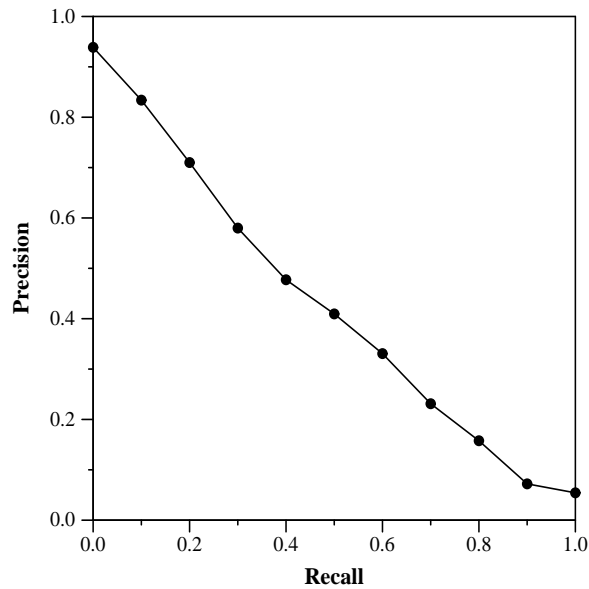
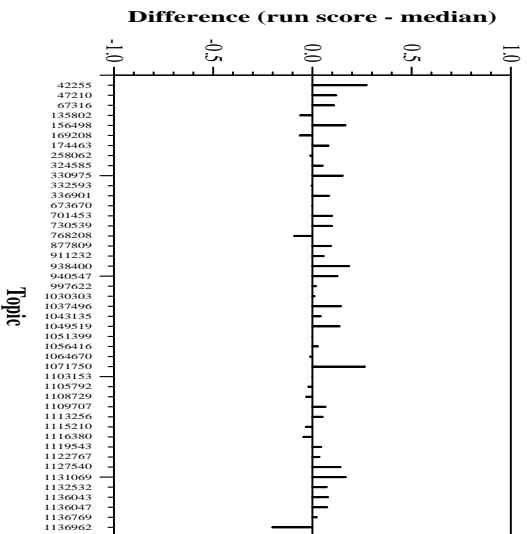


Deep Learning Track results — (BITEM) SIB Text Mining - BiTeM group

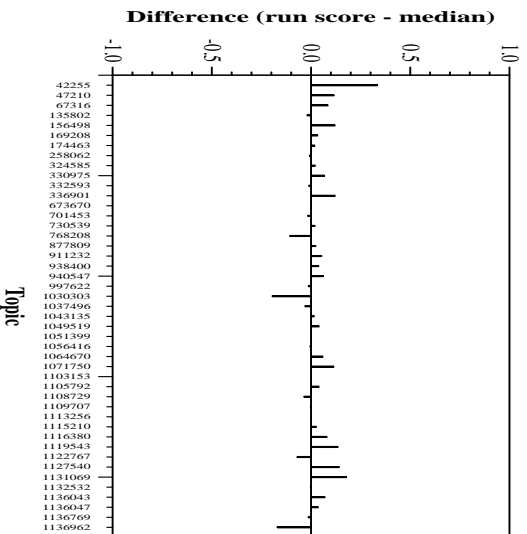
Run Description	
Run ID	roberta-large
Task	Document Ranking
Subtask	Reranking
Topic type	Automatic
Used deep nn model?	No
Type of training	Used passage task training labels
ORCAS data used for something other than training?	No
Training cost?	Training of the passage model of about 26 hours on 8 GPUs (V100)

Overall measures		Document Level Averages	
Number of topics	45		Precision
Total number retrieved	4500	At 5 docs	0.7511
Total relevant	1767	At 10 docs	0.6378
Total relevant retrieved	814	At 15 docs	0.5496
MAP	0.4199	At 20 docs	0.4933
Mean NDCG@10	0.6295	At 30 docs	0.4148
Mean NDCG@100	0.6060	R-Precision	
Mean Reciprocal Rank	0.9185	Exact	0.4423

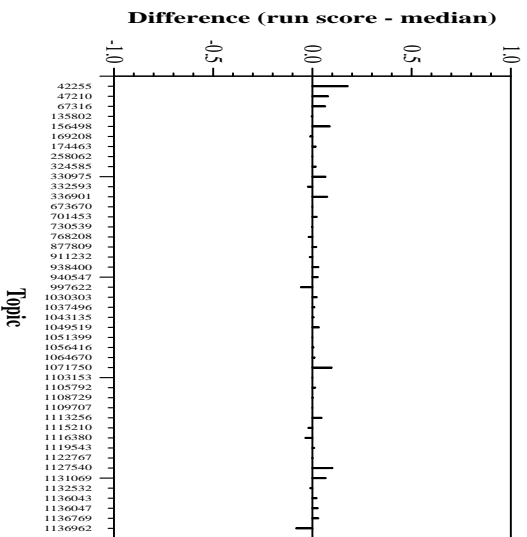




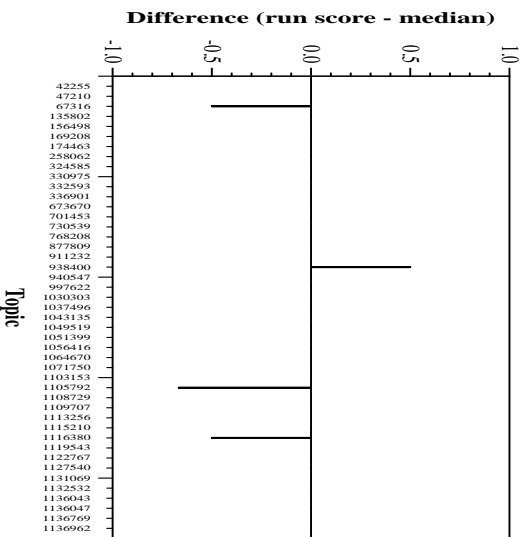
Per-topic difference from median for NDCG@10 for Document Ranking runs



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