

TREC 2012 Crowdsourcing Track, Text Relevance Assessing Task (TRAT) results

Group: (york) York University

Run ID: yorku12cs02

Run type: Secondary

Description of run:

Judgements based on fusion with extracted expansion words from crowdFlower.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	15	1810	219	12	0.554	0.892	0.108	0.446	0.238	0.407
416	1235	45	27	973	217	18	0.598	0.817	0.183	0.402	0.279	0.398
417	2992	75	4	2691	226	71	0.059	0.922	0.078	0.941	0.536	0.500
420	1136	37	11	820	279	26	0.303	0.746	0.254	0.697	0.470	0.443
427	1528	37	22	1264	227	15	0.592	0.848	0.152	0.408	0.260	0.363
432	2503	22	4	2324	157	18	0.196	0.937	0.063	0.804	0.345	0.431
438	1798	162	23	1502	134	139	0.144	0.918	0.082	0.856	0.422	0.448
445	1404	60	18	1096	248	42	0.303	0.815	0.185	0.697	0.419	0.403
446	2020	156	18	1714	150	138	0.118	0.919	0.081	0.882	0.448	0.497
447	1588	16	14	1512	60	2	0.853	0.962	0.038	0.147	0.077	0.486
Average	1826.000	63.700	15.600	1570.600	191.700	48.100	0.372	0.878	0.122	0.628	0.349	0.438

Table 1: This table shows per-topic statistics and overall averages for the run yorku12cs02. The topics are 10 randomly selected topics from the TREC 8 ad-hoc task. A relevant document is positive and a non-relevant document is negative. The true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts are based on an adjudicated set of relevance judgments that differs from the original TREC-8 ad-hoc qrels. The true positive rate (TPR), false positive rate (FPR), true negative rate (TNR), and the false negative rate (FNR) are all smoothed values. Details of the computation of the logistic average misclassification (LAM) rate and the area under the curve (AUC) are given in the track overview paper. Some runs did not report a probability of relevance and thus will have NA for their AUC score.

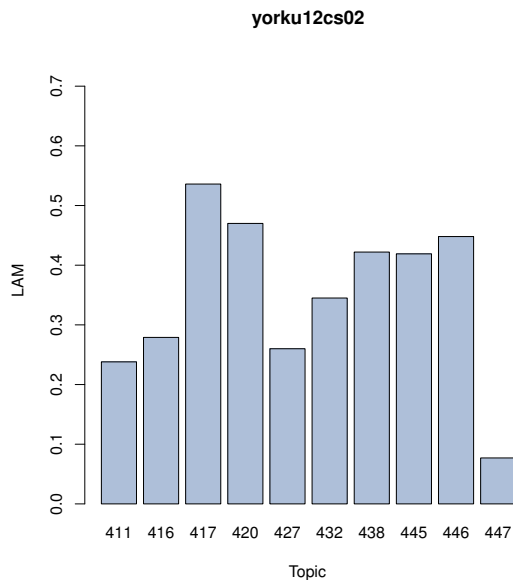


Figure 1: yorku12cs02 LAM

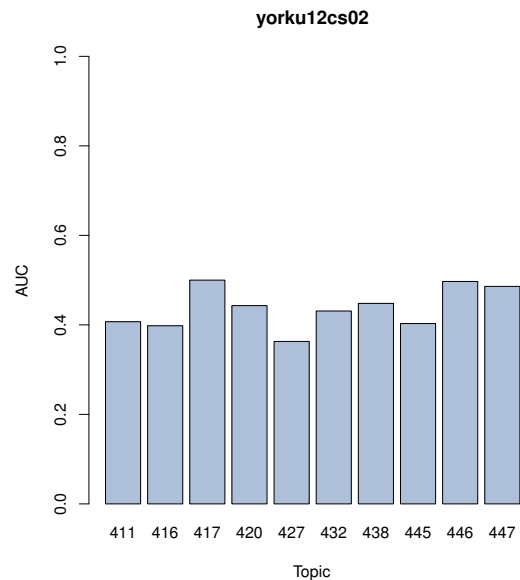


Figure 2: yorku12cs02 AUC