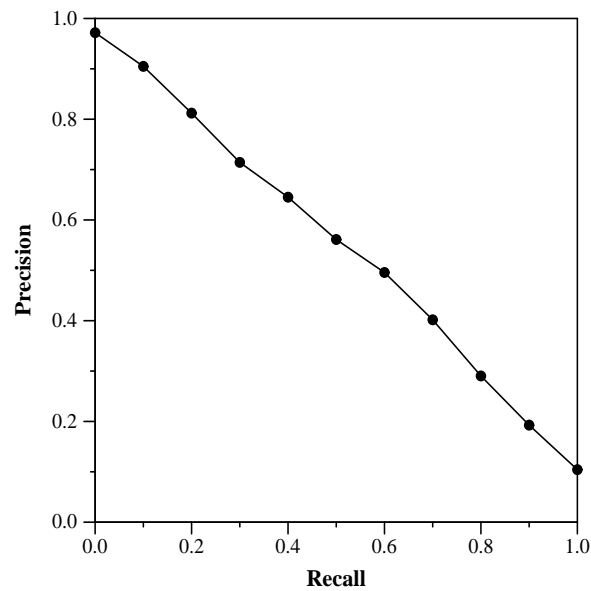
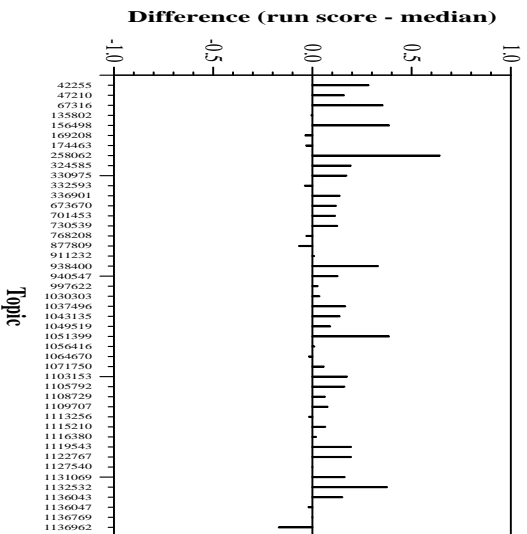


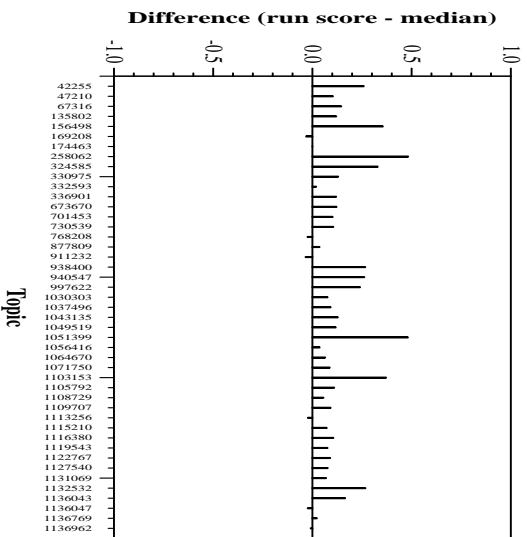
Run Description	
Run ID	d_d2q_rm3_duo
Task	Document Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	monoT5-3B/duoT5-3B training/inference costs

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.8133
Total relevant	1767	At 10 docs	0.7267
Total relevant retrieved	1074	At 15 docs	0.6533
MAP	0.5427	At 20 docs	0.5856
Mean NDCG@10	0.6900	At 30 docs	0.4978
Mean NDCG@100	0.7124	R-Precision	
Mean Reciprocal Rank	0.9476	Exact	0.5390

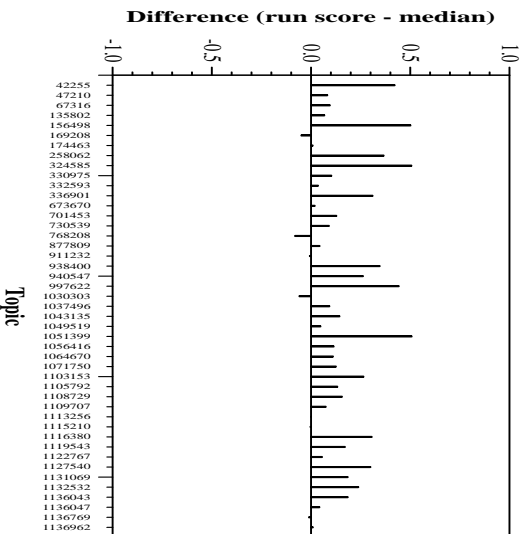




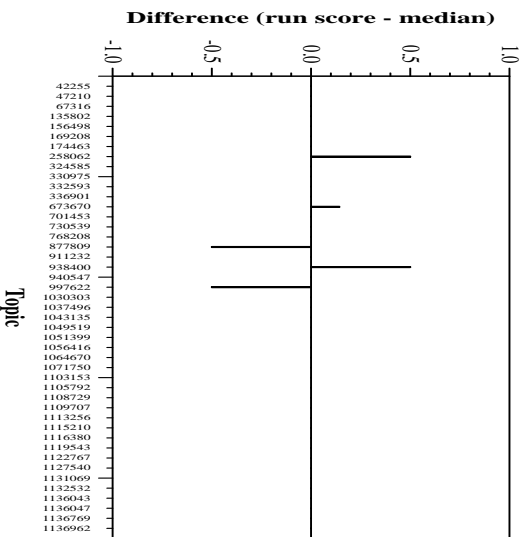
Per-topic difference from median for NDCG@10 for Document Ranking runs



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



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



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