

Round 2 results — Run GUIR_S2_run3 submitted from GUIR_S2

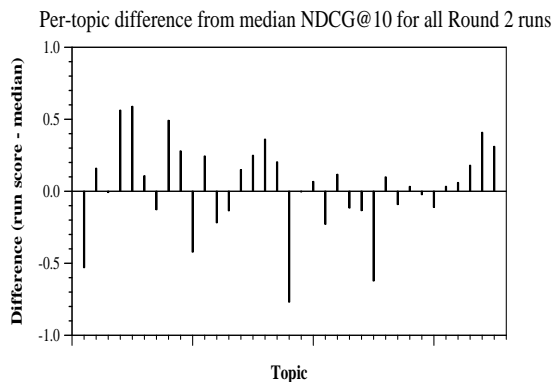
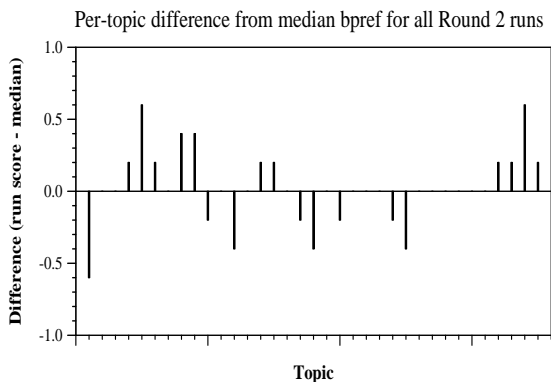
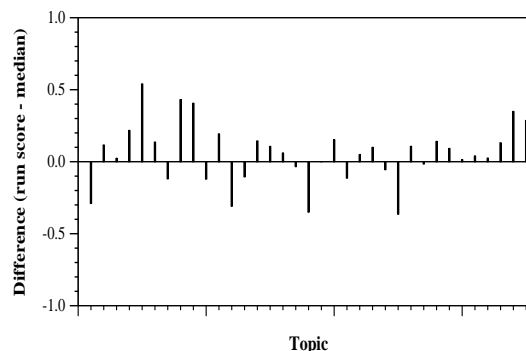
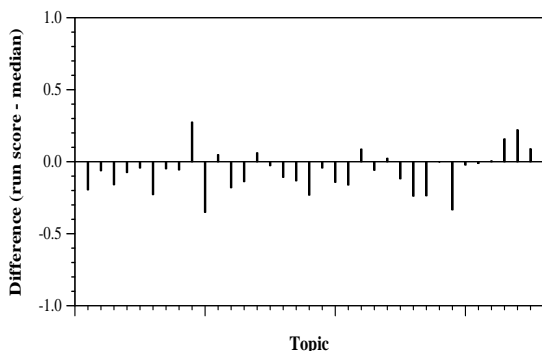
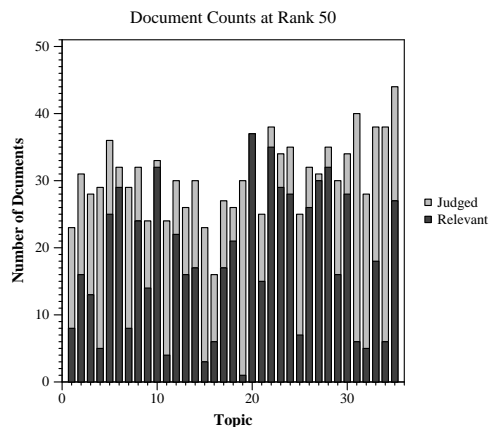
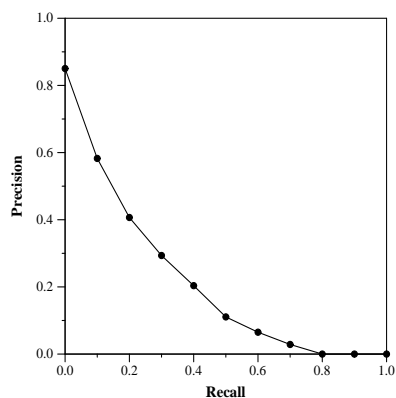
Run Description

SciBERT over paragraphs, trained on medical subset of MS-MARCO. Tuning conducted on judged documents from Round 1.

| Summary Statistics | |
|-------------------------------|--------------|
| Run ID | GUIR_S2_run3 |
| Topic type | feedback |
| Contributed to judgment sets? | no |

| Overall measures | |
|--------------------------|----------------|
| Number of topics | 35 |
| Total number retrieved | 6281 |
| Total relevant | 3002 |
| Total relevant retrieved | 975 |
| MAP | 0.2019 |
| Mean Bpref | 0.3008 |
| Mean NDCG@10 | 0.5286 |
| Mean RBP(p=0.5) | 0.5663 +0.0544 |

| Document Level Averages | |
|-------------------------|-----------|
| | Precision |
| At 5 docs | 0.6229 |
| At 10 docs | 0.5743 |
| At 15 docs | 0.5257 |
| At 20 docs | 0.5029 |
| At 30 docs | 0.4410 |
| R-Precision | |
| Exact | 0.2798 |



Per-topic difference from median P@5 for all Round 2 runs

Per-topic difference from median RBP(p=0.5) for all Round 2 runs