

Round 1 results — Run IRIT_marked_un_pair submitted from IRIT_markers

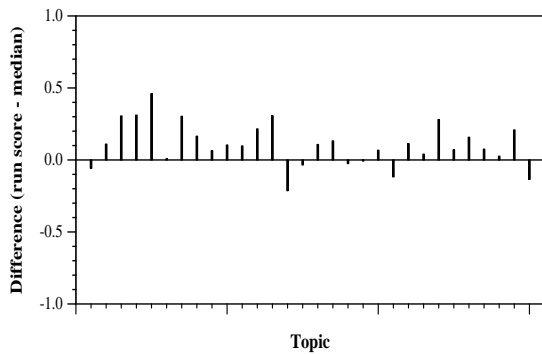
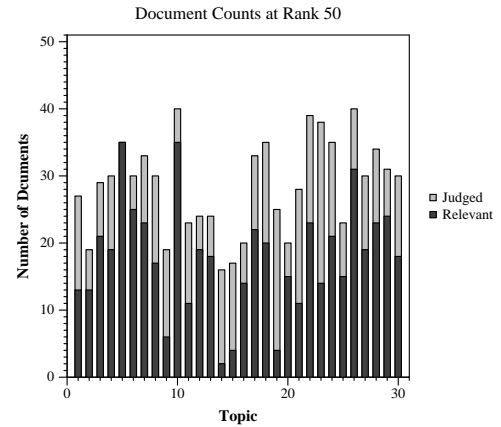
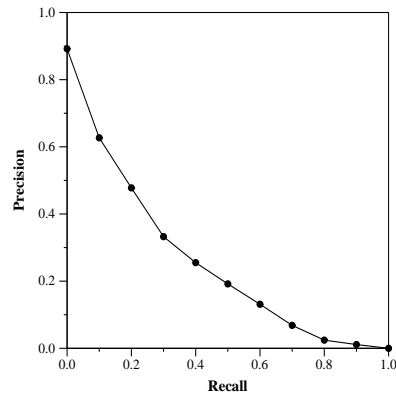
Run Description

We use a BERT-base (12 layers, 768 hidden size) fine-tuned on Ms Marco passage set with a marking strategy that puts focus on exact match signals between query and document terms. We use a full ranking strategy with two stages: in the first stage, we use Anserini Bm25+ RM3 to retrieve top-1000 candidate documents for each topic using an index on the title+abstract of the CORD-19 documents, then we use the fine-tuned BERT to re-rank this list.

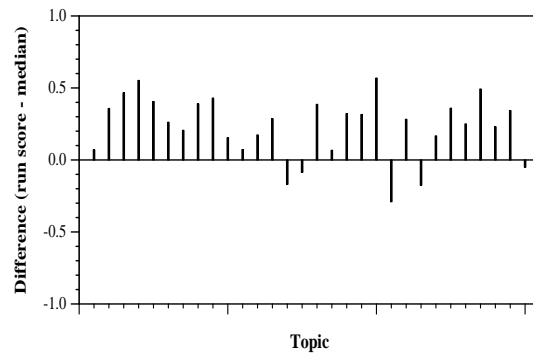
| Summary Statistics | |
|-------------------------------|---------------------|
| Run ID | IRIT_marked_un_pair |
| Topic type | automatic |
| Contributed to judgment sets? | no |

| Overall measures | |
|--------------------------|--------|
| Number of topics | 30 |
| Total number retrieved | 29994 |
| Total relevant | 2352 |
| Total relevant retrieved | 1521 |
| MAP | 0.2481 |
| Mean Bpref | 0.4299 |
| Mean NDCG@10 | 0.5859 |

| Document Level Averages | |
|-------------------------|-----------|
| | Precision |
| At 5 docs | 0.7333 |
| At 10 docs | 0.6200 |
| At 15 docs | 0.5467 |
| At 20 docs | 0.5017 |
| At 30 docs | 0.4378 |
| R-Precision | |
| Exact | 0.3049 |

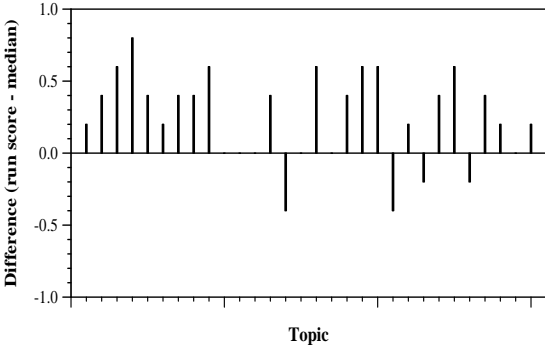


Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs

Round 1 results — Run IRIT_marked_un_pair submitted from IRIT_markers



Per-topic difference from median P@5 for all Round 1 runs