

TREC 2013 Crowdsourcing Track Results

Group: (udel) University of Delaware

Run ID: udelJudge4

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 **udelJudge4** 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	0.091	-0.018	0.373
214	305	0.592	-0.226	0.145
216	387	0.592	0.143	0.254
221	368	0.493	-0.036	0.172
227	246	0.113	-0.278	0.466
230	172	0.218	-0.243	0.441
234	298	0.603	0.140	0.236
243	342	0.235	0.394	0.379
246	202	0.363	0.193	0.376
250	207	0.070	-0.135	0.234
all	2758	0.337	-0.221	0.221

Table 1: This table shows per-topic statistics and overall averages for the run udelJudge4. 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.

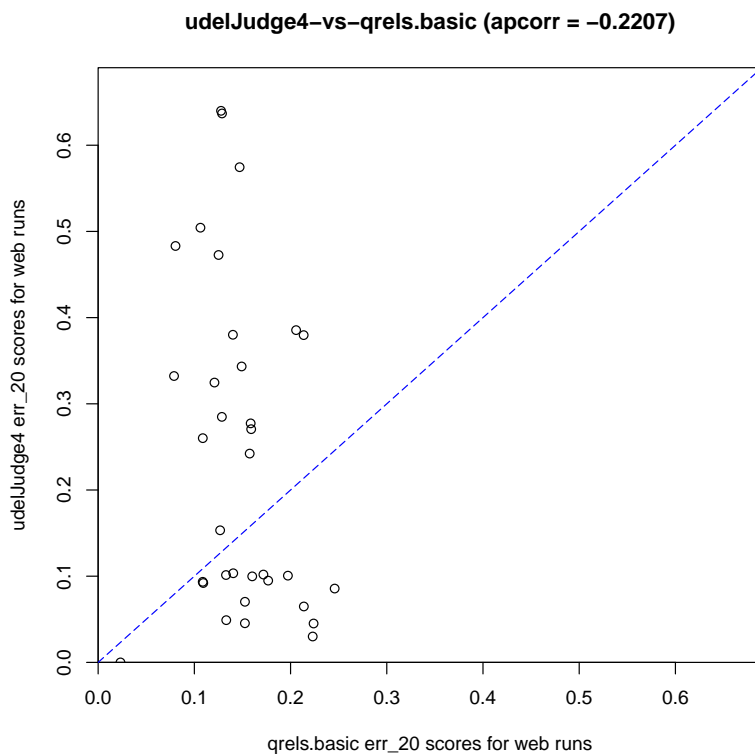


Figure 1: udelJudge4-basic-ERR@20 vs qrels.basic-ERR@20. *qrels.basic* is the TREC 2013 web track qrels reduced to topics 202, 214, 216, 221, 227, 230, 234, 243, 246, and 250.