Robust track results — IBM Research Lab

Summary Statistics

RunID: Juru

Run Description: Automatic run, description only

Kendall of difficulty predictions: 0.276

Topic set: Old 200 New 49 Hard 50 All 249

Total number retrieved: 200,000 49,000 50,000 249,000

Total relevant: 15,350 20,624 16,174 12,124

Total relevant retrieved: 8,845 15,588 11,330 7,142

Mean average precision:
- Old topics: 0.295
- New topics: 0.299
- Hard topics: 0.280
- All topics: 0.275

% topics with no relevant in top 10:
- Old topics: 8.5%
- New topics: 8.2%
- Hard topics: 8.0%
- All topics: 8.4%

Area under MAP curve:
- Exact: 0.0146 0.0331 0.0101 0.0167
- R-Precision: 0.0 0.2 0.4 0.6 0.8 1.0

Recall

Precision

Document Level Averages

- Precision at 5 docs:
  - Old topics: 0.550
  - New topics: 0.543
  - Hard topics: 0.436
  - All topics: 0.549

- Precision at 10 docs:
  - Old topics: 0.484
  - New topics: 0.445
  - Hard topics: 0.368
  - All topics: 0.476

- Precision at 15 docs:
  - Old topics: 0.428
  - New topics: 0.399
  - Hard topics: 0.312
  - All topics: 0.423

- Precision at 20 docs:
  - Old topics: 0.386
  - New topics: 0.373
  - Hard topics: 0.272
  - All topics: 0.383

- Precision at 30 docs:
  - Old topics: 0.331
  - New topics: 0.321
  - Hard topics: 0.243
  - All topics: 0.329

- Precision at 100 docs:
  - Old topics: 0.190
  - New topics: 0.182
  - Hard topics: 0.144
  - All topics: 0.188

- Precision at 200 docs:
  - Old topics: 0.127
  - New topics: 0.119
  - Hard topics: 0.103
  - All topics: 0.125

- Precision at 500 docs:
  - Old topics: 0.072
  - New topics: 0.060
  - Hard topics: 0.062
  - All topics: 0.069

- Precision at 1000 docs:
  - Old topics: 0.044
  - New topics: 0.032
  - Hard topics: 0.040
  - All topics: 0.042

% topics with no relevant in top 10:
- Old topics: 8.5%
- New topics: 8.2%
- Hard topics: 8.0%
- All topics: 8.4%

Difference from median in average precision per topic (sorted by difference)