

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

Group: (york) York University

Run ID: yorku12cs01

Run type: Secondary

Description of run:

Judgements partially based on fusion with TF and partially from CrowdFlower.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	20	1748	281	7	0.732	0.861	0.139	0.268	0.195	0.500
416	1235	45	32	968	222	13	0.707	0.813	0.187	0.293	0.236	0.471
417	2992	75	23	2738	179	52	0.309	0.938	0.062	0.691	0.277	0.489
420	1136	37	10	833	266	27	0.276	0.758	0.242	0.724	0.478	0.462
427	1528	37	29	1250	241	8	0.776	0.838	0.162	0.224	0.191	0.402
432	2503	22	1	2354	127	21	0.065	0.949	0.051	0.935	0.468	0.470
438	1798	162	96	1338	298	66	0.592	0.818	0.182	0.408	0.282	0.443
445	1404	60	21	1050	294	39	0.352	0.781	0.219	0.648	0.418	0.446
446	2020	156	11	1729	135	145	0.073	0.927	0.073	0.927	0.499	0.489
447	1588	16	16	1368	204	0	0.971	0.870	0.130	0.029	0.063	0.498
Average	1826.000	63.700	25.900	1537.600	224.700	37.800	0.485	0.855	0.145	0.515	0.311	0.467

Table 1: This table shows per-topic statistics and overall averages for the run yorku12cs01. 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.

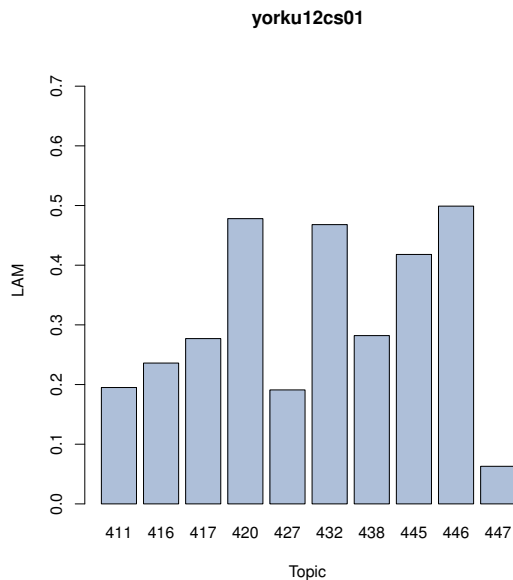


Figure 1: yorku12cs01 LAM

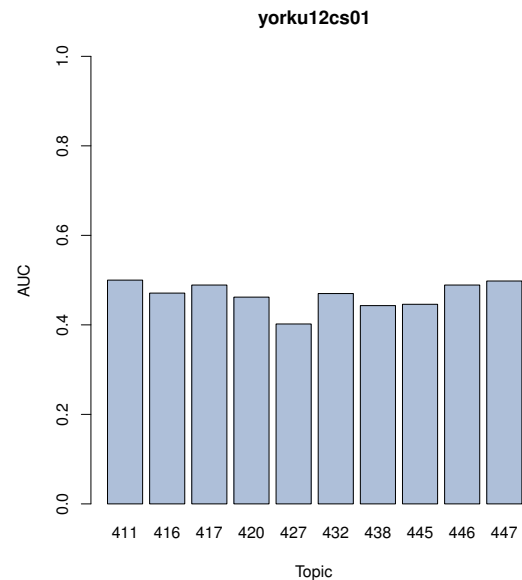


Figure 2: yorku12cs01 AUC