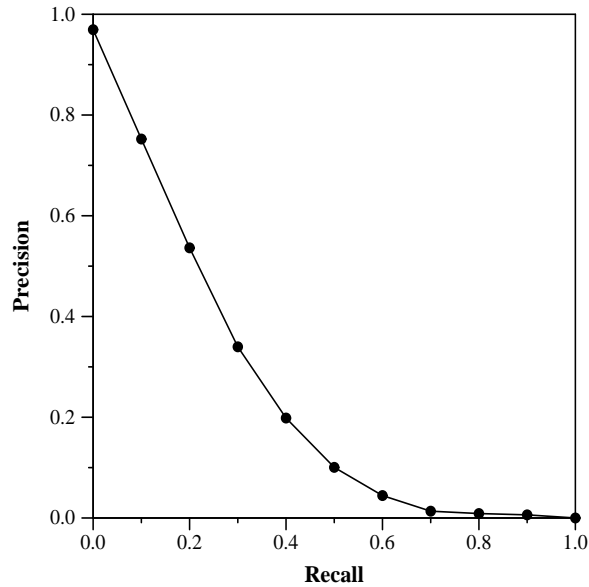
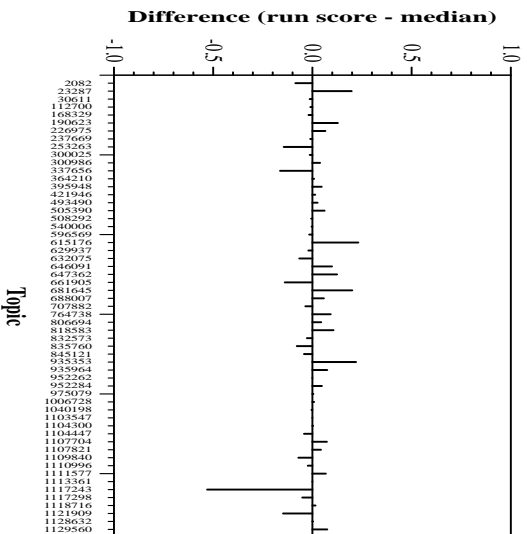


Deep Learning Track results — (CIP) Chinese Information Processing Lab

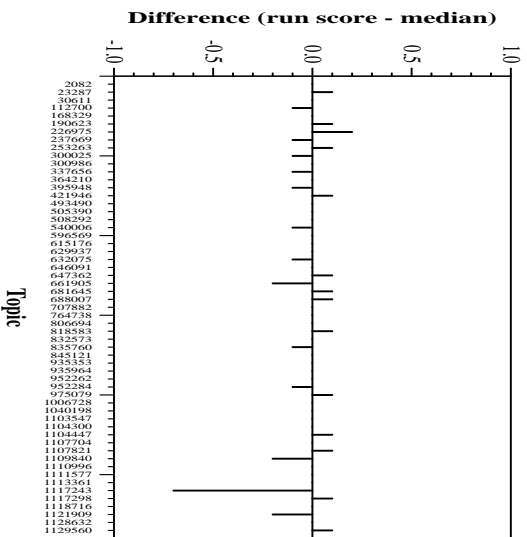
Run Description	
Run ID:	CIP_run3
Task:	Document Ranking
Subtask:	Reranking the official top 100
Topic type:	Automatic
Single-stage retrieval?	No
Dense retrieval?	No
Used deep nn model?	Pre-trained model
Type of training:	This year's MS MARCO training data; Previous year's MS MARCO training data
Pre-processing/indexing cost:	We use the official top-100 candidates, so there is no cost before the query arrives.
Query processing cost:	The inference time of BERT-large for the 477 test queries is about 8 hours with one RTX 3090 GPU.

Overall measures		Document Level Averages	
Number of topics	57		Precision
Total number retrieved	5700	At 5 docs	0.8702
Total relevant	8203	At 10 docs	0.8175
Total relevant retrieved	2086	At 15 docs	0.7637
MAP	0.2457	At 20 docs	0.7184
Mean NDCG@10	0.6668	At 30 docs	0.6327
Mean Reciprocal Rank	0.9567		
		R-Precision	
		Exact	0.3003

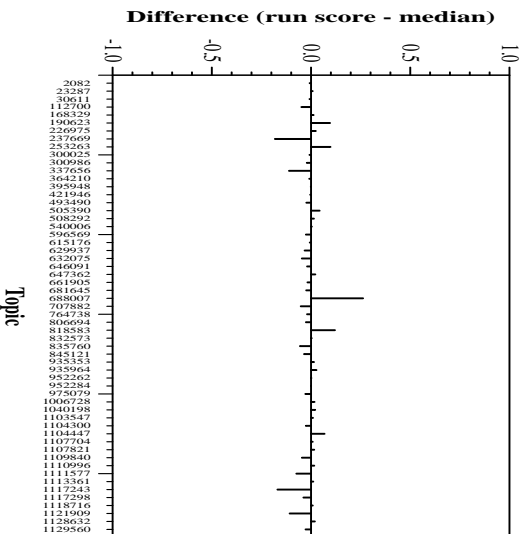




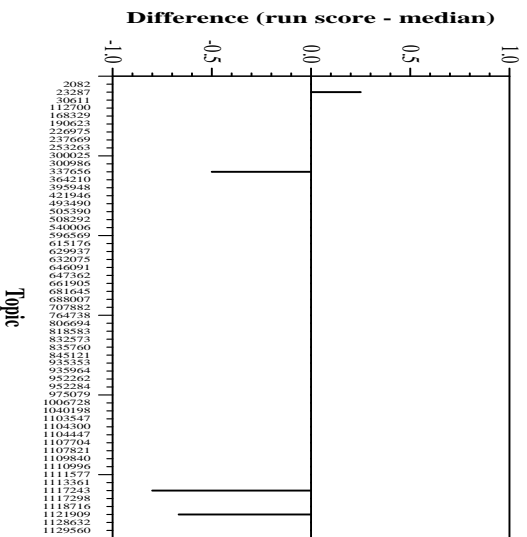
Per-topic difference from median for NDCG@10 for Document Ranking runs



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



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



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