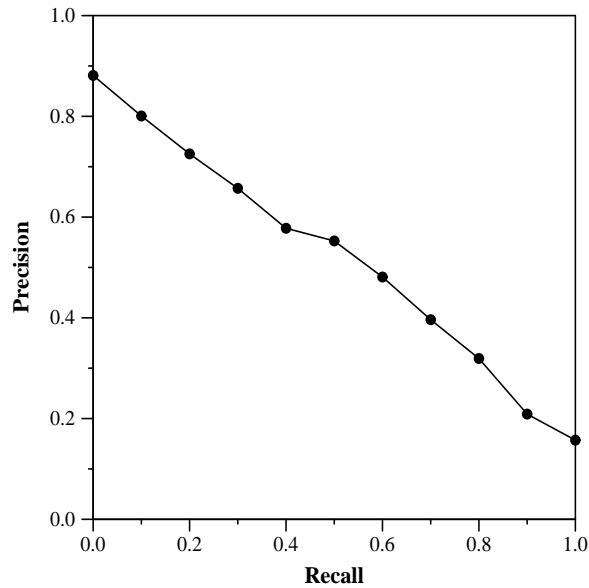
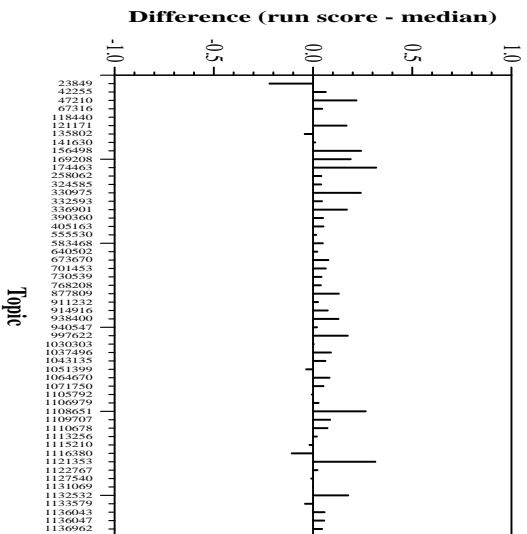
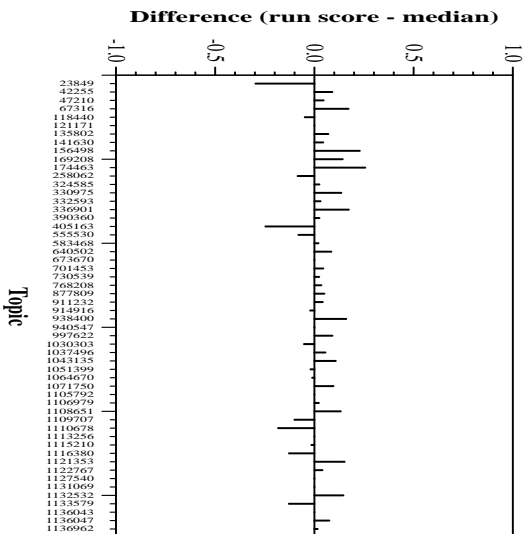


Deep Learning Track results — (QU) Qatar University

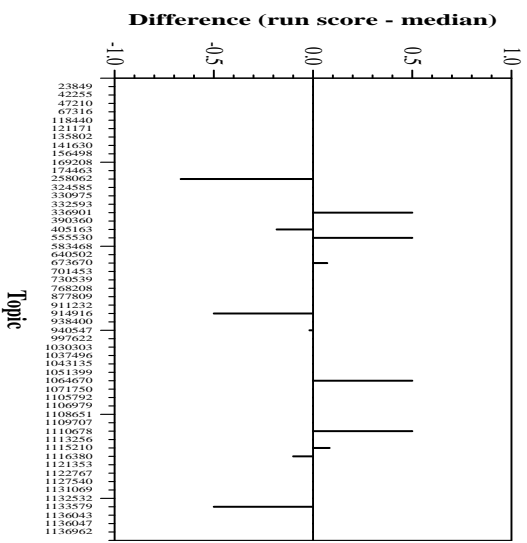
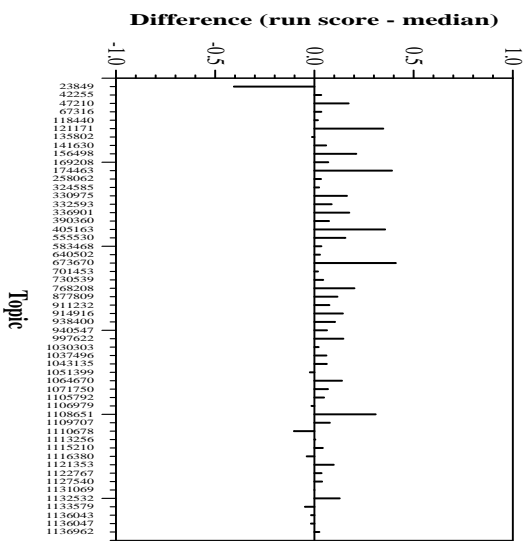
Run Description	
Run ID	bigIR-T5-BERT-F
Task	Passage Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	We used an already pre-trained BERT large model with the MS-Marco passages training data.

Overall measures		Document Level Averages	
Number of topics	54		Precision
Total number retrieved	54000	At 5 docs	0.6444
Total relevant	1666	At 10 docs	0.5444
Total relevant retrieved	1268	At 15 docs	0.4790
MAP	0.5101	At 20 docs	0.4306
Mean NDCG@10	0.7073	At 30 docs	0.3537
Mean NDCG@1000	0.7420	R-Precision	
Mean Reciprocal Rank	0.8478	Exact	0.4875





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



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