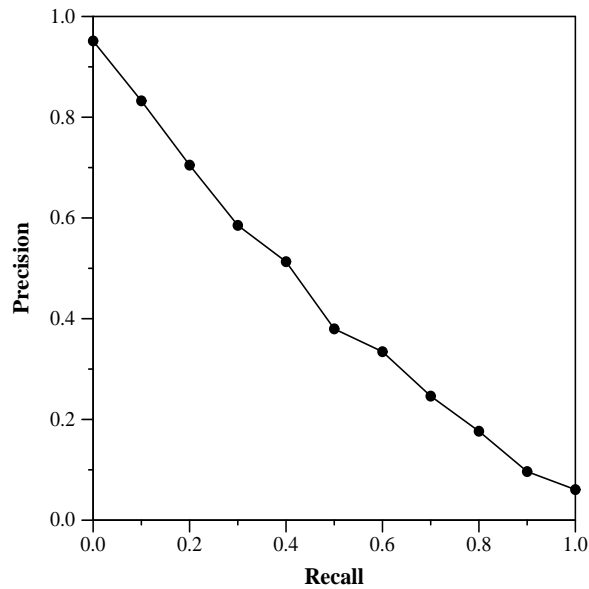
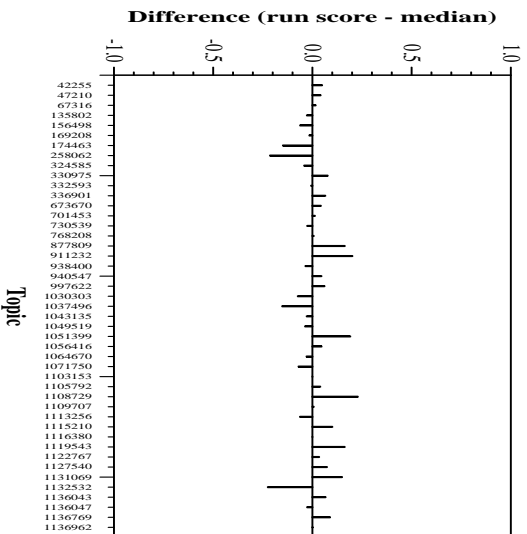


Deep Learning Track results — (anserini) Anserini Gaggles

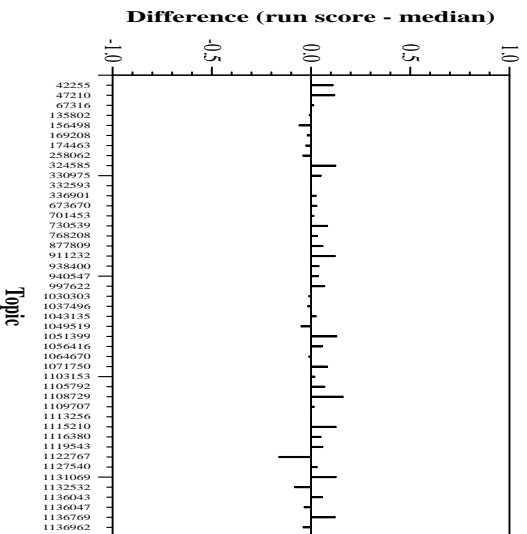
Run Description	
Run ID	d_d2q_bm25
Task	Document Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	Inference cost of T5-based expansion on entire corpus.

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.7289
Total relevant	1767	At 10 docs	0.6089
Total relevant retrieved	875	At 15 docs	0.5333
MAP	0.4230	At 20 docs	0.4789
Mean NDCG@10	0.5885	At 30 docs	0.4000
Mean NDCG@100	0.6117	R-Precision	
Mean Reciprocal Rank	0.9369	Exact	0.4441

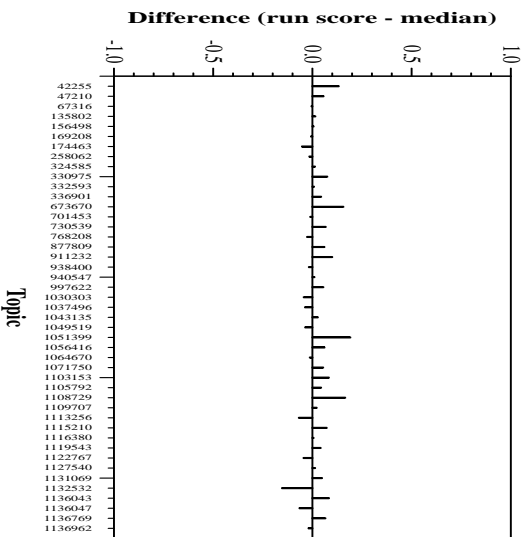




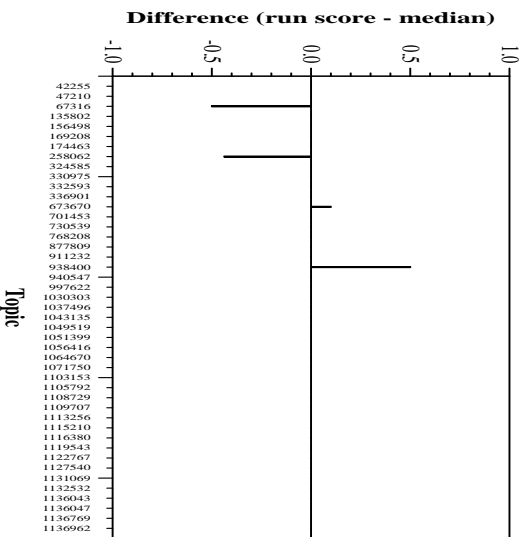
Per-topic difference from median for NDCG@10 for Document Ranking runs



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