

## Round 4 results — Run active\_learning submitted from risklick

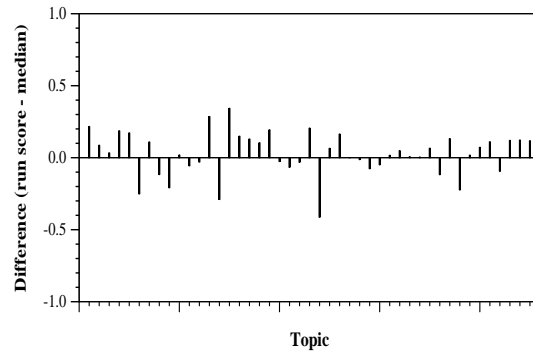
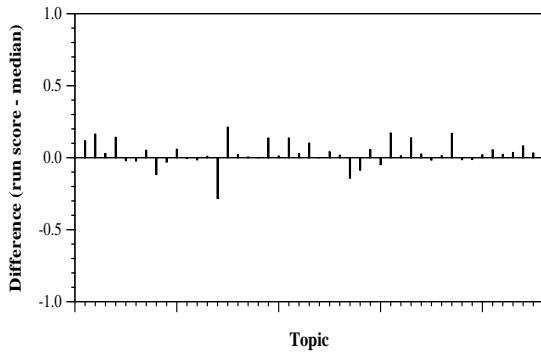
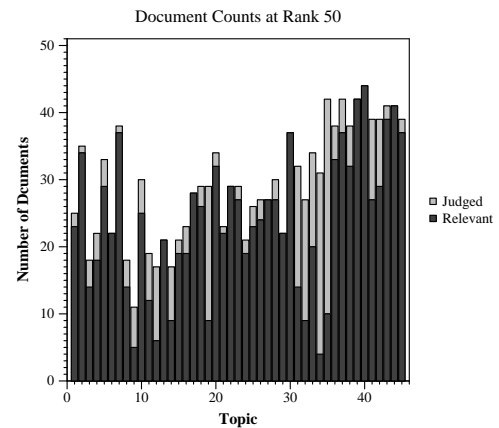
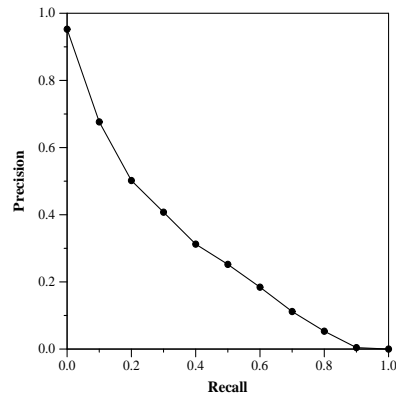
### Run Description

Manually judging retrieved publications return by a basic IR model

Summary Statistics	
Run ID	active_learning
Topic type	manual
Contributed to judgment sets?	yes

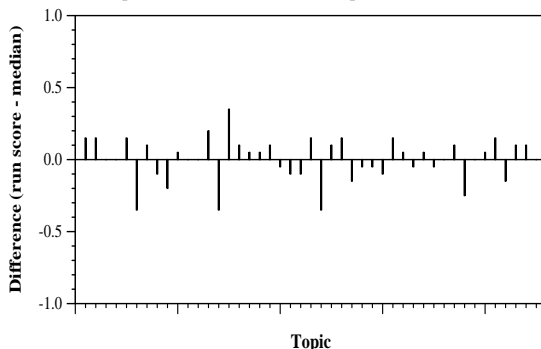
Overall measures	
Number of topics	45
Total number retrieved	39716
Total relevant	5824
Total relevant retrieved	3612
MAP	0.2946
Mean Bpref	0.5464
Mean NDCG@20	0.6551
Mean RBP(p=0.5)	0.8056 +0.0244

Document Level Averages	
	Precision
At 5 docs	0.8578
At 10 docs	0.8000
At 15 docs	0.7052
At 20 docs	0.6533
At 30 docs	0.5748
R-Precision	
Exact	0.3414

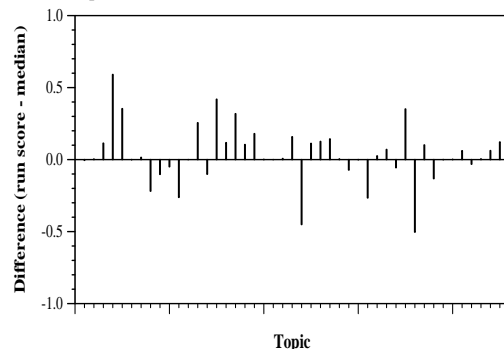


Per-topic difference from median bpref for all Round 4 runs

Per-topic difference from median NDCG@20 for all Round 4 runs



Per-topic difference from median P@20 for all Round 4 runs



Per-topic difference from median RBP(p=0.5) for all Round 4 runs