

Programming Assignment 2

(Team 3)

Team Members:

	Names	UNH ID	Webcat Username	Assignment No
1.	Amith Ramanagar Chandrashekar	928101607	ar1184	2
2.	Vaughan Coder	966427992	vwc1010	2
3.	Sepideh Koohfar	928334858	sk1015	2
4.	Pooja Himanshu Oza	900539783	pho1003	2

Q2) TREC_EVAL Results:

The results from the Run file for the Basic search and Custom search

Lucene Default Scoring Function

MAP: 0.6016

P@R: 0.5966

NDCG@20: 0.7696

Custom Scoring Function

MAP: 0.5095

P@R: 0.5257

NDCG@20: 0.6740

Team-3 Results:

Lucene Default Scoring Function

MAP: 0.4763557390826916

P@R: 0.5963763117196297

NDCG@20: 0.45207493345872407

Custom Scoring Function

MAP: 0.32668502318225295

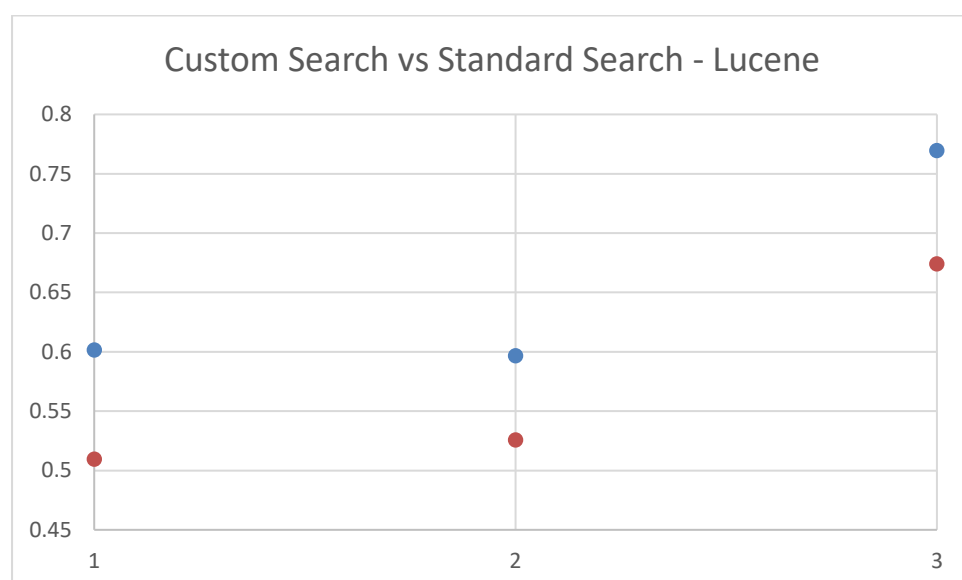
P@R: 0.5260761440065927

NDCG@20: 0.43956934365099

Q5) We are uncertain as to the root cause of the difference between our NDCG@20 numbers and the trec_eval numbers. Our formulas do not appear significantly different except for the “If $i < b$ ” = 1 Section from the paper by Jarvelin and Kekalainen. We believe this is covered when we add 1 in the logarithm of the denominator to prevent the division by 0.

Q6)

	Method	Comparison	Mean	Error
1	map	Standard	0.6016	0.001015
2	Rprec	Standard	0.5966	0.001165
3	ndcg_cut_20	Standard	0.7696	0.001052
1	map	Custom	0.5095	0.001215
2	Rprec	Custom	0.5257	0.001212
3	ndcg_cut_20	Custom	0.674	0.001239



With the standard error as low as it is, it is almost impossible to see the markers. With such a low standard error and a noticeable difference in all cases between the standard functionality and the custom, the standard Lucene comparison is superior. With the very limited rule set in the custom comparator, it is obvious that there would be a difference in the two method. What we find

interesting is the level of diminishing returns the standard function has over the clearly rudimentary custom comparator.