Deep Learning Track results — (NLM) U.S. National Library of Medicine
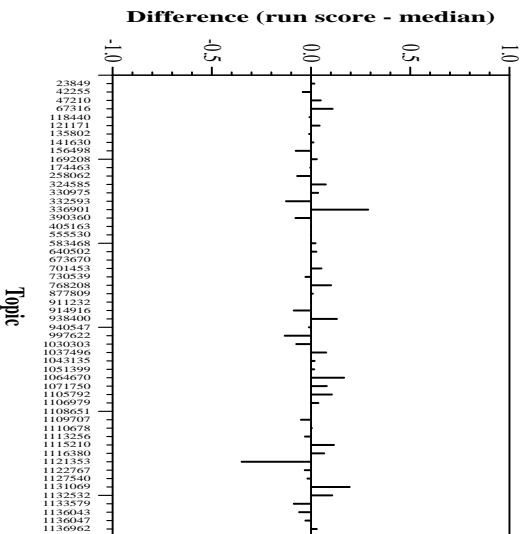
Run Description	
Run ID	nlm-ens-bst-3
Task	Passage Ranking
Subtask	Full ranking
Topic type	Automatic
Used deep nn model?	Yes
Type of training	Unsupervised data/model; Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	$\tilde{3}0$ hours of GPU compute time (1 high-end GPU) * multiple experiments/configs $\tilde{1}2$ hours of CPU/40G-RAM x multiple experiments/configs

Overall measures	
Number of topics	54
Total number retrieved	53946
Total relevant	1666
Total relevant retrieved	1067
MAP	0.4526
Mean NDCG@10	0.6803
Mean NDCG@1000	0.6851
Mean Reciprocal Rank	0.8491

Document Level Averages	
	Precision
At 5 docs	0.6185
At 10 docs	0.5056
At 15 docs	0.4383
At 20 docs	0.4046
At 30 docs	0.3235
R-Precision	
Exact	0.4474

Deep Learning Track results — (NLM) U.S. National Library of Medicine
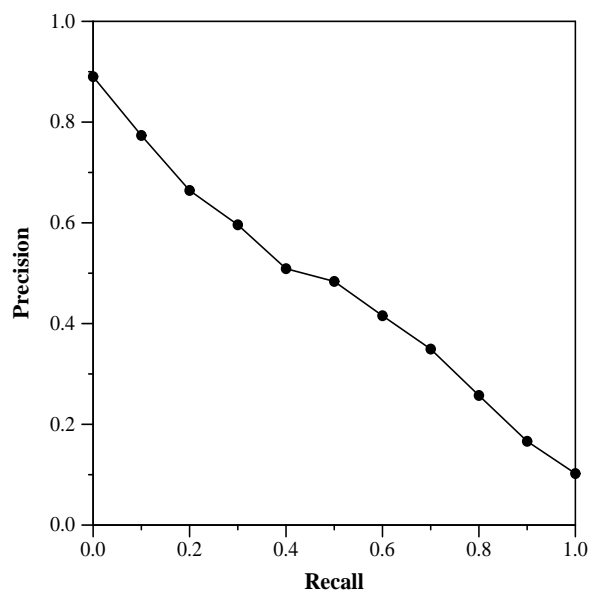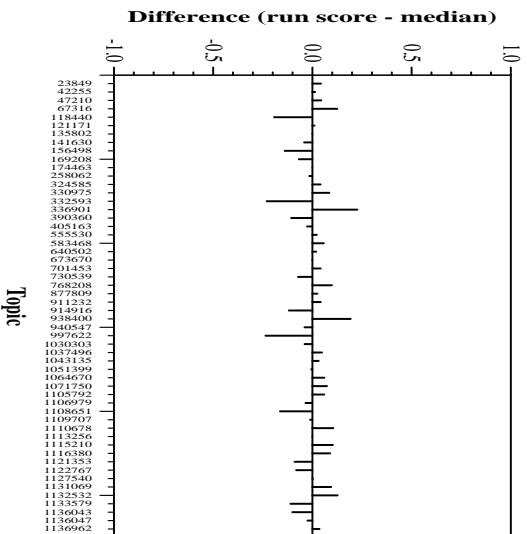
Per-topic difference from median for
NDCG@10 for Passage Ranking runs

Difference (run score - median)

-1.0   -0.5   0.0   0.5   1.0

Topic

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

Difference (run score - median)

-1.0   -0.5   0.0   0.5   1.0
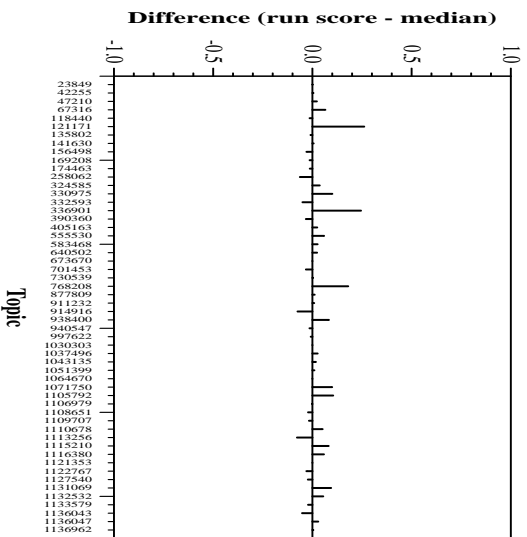
Topic

Per-topic difference from median for
NDCG@1000 for Passage Ranking runs

Difference (run score - median)

-1.0   -0.5   0.0   0.5   1.0
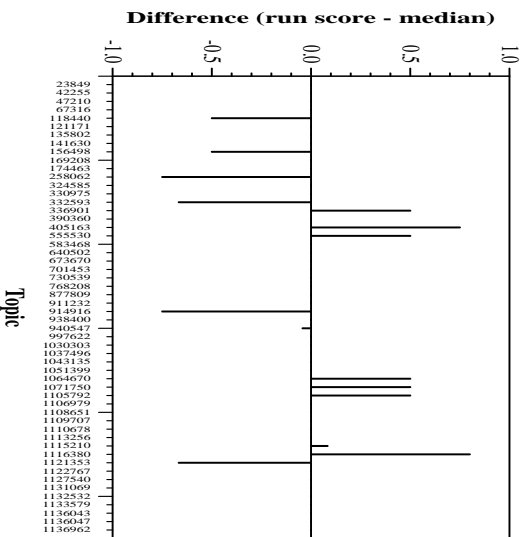
Topic

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

Difference (run score - median)

-1.0   -0.5   0.0   0.5   1.0

Topic

23849
42255
47210
67316
118440
121171
135802
141630
156498
169208
174463
258062
324585
330975
332593
336901
390360
405163
555530
583468
640502
673670
701453
730539
768208
877809
911232
914916
938400
940547
997622
1030303
1037496
1043135
1051399
1064670
1071750
1105792
1106979
1108651
1109707
1110678
1113256
1115210
1116380
1121353
1122767
1127540
1131069
1132532
1133579
1136043
1136047
1136962