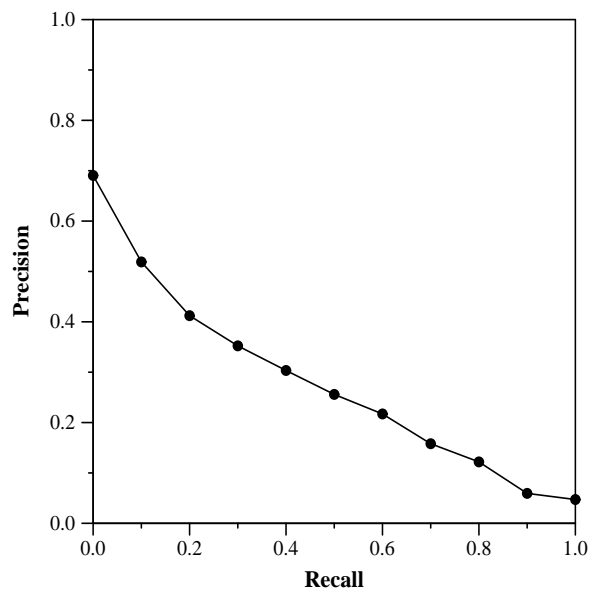
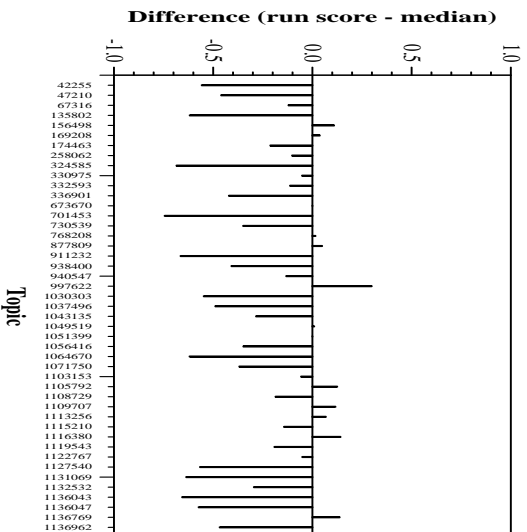


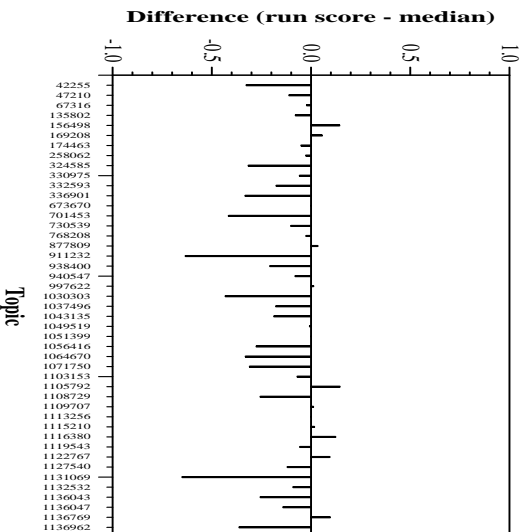
Run Description	
Run ID	mpii_run3
Task	Document Ranking
Subtask	Reranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels; Used document task training labels
ORCAS data used for something other than training?	No
Training cost?	Training was on a Google TPU v3: 11 hours on the passage dataset followed by 10 minutes on the document dataset.

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.4267
Total relevant	1767	At 10 docs	0.3933
Total relevant retrieved	814	At 15 docs	0.3748
MAP	0.2587	At 20 docs	0.3389
Mean NDCG@10	0.3286	At 30 docs	0.2926
Mean NDCG@100	0.4474	R-Precision	
Mean Reciprocal Rank	0.6388	Exact	0.2936

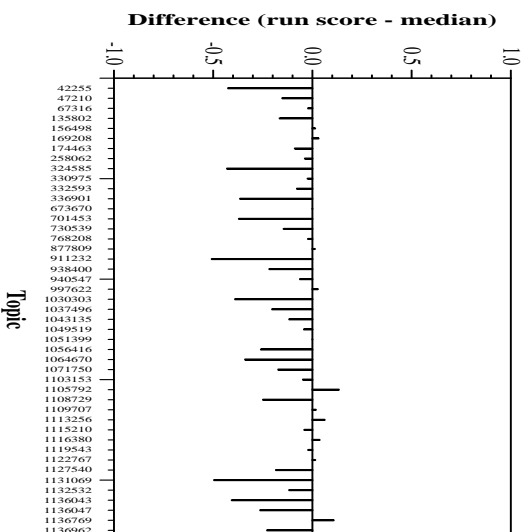




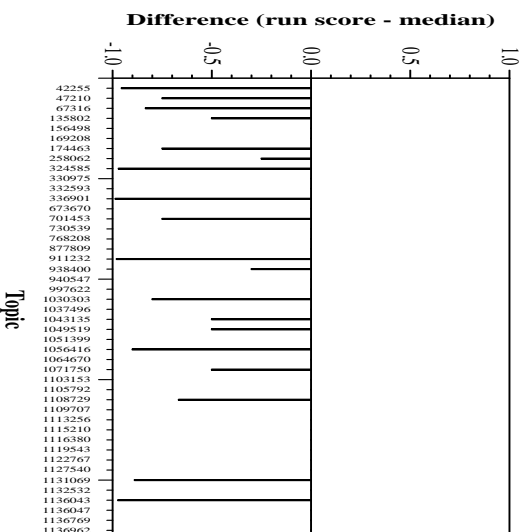
Per-topic difference from median for NDCG@10 for Document Ranking runs



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