

TREC 2013 Crowdsourcing Track Results

Group: (udel) University of Delaware

Run ID: udelJudge6

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 **udelJudge6** 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	0.036	-0.019	0.410
214	305	0.630	0.163	0.174
216	387	0.571	-0.073	0.228
221	368	0.545	0.223	0.254
227	246	0.114	-0.223	0.380
230	172	0.267	-0.078	0.263
234	298	0.580	0.045	0.275
243	342	0.228	0.256	0.277
246	202	0.444	0.291	0.260
250	207	0.114	-0.334	0.355
all	2758	0.353	-0.173	0.233

Table 1: This table shows per-topic statistics and overall averages for the run udelJudge6. 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.

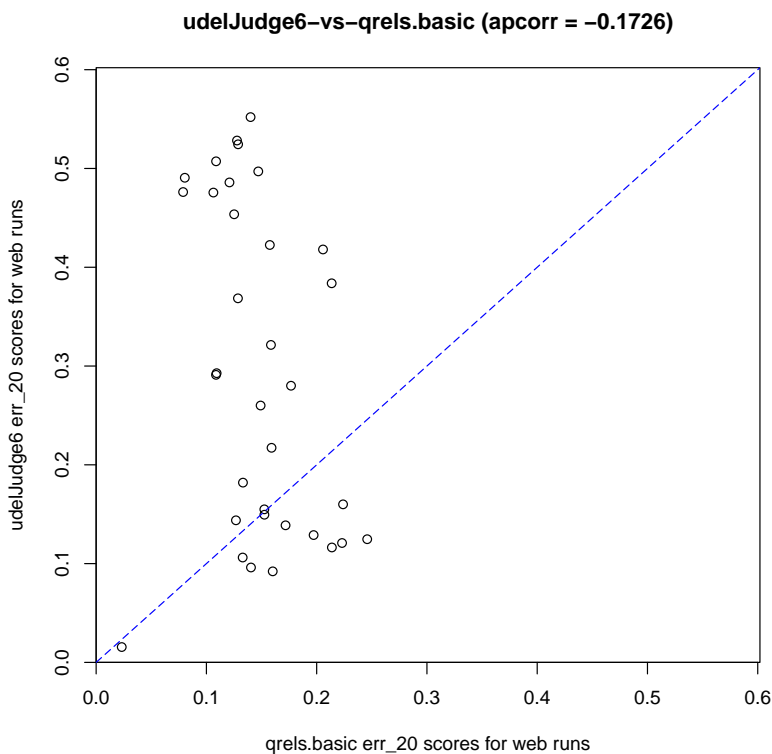


Figure 1: udelJudge6-basic-ERR@20 vs qrels.basic-ERR@20. *qrels.basic* is the TREC 2013 web track qrels reduced to topics 202, 214, 216, 221, 227, 230, 234, 243, 246, and 250.