

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

Group: (york) York University

Run ID: yorku12cs03

Run type: Primary

Description of run:

Judgements based on fusion with expansion words from term proximity.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	19	1757	272	8	0.696	0.866	0.134	0.304	0.206	0.474
416	1235	45	37	964	226	8	0.815	0.810	0.190	0.185	0.187	0.465
417	2992	75	48	2730	187	27	0.638	0.936	0.064	0.362	0.165	0.451
420	1136	37	24	754	345	13	0.645	0.686	0.314	0.355	0.334	0.432
427	1528	37	31	1196	295	6	0.829	0.802	0.198	0.171	0.184	0.464
432	2503	22	6	2159	322	16	0.283	0.870	0.130	0.717	0.381	0.503
438	1798	162	116	1295	341	46	0.715	0.791	0.209	0.285	0.245	0.463
445	1404	60	52	998	346	8	0.861	0.742	0.258	0.139	0.192	0.509
446	2020	156	111	1676	188	45	0.710	0.899	0.101	0.290	0.176	0.491
447	1588	16	15	1356	216	1	0.912	0.862	0.138	0.088	0.111	0.534
Average	1826.000	63.700	45.900	1488.500	273.800	17.800	0.710	0.826	0.174	0.290	0.218	0.479

Table 1: This table shows per-topic statistics and overall averages for the run yorku12cs03. 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.

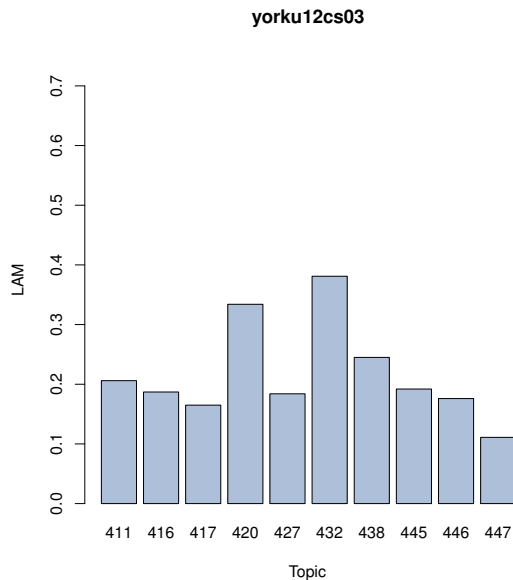


Figure 1: yorku12cs03 LAM

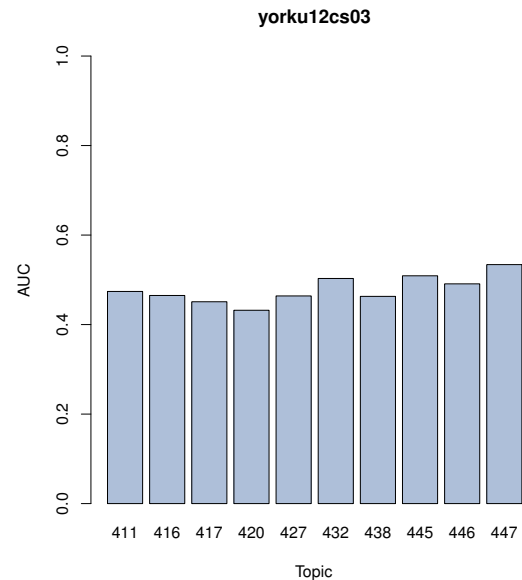


Figure 2: yorku12cs03 AUC