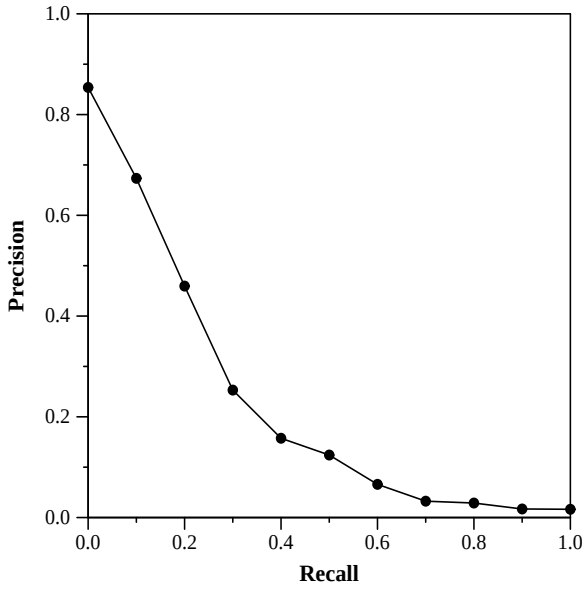
Deep Learning Track results — (IHSM) IHS Markit

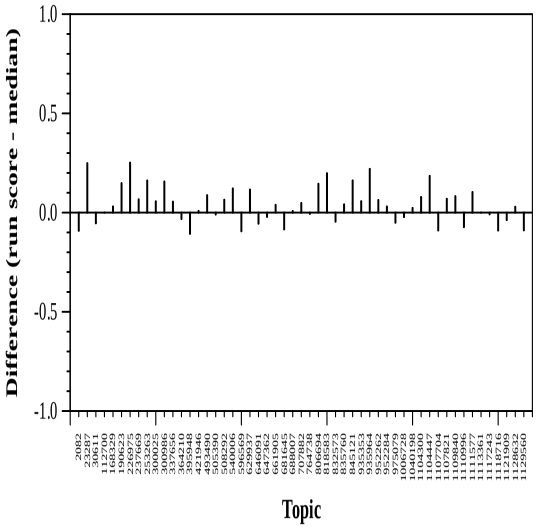
Run Description	
Run ID:	ihsm_bicolbert
Task:	Passage Ranking
Subtask:	Reranking the official top 100
Topic type:	Automatic
Single-stage retrieval?	Yes
Dense retrieval?	No
Used deep nn model?	Pre-trained model
Type of training:	Previous year's MS MARCO training data
Pre-processing/indexing cost:	Vectorization cost: 2 ms per candidate in vectorization of passages with batch size of 256 and with maximum length set to 160 tokens (using single RTX 2080 TI). ∼5 hours to vectorize MSMarco index. Index storage cost: n vectors of 256 binary values per candidate, where n is the number of tokens in a candidate. 18 GB for MSMarco index.
Query processing cost:	19 ms to vectorize single query + 38 ms to rank 1000 candidates on a single RTX 2080 TI.

Overall measures	
Number of topics	53
Total number retrieved	5300
Total relevant	3427
Total relevant retrieved	879
MAP	0.2111
Mean NDCG@10	0.6393
Mean Reciprocal Rank	0.7962

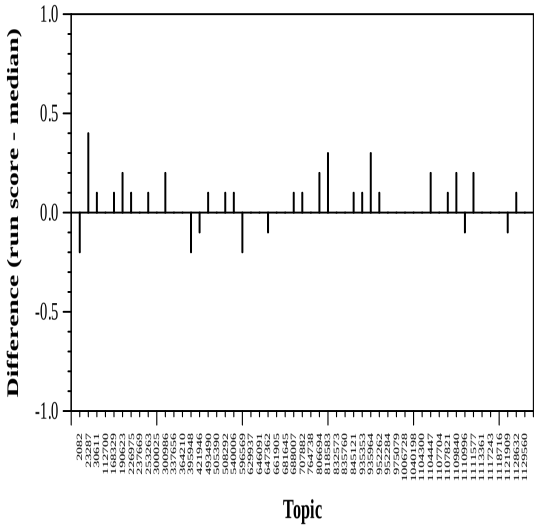
Document Level Averages	
	Precision
At 5 docs	0.6075
At 10 docs	0.5547
At 15 docs	0.4994
At 20 docs	0.4575
At 30 docs	0.3811

R-Precision	
Exact	0.2680

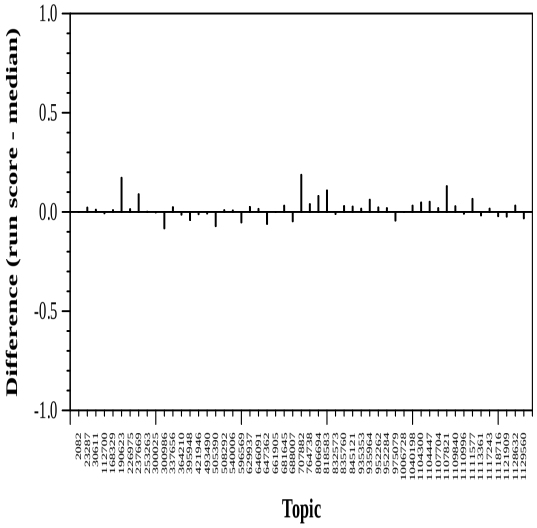




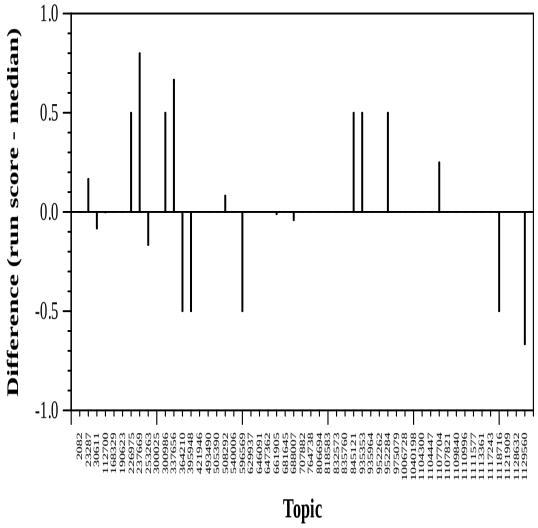
Per-topic difference from median for NDCG@10 for Passage Ranking runs



Per-topic difference from median for Prec@10 for Passage Ranking runs



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



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