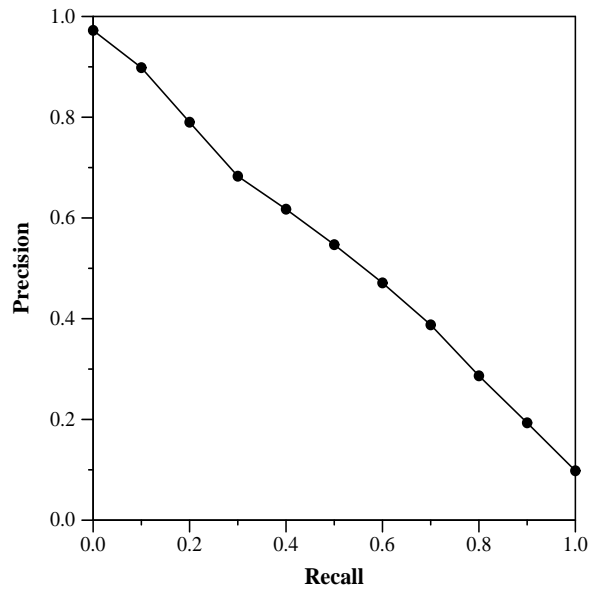
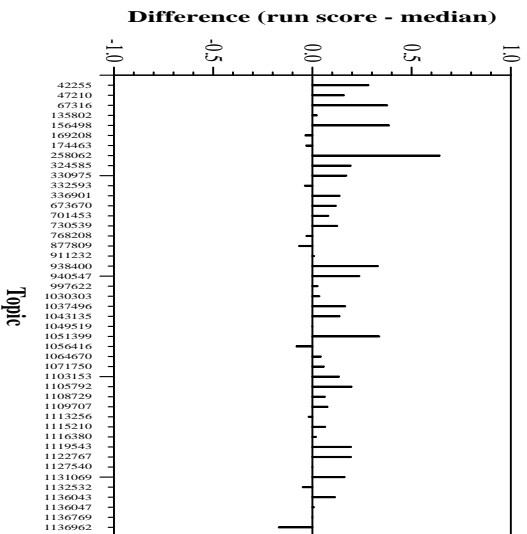


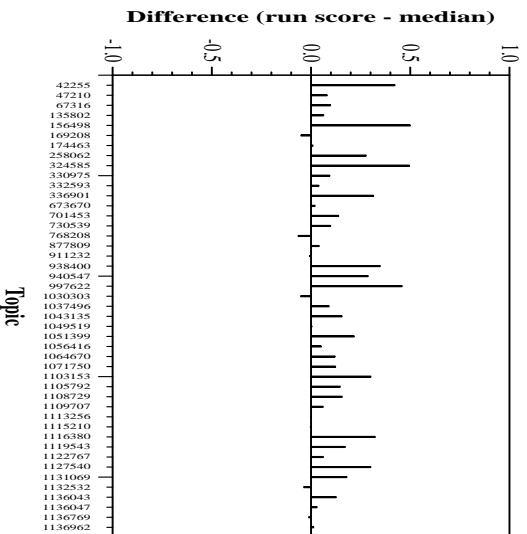
Run Description	
Run ID	d_rm3_duo
Task	Document Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	monoT5-3B/duoT5-3B training/inference costs

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.8044
Total relevant	1767	At 10 docs	0.7222
Total relevant retrieved	1025	At 15 docs	0.6370
MAP	0.5270	At 20 docs	0.5689
Mean NDCG@10	0.6794	At 30 docs	0.4815
Mean NDCG@100	0.6930	R-Precision	
Mean Reciprocal Rank	0.9476	Exact	0.5206

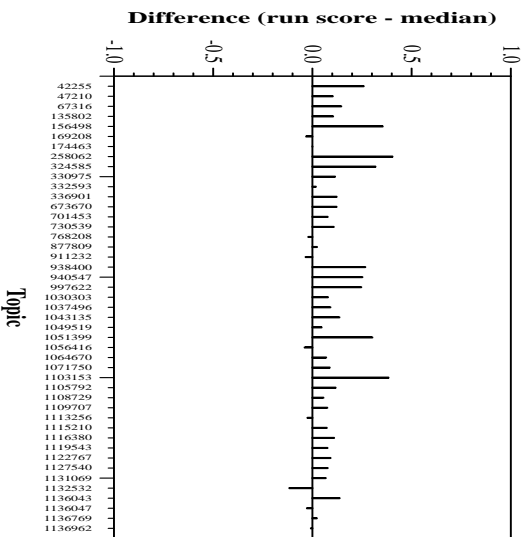




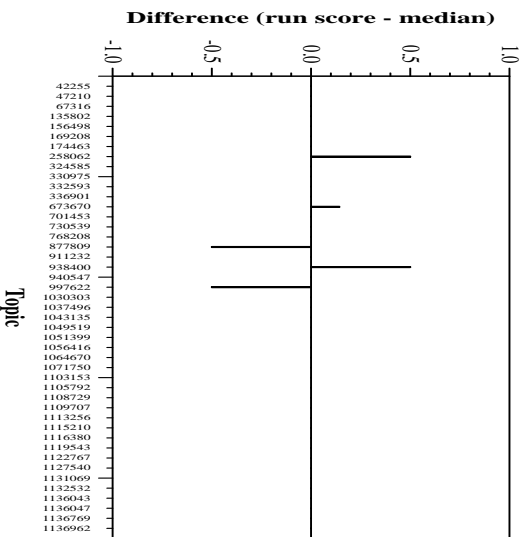
Per-topic difference from median for NDCG@10 for Document Ranking runs



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