

TREC 2013 Crowdsourcing Track Results

Group: (udel) University of Delaware

Run ID: udelJudge3

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 **udelJudge3** 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	0.017	0.608
214	305	0.601	-0.183	0.274
216	387	0.588	0.193	0.424
221	368	0.489	0.015	0.250
227	246	0.113	-0.286	0.487
230	172	0.209	-0.260	0.465
234	298	0.604	0.155	0.433
243	342	0.255	0.314	0.440
246	202	0.364	0.022	0.463
250	207	0.067	-0.201	0.361
all	2758	0.338	-0.173	0.330

Table 1: This table shows per-topic statistics and overall averages for the run udelJudge3. 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.

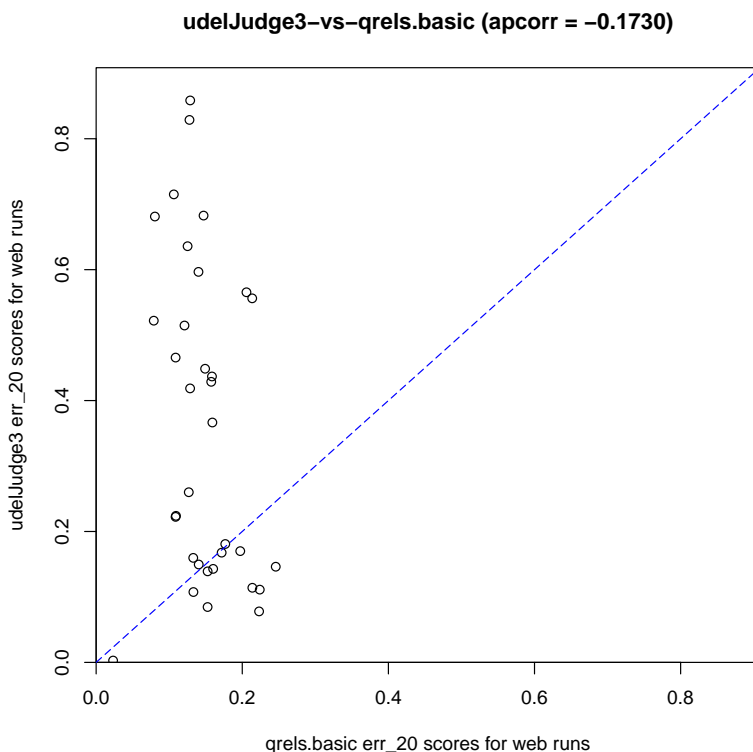


Figure 1: udelJudge3-basic-ERR@20 vs qrels.basic-ERR@20. *qrels.basic* is the TREC 2013 web track qrels reduced to topics 202, 214, 216, 221, 227, 230, 234, 243, 246, and 250.