## Run Description

End-to-end retrieval with siamese BERT encoder trained on NLI data (SNLI, Multi-NLI and a new dataset with artificial inference examples). We used the narratives as queries.

| Summary Statistics | |
|---|---|
| Run ID | factum-1 |
| Topic type | automatic |
| Contributed to judgment sets? | yes |

| Overall measures | |
|---|---|
| Number of topics | 30 |
| Total number retrieved | 27989 |
| Total relevant | 2352 |
| Total relevant retrieved | 629 |
| MAP | 0.0590 |
| Mean Bpref | 0.2014 |
| Mean NDCG@10 | 0.1713 |

| Document Level Averages | |
|---|---|
| | Precision |
| At 5 docs | 0.2667 |
| At 10 docs | 0.2200 |
| At 15 docs | 0.1867 |
| At 20 docs | 0.1750 |
| At 30 docs | 0.1522 |
| R-Precision | |
| Exact | 0.1021 |





Document Counts at Rank 50
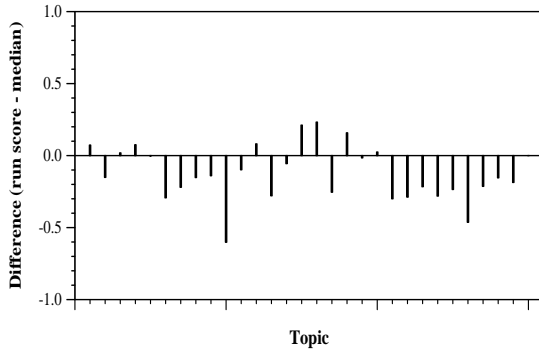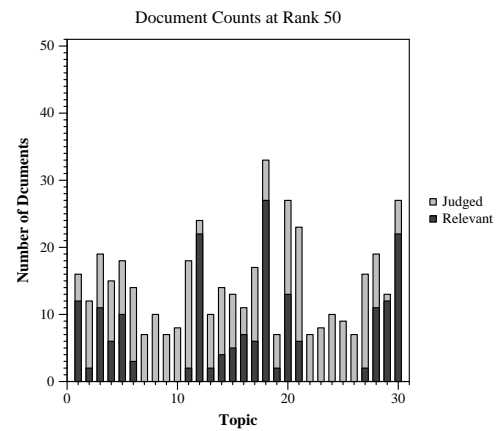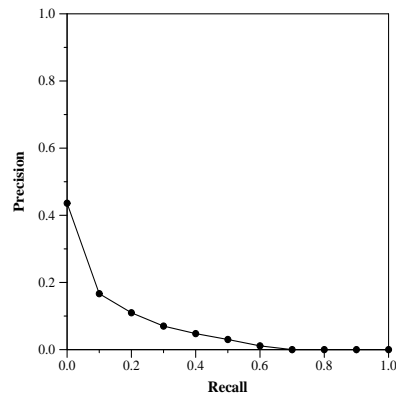


Per-topic difference from median bpref for all Round 1 runs
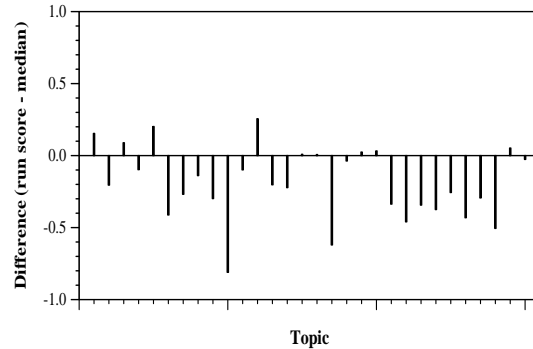


Per-topic difference from median NDCG@10 for all Round 1 runs
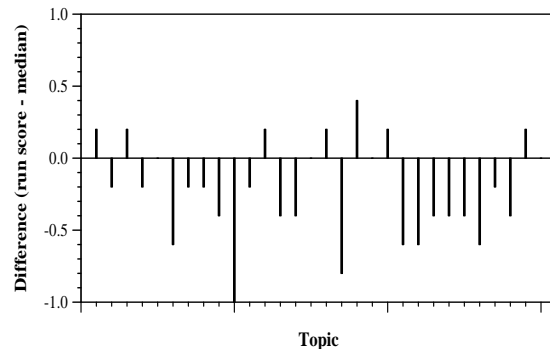


Per-topic difference from median P@5 for all Round 1 runs