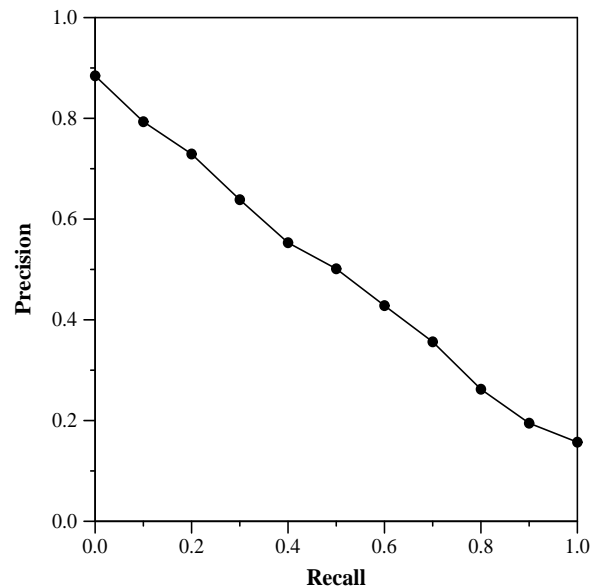
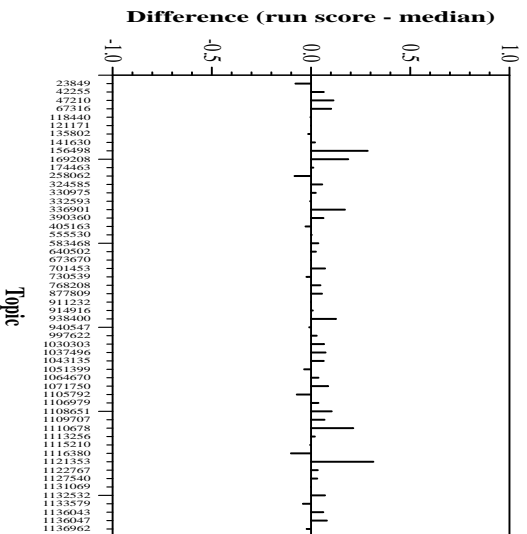
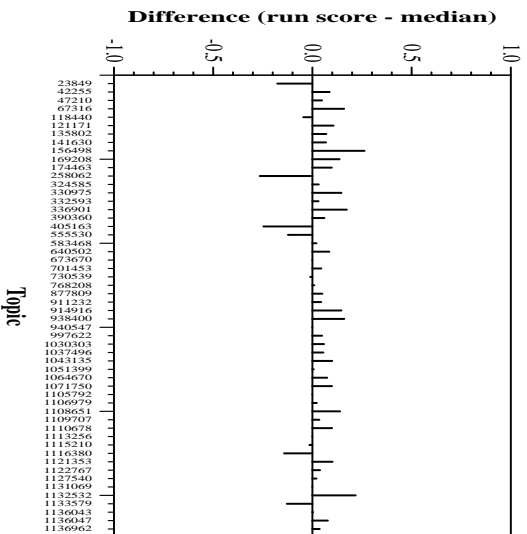


Deep Learning Track results — (QU) Qatar University

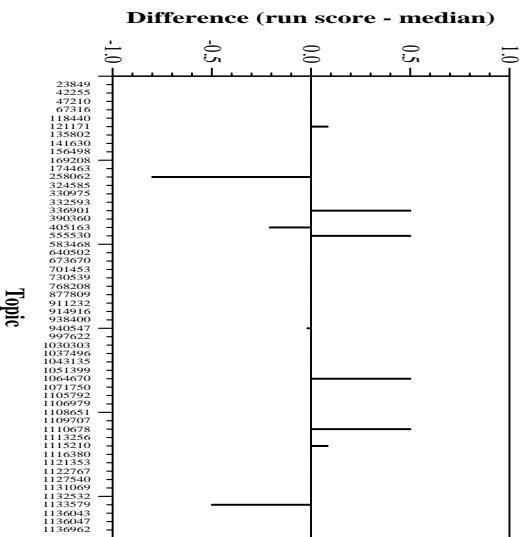
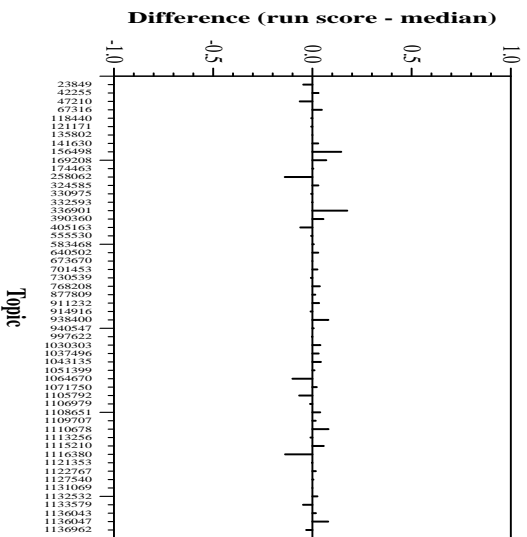
Run Description	
Run ID	bigIR-BERT-R
Task	Passage Ranking
Subtask	Reranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	We used an already pre-trained BERT large model with the MS-Marco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

Overall measures		Document Level Averages	
Number of topics	54		Precision
Total number retrieved	50024	At 5 docs	0.6778
Total relevant	1666	At 10 docs	0.5556
Total relevant retrieved	959	At 15 docs	0.4827
MAP	0.4845	At 20 docs	0.4204
Mean NDCG@10	0.7201	At 30 docs	0.3414
Mean NDCG@1000	0.6732	R-Precision	
Mean Reciprocal Rank	0.8562	Exact	0.4843





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



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