

TREC 2013 Crowdsourcing Track Results

Group: (udel) University of Delaware

Run ID: udelJudge2

Task: Standard

Run type: Secondary

Description of run:

Fully automated run using 3 search engines on the pool of docs. Rules based on the position of a document for a topic in each of the 3 systems' ranking determines whether a document is relevant or not. Each udelJudgeX run has rules slightly different from the other udelJudgeY runs.

Results

Evaluation of crowd qrel **udelJudge2** for the Basic task of the TREC 2013 Crowdsourcing Track. The 10 topics for the basic task were randomly selected from the TREC 2013 web track ad-hoc task.

Topic	#Docs	GAP	τ_{AP} (APCorr)	RMSE
202	231	0.333	0.011	0.436
214	305	0.594	-0.246	0.252
216	387	0.578	0.049	0.289
221	368	0.499	-0.003	0.233
227	246	0.112	-0.174	0.386
230	172	0.207	-0.297	0.331
234	298	0.578	0.134	0.291
243	342	0.237	0.224	0.293
246	202	0.357	0.011	0.292
250	207	0.064	-0.270	0.325
all	2758	0.356	-0.227	0.248

Table 1: This table shows per-topic statistics and overall averages for the run udelJudge2. The metrics GAP, ERR@20, AP-correlation and RMSE are listed for each topic. Note that for row *all*, (i) GAP is the mean gap over all 10 topics, (ii) APCorr and RMSE depend on the ranking of runs induced by the mean ERR@20 for *all* the 10 topics.

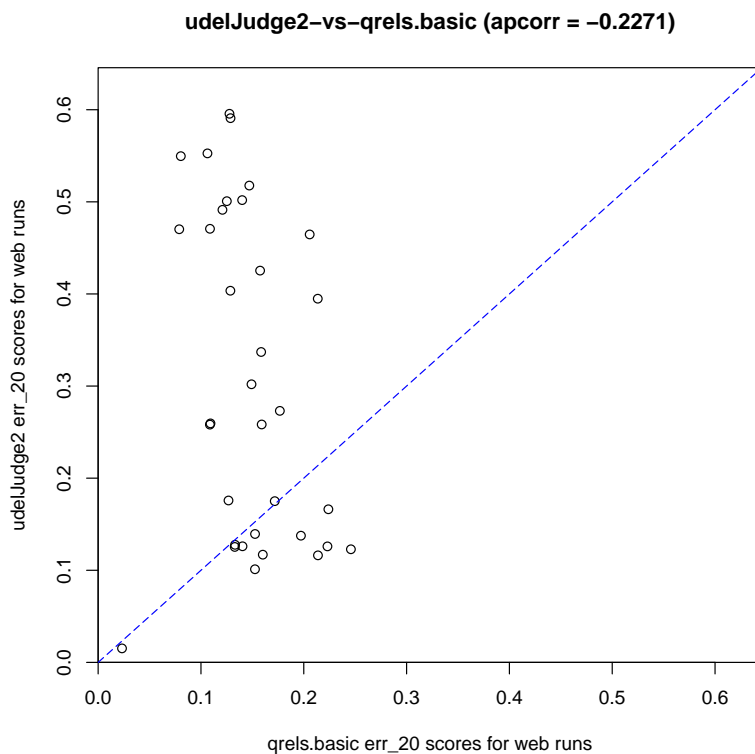


Figure 1: udelJudge2-basic-ERR@20 vs qrels.basic-ERR@20. *qrels.basic* is the TREC 2013 web track qrels reduced to topics 202, 214, 216, 221, 227, 230, 234, 243, 246, and 250.