Programming Assignment 3

Team: 7

Team member:

Name. UNH id Webcat name

Kunpeng Xie 949610377 kx1005

Xin Liu 910991617 xl1044

1. Implement TF-IDF

Which variant perform best ?

For these three variants, only anc.pnc could be analyzed by trec_eval. The other two had run files, but they gave errors when using trec_eval tools to analyze. This may because there were some issues with our implementation of these two variants.

Lucene

	MAP	Precision@R	NDCG@20
Mean	0.59999	0.596	0.7369
Standard Error	0.243965	0.230038	0.194078

TF-IDF (anc.apc)

	MAP	Precision@R	NDCG@20
Mean	0.11695	0.24592	0.17295
Standard Error	0.13398	0.13101	0.162231

From the table above, it was clearly to find out Lucene got a higher mean for MAP, Precision@R and NDCG@20 scores. So we could conclude that Lucenes scoring was doing better than the anc.pnc variants.

The Standard Error

From the table above, we could saw that Lucene got higher standard error than TF-IDF. This may suggest that TF-IDF's results were more stable. And the difference was significant.

Spearman's Rank Correlation Coefficient

SpearMan Rank Correlation(lucene, anc.apc) = 8.331142

SpearMan Rank Correlation(lucene, lnc_ltn) = 40572.914

SpearMan Rank Correlation(lucene, bnn_bnn) = 0

From the above, we know that lnc.ltn is the closest to lucene default. However, we thought the results from above may have some problems, because it looked a little bit unusual, but we checked the implementation of SpearMan's Rank Correlation for many times, we believed our implementation is correct.

TF-IDF for Section Queries

We got the pageSkeleton for each page, and extracted the heading of this pageSkeleton, combined it with page name as keyword. Repeat task 1, then we get the .run file as result.

However, we don't know why our output files couldn't get any result when we ran trec_eval with hierarchical qurels file. It said that "trec_eval: No queries with both results and relevance info" .