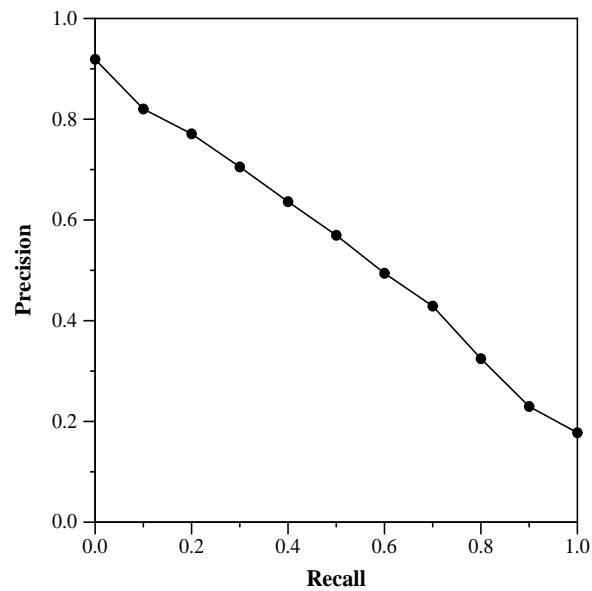
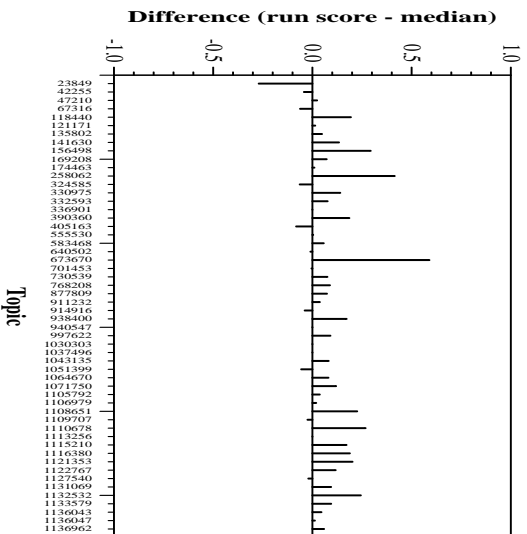


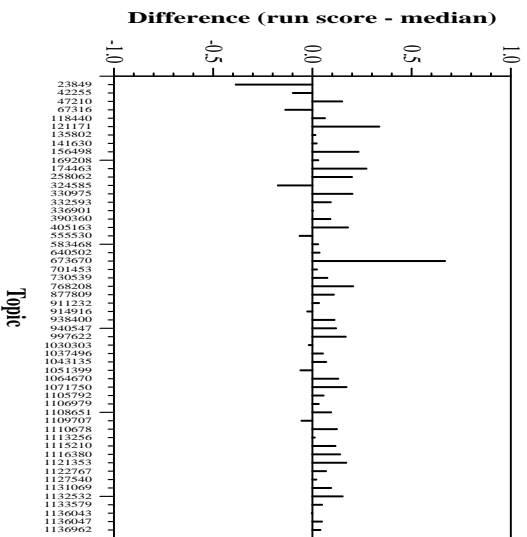
Run Description	
Run ID	p_bm25rm3_duo
Task	Passage Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	monoT5-3B/duoT5-3B training/inference costs

Overall measures		Document Level Averages	
Number of topics	54		Precision
Total number retrieved	54000	At 5 docs	0.6926
Total relevant	1666	At 10 docs	0.5963
Total relevant retrieved	1119	At 15 docs	0.5198
MAP	0.5355	At 20 docs	0.4565
Mean NDCG@10	0.7583	At 30 docs	0.3648
Mean NDCG@1000	0.7387	R-Precision	
Mean Reciprocal Rank	0.8759	Exact	0.5088

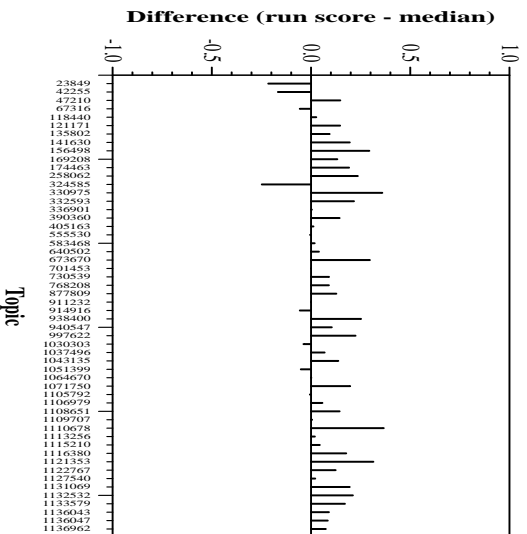




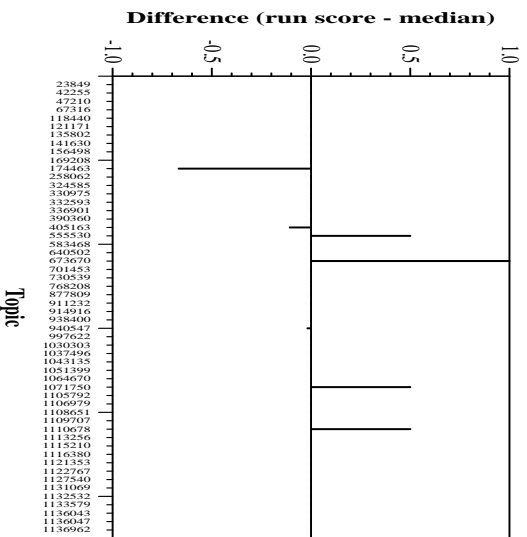
Per-topic difference from median for NDCG@10 for Passage Ranking runs



Per-topic difference from median for NDCG@1000 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