Robust track results — IBM Research Lab

Summary Statistics

Run ID: JuruDes
Run Description: Automatic run, description only

Kendall of difficulty predictions: 0.459

Topic set
Old: 200
New: 49
Hard: 50
All: 249

Total number retrieved
Old: 20,000
New: 4,900
Hard: 5,000
All: 24,900

Total relevant
Old: 15,350
New: 2,062
Hard: 2,125
All: 17,547

Total relevant retrieved
Old: 8,253
New: 1,146
Hard: 1,146
All: 10,545

Mean average precision
Old: 0.262
New: 0.289
Hard: 0.146
All: 0.268

% topics with no relevant in top 10
Old: 9.5%
New: 12.2%
Hard: 10.0%
All: 10.0%

Area under MAP(X) curve
Old: 0.009
New: 0.022
Hard: 0.007
All: 0.011

Document level averages

<table>
<thead>
<tr>
<th>Precision</th>
<th>Old</th>
<th>New</th>
<th>Hard</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>0.536</td>
<td>0.486</td>
<td>0.412</td>
<td>0.526</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>0.473</td>
<td>0.423</td>
<td>0.348</td>
<td>0.463</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>0.412</td>
<td>0.393</td>
<td>0.293</td>
<td>0.408</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>0.370</td>
<td>0.357</td>
<td>0.259</td>
<td>0.368</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>0.321</td>
<td>0.307</td>
<td>0.230</td>
<td>0.319</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>0.181</td>
<td>0.181</td>
<td>0.137</td>
<td>0.181</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>0.122</td>
<td>0.115</td>
<td>0.097</td>
<td>0.120</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.068</td>
<td>0.057</td>
<td>0.059</td>
<td>0.066</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.042</td>
<td>0.031</td>
<td>0.038</td>
<td>0.040</td>
</tr>
</tbody>
</table>

R-Precision

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 0.295</th>
<th>New: 0.306</th>
<th>Hard: 0.189</th>
<th>All: 0.297</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>0.043</td>
<td>0.044</td>
<td>0.028</td>
<td>0.044</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>0.122</td>
<td>0.123</td>
<td>0.069</td>
<td>0.123</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>0.235</td>
<td>0.236</td>
<td>0.147</td>
<td>0.237</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>0.321</td>
<td>0.321</td>
<td>0.226</td>
<td>0.322</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>0.393</td>
<td>0.393</td>
<td>0.294</td>
<td>0.395</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>0.636</td>
<td>0.637</td>
<td>0.473</td>
<td>0.638</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>0.690</td>
<td>0.691</td>
<td>0.534</td>
<td>0.693</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.735</td>
<td>0.736</td>
<td>0.601</td>
<td>0.738</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.768</td>
<td>0.769</td>
<td>0.653</td>
<td>0.770</td>
</tr>
</tbody>
</table>

% of topics with no relevant in top 10

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 9.5%</th>
<th>New: 12.2%</th>
<th>Hard: 10.0%</th>
<th>All: 10.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>8.1%</td>
<td>11.0%</td>
<td>9.2%</td>
<td>9.5%</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>7.3%</td>
<td>10.1%</td>
<td>8.6%</td>
<td>9.2%</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>6.3%</td>
<td>9.1%</td>
<td>7.5%</td>
<td>8.0%</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>5.5%</td>
<td>8.1%</td>
<td>6.4%</td>
<td>7.0%</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>4.7%</td>
<td>7.1%</td>
<td>5.5%</td>
<td>6.0%</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>1.8%</td>
<td>3.6%</td>
<td>2.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>1.3%</td>
<td>2.5%</td>
<td>1.5%</td>
<td>1.7%</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.7%</td>
<td>1.3%</td>
<td>0.7%</td>
<td>0.8%</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.6%</td>
<td>1.1%</td>
<td>0.6%</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

