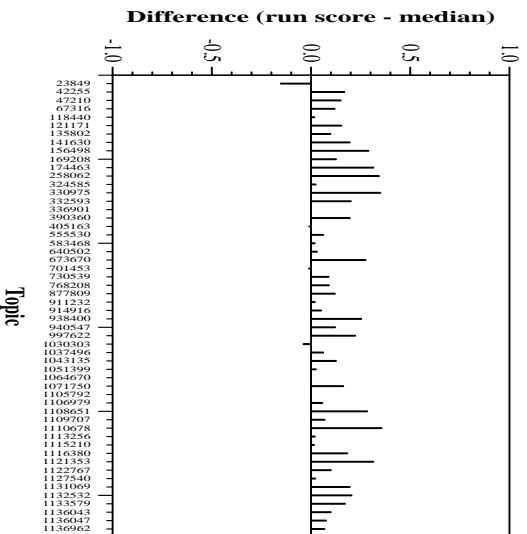
Run Description	
Run ID	p_d2q_bm25_duo
Task	Passage Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	monoT5-3B/duoT5-3B training/inference costs

Overall measures		Document Level Averages	
Number of topics	54		Precision
Total number retrieved	54000	At 5 docs	0.7333
Total relevant	1666	At 10 docs	0.6259
Total relevant retrieved	1177	At 15 docs	0.5457
MAP	0.5609	At 20 docs	0.4796
Mean NDCG@10	0.7837	At 30 docs	0.3852
Mean NDCG@1000	0.7539	R-Precision	
Mean Reciprocal Rank	0.8798	Exact	0.5339

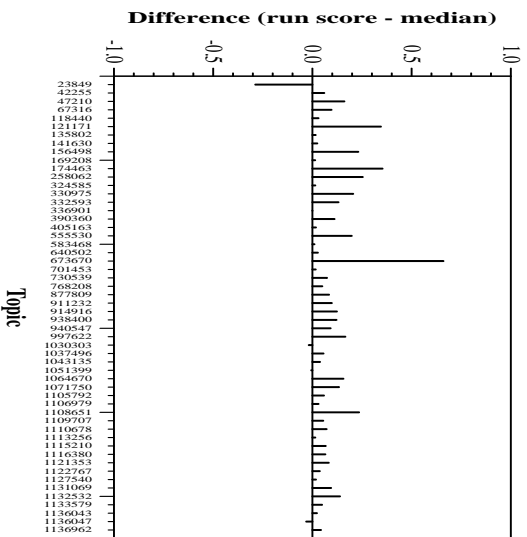




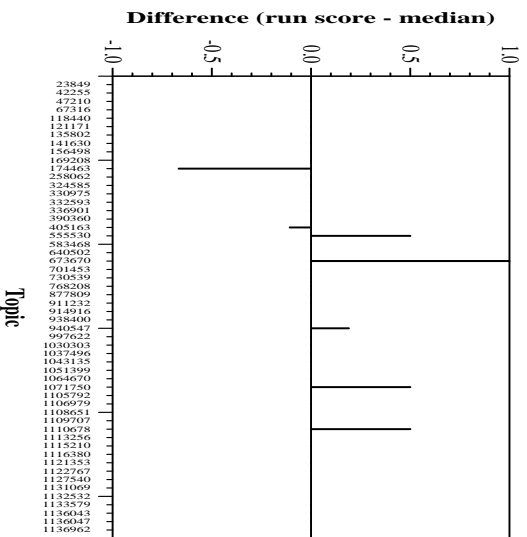
Per-topic difference from median for NDCG@10 for Passage Ranking runs



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