

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

Group: (UIowaS) The University of Iowa

Run ID: UIowaS03r

Run type: Secondary

Description of run:

Used k-means text clustering. Each cluster ranked by an average of each document's weighted score. Sent to the crowd in ranked order by cluster, then by a random sample of documents within the cluster.

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	20	2023	6	7	0.732	0.997	0.003	0.268	0.033	NA
416	1235	45	42	1179	11	3	0.924	0.990	0.010	0.076	0.028	NA
417	2992	75	55	2908	9	20	0.730	0.997	0.003	0.270	0.034	NA
420	1136	37	26	1084	15	11	0.697	0.986	0.014	0.303	0.073	NA
427	1528	37	29	1469	22	8	0.776	0.985	0.015	0.224	0.062	NA
432	2503	22	2	2465	16	20	0.109	0.993	0.007	0.891	0.190	NA
438	1798	162	70	1622	14	92	0.433	0.991	0.009	0.567	0.098	NA
445	1404	60	39	1337	7	21	0.648	0.994	0.006	0.352	0.052	NA
446	2020	156	75	1853	11	81	0.481	0.994	0.006	0.519	0.076	NA
447	1588	16	16	1562	10	0	0.971	0.993	0.007	0.029	0.014	NA
Average	1826.000	63.700	37.400	1750.200	12.100	26.300	0.650	0.992	0.008	0.350	0.066	NA

Table 1: This table shows per-topic statistics and overall averages for the run UIowaS03r. 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.

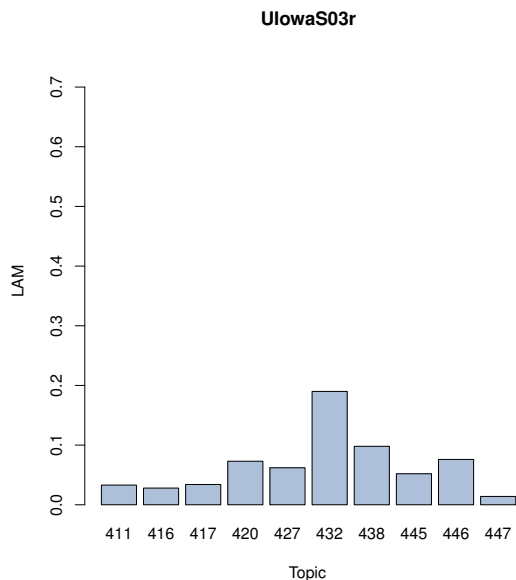


Figure 1: UIowaS03r LAM