

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

Group: (york) York University

Run ID: yorku12cs04

Run type: Secondary

Description of run:

Judgements based on fusion with expansion words from pseudo feedback and partially from crowdFlower.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	21	1754	275	6	0.768	0.864	0.136	0.232	0.179	0.459
416	1235	45	34	958	232	11	0.750	0.805	0.195	0.250	0.221	0.480
417	2992	75	41	2716	201	34	0.546	0.931	0.069	0.454	0.199	0.429
420	1136	37	21	796	303	16	0.566	0.724	0.276	0.434	0.351	0.476
427	1528	37	31	1200	291	6	0.829	0.805	0.195	0.171	0.183	0.415
432	2503	22	6	2198	283	16	0.283	0.886	0.114	0.717	0.364	0.454
438	1798	162	103	1315	321	59	0.635	0.804	0.196	0.365	0.273	0.457
445	1404	60	51	1003	341	9	0.844	0.746	0.254	0.156	0.200	0.480
446	2020	156	106	1634	230	50	0.678	0.876	0.124	0.322	0.205	0.478
447	1588	16	15	1381	191	1	0.912	0.878	0.122	0.088	0.104	NA
Average	1826.000	63.700	42.900	1495.500	266.800	20.800	0.681	0.832	0.168	0.319	0.228	0.465

Table 1: This table shows per-topic statistics and overall averages for the run yorku12cs04. 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.

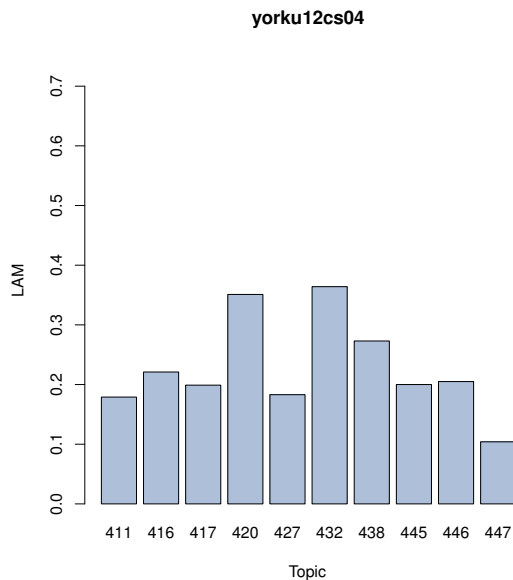


Figure 1: yorku12cs04 LAM

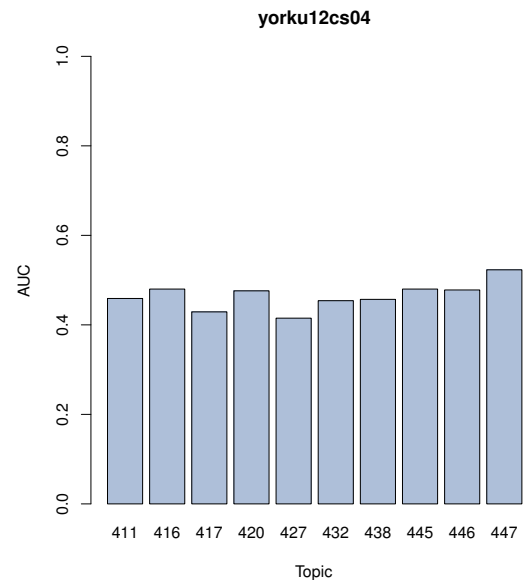


Figure 2: yorku12cs04 AUC