

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

Group: (UIowaS) The University of Iowa

Run ID: UIowaS02r

Run type: Primary

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 first by cluster, then by order 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 | 43 | 1179 | 11 | 2 | 0.946 | 0.990 | 0.010 | 0.054 | 0.023 | NA |
| 417 | 2992 | 75 | 62 | 2907 | 10 | 13 | 0.822 | 0.996 | 0.004 | 0.178 | 0.027 | NA |
| 420 | 1136 | 37 | 31 | 1077 | 22 | 6 | 0.829 | 0.980 | 0.020 | 0.171 | 0.062 | NA |
| 427 | 1528 | 37 | 32 | 1469 | 22 | 5 | 0.855 | 0.985 | 0.015 | 0.145 | 0.048 | NA |
| 432 | 2503 | 22 | 8 | 2463 | 18 | 14 | 0.370 | 0.993 | 0.007 | 0.630 | 0.102 | NA |
| 438 | 1798 | 162 | 109 | 1617 | 19 | 53 | 0.672 | 0.988 | 0.012 | 0.328 | 0.071 | NA |
| 445 | 1404 | 60 | 53 | 1335 | 9 | 7 | 0.877 | 0.993 | 0.007 | 0.123 | 0.031 | NA |
| 446 | 2020 | 156 | 101 | 1850 | 14 | 55 | 0.646 | 0.992 | 0.008 | 0.354 | 0.061 | NA |
| 447 | 1588 | 16 | 16 | 1561 | 11 | 0 | 0.971 | 0.993 | 0.007 | 0.029 | 0.015 | NA |
| Average | 1826.000 | 63.700 | 47.500 | 1748.100 | 14.200 | 16.200 | 0.772 | 0.991 | 0.009 | 0.228 | 0.047 | NA |

Table 1: This table shows per-topic statistics and overall averages for the run UIowaS02r. 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.

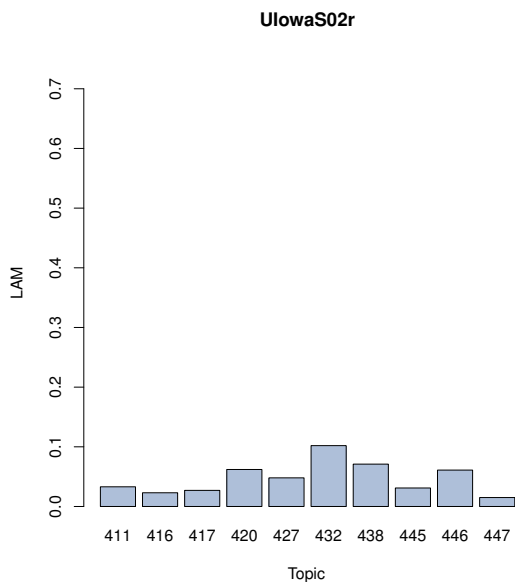


Figure 1: UIowaS02r LAM