

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

Group: (SetuServ) SetuServ

Run ID: SSNoEC

Run type: Secondary

Description of run:

Crowdsourced using Skierarchy approach for all pairs (no error correction & no plurality)

## Results

Topic	#Docs	#Rel	TP	TN	FP	FN	TPR	TNR	FPR	FNR	LAM	AUC
411	2056	27	24	1967	62	3	0.875	0.969	0.031	0.125	0.063	0.951
416	1235	45	41	1177	13	4	0.902	0.989	0.011	0.098	0.034	0.953
417	2992	75	59	2813	104	16	0.783	0.964	0.036	0.217	0.092	0.885
420	1136	37	34	1055	44	3	0.908	0.960	0.040	0.092	0.061	0.961
427	1528	37	29	1476	15	8	0.776	0.990	0.010	0.224	0.052	0.928
432	2503	22	9	2428	53	13	0.413	0.978	0.022	0.587	0.150	0.716
438	1798	162	147	1504	132	15	0.905	0.919	0.081	0.095	0.088	0.934
445	1404	60	47	1325	19	13	0.779	0.986	0.014	0.221	0.061	0.901
446	2020	156	91	1798	66	65	0.583	0.964	0.036	0.417	0.140	0.811
447	1588	16	14	1571	1	2	0.853	0.999	0.001	0.147	0.013	0.969
Average	1826.000	63.700	49.500	1711.400	50.900	14.200	0.778	0.972	0.028	0.222	0.075	0.901

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

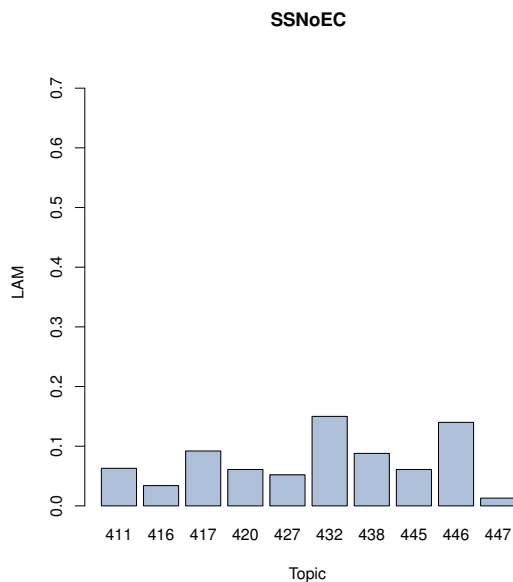


Figure 1: SSNoEC LAM

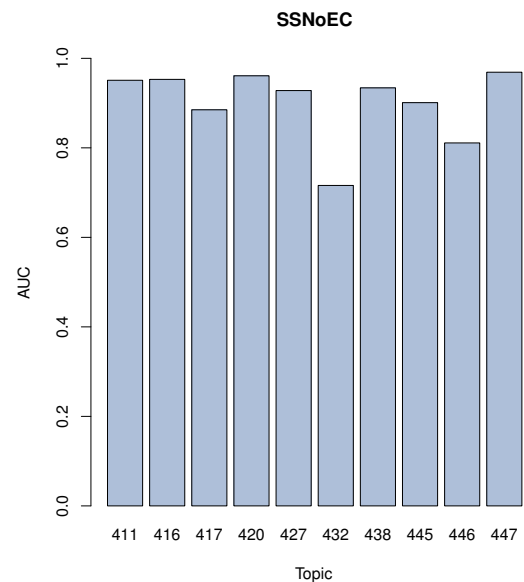


Figure 2: SSNoEC AUC