

Round 5 results — Run rk_ir_trf_logit_rr submitted from risklick

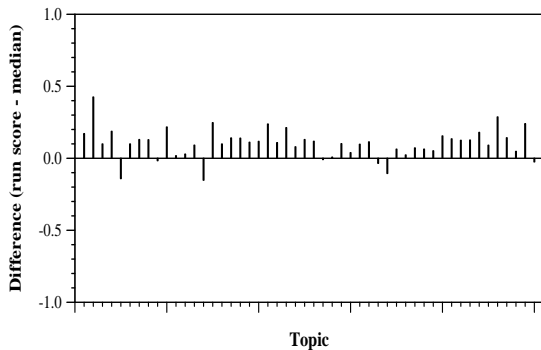
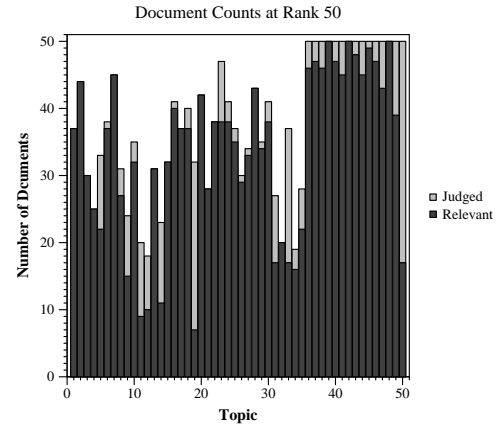
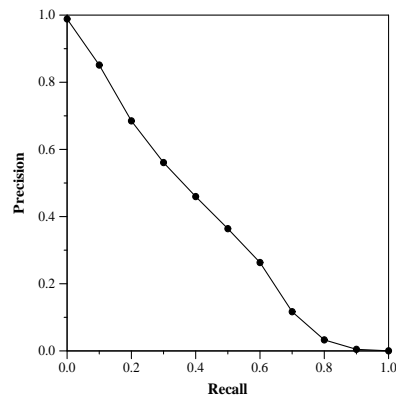
Run Description

A combination of information retrieval (BM25, DFR, LM Dirichlet), neural language (BERT, RoBERTa, XLNet) and logistic regression models trained on round 4 QREL. Reciprocal-rank fusion with a k=60 was used to combined the runs.

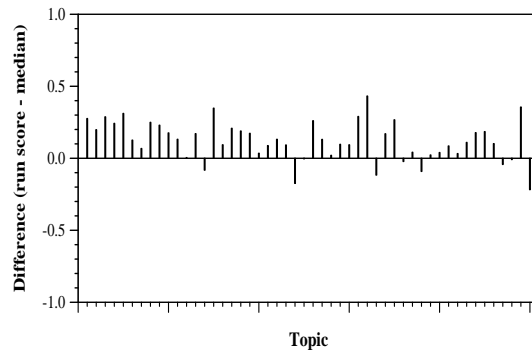
| Summary Statistics | |
|-------------------------------|--------------------|
| Run ID | rk_ir_trf_logit_rr |
| Topic type | feedback |
| Contributed to judgment sets? | yes |

| Overall measures | |
|--------------------------|----------------|
| Number of topics | 50 |
| Total number retrieved | 45594 |
| Total relevant | 10910 |
| Total relevant retrieved | 6939 |
| MAP | 0.3789 |
| Mean Bpref | 0.5659 |
| Mean NDCG@20 | 0.7956 |
| Mean RBP(p=0.5) | 0.8800 +0.0004 |

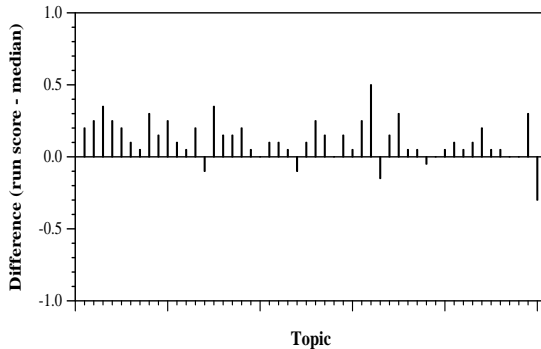
| Document Level Averages | |
|-------------------------|-----------|
| | Precision |
| At 5 docs | 0.9320 |
| At 10 docs | 0.8880 |
| At 15 docs | 0.8547 |
| At 20 docs | 0.8260 |
| At 30 docs | 0.7613 |
| R-Precision | |
| Exact | 0.4209 |



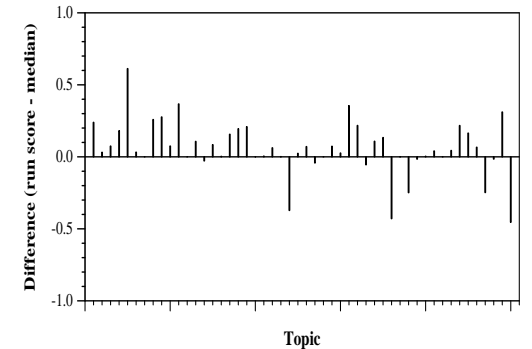
Per-topic difference from median bpref for all Round 4 runs



Per-topic difference from median NDCG@20 for all Round 4 runs



Per-topic difference from median P@20 for all Round 4 runs



Per-topic difference from median RBP(p=0.5) for all Round 4 runs