

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

Group: (SetuServ) SetuServ

Run ID: SSECML75pct

Run type: Secondary

Description of run:

Post Error Correction (likely errors are identified using M/L for document with confidence score of >75%), counting “not sures” as relevant documents

Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	22	1973	56	5	0.804	0.972	0.028	0.196	0.077	0.949
416	1235	45	41	1178	12	4	0.902	0.990	0.010	0.098	0.033	0.965
417	2992	75	60	2808	109	15	0.796	0.962	0.038	0.204	0.091	0.891
420	1136	37	34	1051	48	3	0.908	0.956	0.044	0.092	0.064	0.961
427	1528	37	30	1478	13	7	0.803	0.991	0.009	0.197	0.045	0.942
432	2503	22	9	2418	63	13	0.413	0.974	0.026	0.587	0.162	0.714
438	1798	162	149	1503	133	13	0.917	0.918	0.082	0.083	0.082	0.940
445	1404	60	49	1325	19	11	0.811	0.986	0.014	0.189	0.055	0.918
446	2020	156	93	1796	68	63	0.596	0.963	0.037	0.404	0.139	0.821
447	1588	16	15	1571	1	1	0.912	0.999	0.001	0.088	0.010	0.969
Average	1826.000	63.700	50.200	1710.100	52.200	13.500	0.786	0.971	0.029	0.214	0.076	0.907

Table 1: This table shows per-topic statistics and overall averages for the run SSECML75pct. 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.

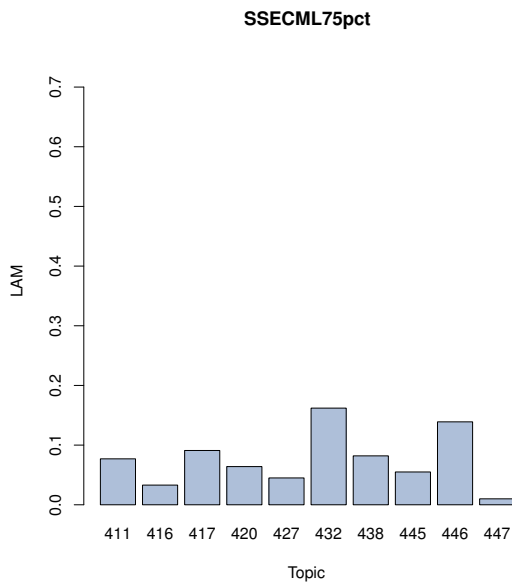


Figure 1: SSECML75pct LAM

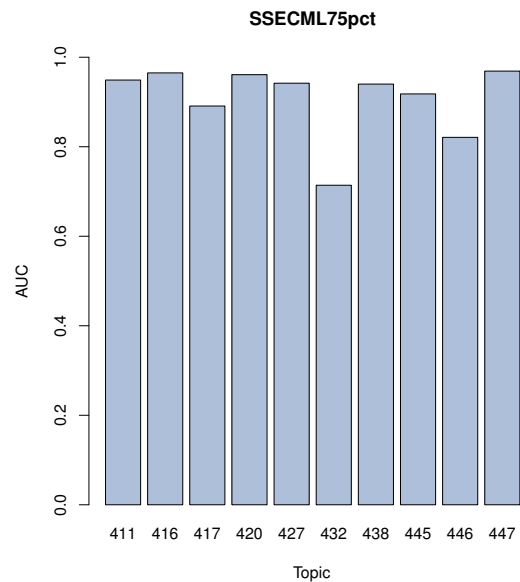


Figure 2: SSECML75pct AUC