## Run Description

The initial results are retrieved with SLEDGE on the whole collections on history, with BM25 and query over fulltext, re-ranking using SciBERT on title and abstract with the question. Then the results are reranked with an inverse-time-decay mechanism and a time score is used as an additional term for reranking. The time score of documents published before 2010 is assigned as 0 and the subsequent published documents is assigned with a scaled score. The weight between the score from SLEDGE and time term is tuned for better performance based on the evaluation from round 1.

| Summary Statistics | |
| --- | --- |
| Run ID | Emory_IRLab_Run1 |
| Topic type | automatic |
| Contributed to judgment sets? | yes |

| Overall measures | |
| --- | --- |
| Number of topics | 35 |
| Total number retrieved | 15080 |
| Total relevant | 3002 |
| Total relevant retrieved | 1240 |
| MAP | 0.2247 |
| Mean Bpref | 0.3539 |
| Mean NDCG@10 | 0.5832 |
| Mean RBP(p=0.5) | 0.6338 +0.0020 |

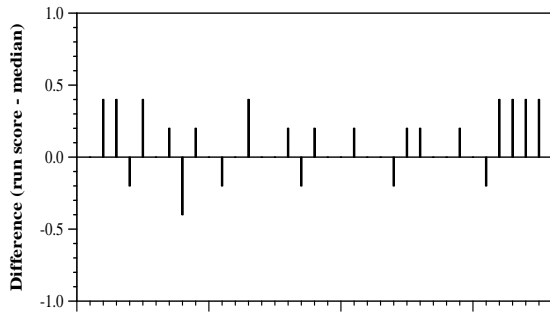| Document Level Averages | |
| --- | --- |
| | Precision |
| At 5 docs | 0.6971 |
| At 10 docs | 0.6286 |
| At 15 docs | 0.5657 |
| At 20 docs | 0.5129 |
| At 30 docs | 0.4429 |
| R-Precision | |
| Exact | 0.2795 |





Document Counts at Rank 50
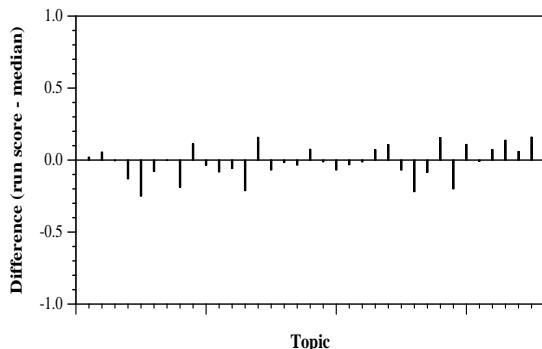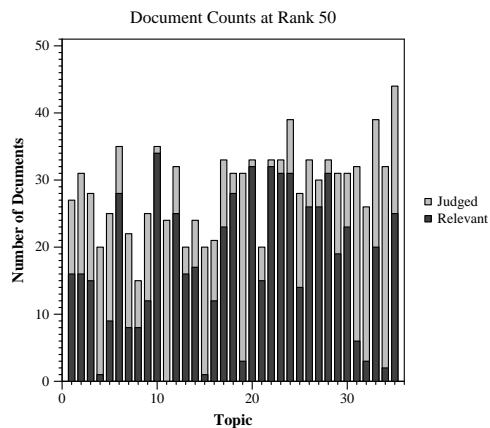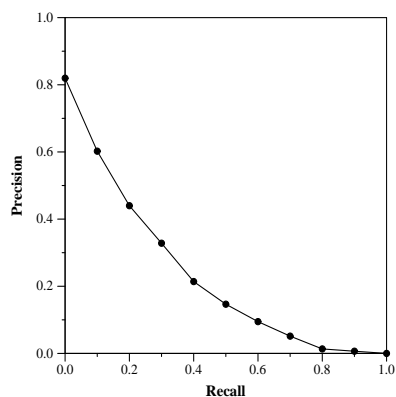


Per-topic difference from median bpref for all Round 2 runs
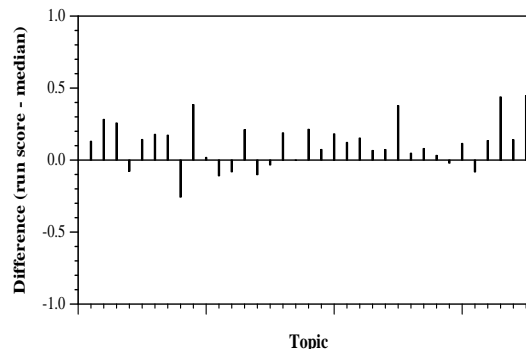


Per-topic difference from median NDCG@10 for all Round 2 runs

Per-topic difference from median P@5 for all Round 2 runs
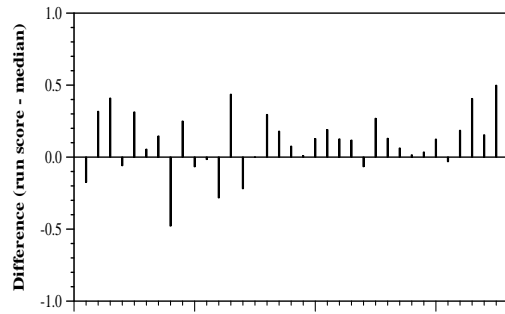
Per-topic difference from median RBP(p=0.5) for all Round 2 runs