

TREC 2012 Crowdsourcing Track, Text Relevance Assessing Task (TRAT) results
Group: (pris411) PRIS Lab at Beijing University of Posts and Telecommunications
Run ID: BUPTPRISZHS

Run type: Primary

Description of run:

This is our first run. Results are obtained on the CrowdFlower platform.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	4	2015	14	23	0.161	0.993	0.007	0.839	0.162	NA
416	1235	45	11	1160	30	34	0.250	0.974	0.026	0.750	0.219	NA
417	2992	75	10	2863	54	65	0.138	0.981	0.019	0.862	0.256	NA
420	1136	37	11	1086	13	26	0.303	0.988	0.012	0.697	0.145	NA
427	1528	37	3	1454	37	34	0.092	0.975	0.025	0.908	0.335	NA
432	2503	22	1	2471	10	21	0.065	0.996	0.004	0.935	0.198	NA
438	1798	162	19	1580	56	143	0.120	0.965	0.035	0.880	0.339	NA
445	1404	60	1	1334	10	59	0.025	0.992	0.008	0.975	0.358	NA
446	2020	156	44	1795	69	112	0.283	0.963	0.037	0.717	0.238	NA
447	1588	16	11	1535	37	5	0.676	0.976	0.024	0.324	0.098	NA
Average	1826.000	63.700	11.500	1729.300	33.000	52.200	0.211	0.980	0.020	0.789	0.235	NA

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

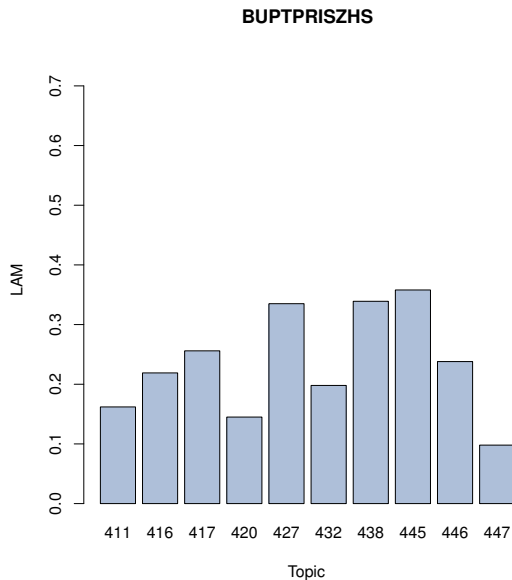


Figure 1: BUPTPRISZHS LAM