R-Precision

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 0.295</th>
<th>New: 0.306</th>
<th>Hard: 0.189</th>
<th>All: 0.297</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>0.043</td>
<td>0.044</td>
<td>0.028</td>
<td>0.044</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>0.122</td>
<td>0.123</td>
<td>0.069</td>
<td>0.123</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>0.235</td>
<td>0.236</td>
<td>0.147</td>
<td>0.237</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>0.321</td>
<td>0.321</td>
<td>0.226</td>
<td>0.322</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>0.393</td>
<td>0.393</td>
<td>0.294</td>
<td>0.395</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>0.636</td>
<td>0.637</td>
<td>0.473</td>
<td>0.638</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>0.690</td>
<td>0.691</td>
<td>0.534</td>
<td>0.693</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.735</td>
<td>0.736</td>
<td>0.601</td>
<td>0.738</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.768</td>
<td>0.769</td>
<td>0.653</td>
<td>0.770</td>
</tr>
</tbody>
</table>

% of topics with no relevant in top 10

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 9.5%</th>
<th>New: 12.2%</th>
<th>Hard: 10.0%</th>
<th>All: 10.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>8.1%</td>
<td>11.0%</td>
<td>9.2%</td>
<td>9.5%</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>7.3%</td>
<td>10.1%</td>
<td>8.6%</td>
<td>9.2%</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>6.3%</td>
<td>9.1%</td>
<td>7.5%</td>
<td>8.0%</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>5.5%</td>
<td>8.1%</td>
<td>6.4%</td>
<td>7.0%</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>4.7%</td>
<td>7.1%</td>
<td>5.5%</td>
<td>6.0%</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>1.8%</td>
<td>3.6%</td>
<td>2.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>1.3%</td>
<td>2.5%</td>
<td>1.5%</td>
<td>1.7%</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.7%</td>
<td>1.3%</td>
<td>0.7%</td>
<td>0.8%</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.6%</td>
<td>1.1%</td>
<td>0.6%</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

R-Precision

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 0.295</th>
<th>New: 0.306</th>
<th>Hard: 0.189</th>
<th>All: 0.297</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>0.043</td>
<td>0.044</td>
<td>0.028</td>
<td>0.044</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>0.122</td>
<td>0.123</td>
<td>0.069</td>
<td>0.123</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>0.235</td>
<td>0.236</td>
<td>0.147</td>
<td>0.237</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>0.321</td>
<td>0.321</td>
<td>0.226</td>
<td>0.322</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>0.393</td>
<td>0.393</td>
<td>0.294</td>
<td>0.395</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>0.636</td>
<td>0.637</td>
<td>0.473</td>
<td>0.638</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>0.690</td>
<td>0.691</td>
<td>0.534</td>
<td>0.693</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.735</td>
<td>0.736</td>
<td>0.601</td>
<td>0.738</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.768</td>
<td>0.769</td>
<td>0.653</td>
<td>0.770</td>
</tr>
</tbody>
</table>

% of topics with no relevant in top 10

<table>
<thead>
<tr>
<th>Exact</th>
<th>Old: 9.5%</th>
<th>New: 12.2%</th>
<th>Hard: 10.0%</th>
<th>All: 10.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 5 docs</td>
<td>8.1%</td>
<td>11.0%</td>
<td>9.2%</td>
<td>9.5%</td>
</tr>
<tr>
<td>At 10 docs</td>
<td>7.3%</td>
<td>10.1%</td>
<td>8.6%</td>
<td>9.2%</td>
</tr>
<tr>
<td>At 15 docs</td>
<td>6.3%</td>
<td>9.1%</td>
<td>7.5%</td>
<td>8.0%</td>
</tr>
<tr>
<td>At 20 docs</td>
<td>5.5%</td>
<td>8.1%</td>
<td>6.4%</td>
<td>7.0%</td>
</tr>
<tr>
<td>At 30 docs</td>
<td>4.7%</td>
<td>7.1%</td>
<td>5.5%</td>
<td>6.0%</td>
</tr>
<tr>
<td>At 100 docs</td>
<td>1.8%</td>
<td>3.6%</td>
<td>2.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td>At 200 docs</td>
<td>1.3%</td>
<td>2.5%</td>
<td>1.5%</td>
<td>1.7%</td>
</tr>
<tr>
<td>At 500 docs</td>
<td>0.7%</td>
<td>1.3%</td>
<td>0.7%</td>
<td>0.8%</td>
</tr>
<tr>
<td>At 1000 docs</td>
<td>0.6%</td>
<td>1.1%</td>
<td>0.6%</td>
<td>0.7%</td>
</tr>
</tbody>
</table>