

NAME	UNH ID	WEBCAT ID	ASSIGNMENT NUM
AJESH VIJAYARAGAVAN	999276672	avv1004@wildcats.unh.edu	3
NITHIN SIVAKUMAR	985392194	ns1077@wildcats.unh.edu	3
PEIHAO YUAN	999560760	py1004@wildcats.unh.edu	3

GitHub repository:- <https://github.com/Team54ever/a3>

Q1)

Results for lucene default ranking

```
nithins-mbp:trec_eval.9.0 Nithin$ ./trec_eval -m "set_P" -m "set_F" -m "P_2" -m
"Rprec" -m "map" -m "ndcg_cut.20" -c
/Users/Nithin/Desktop/test200/train.test200.cbor.article.qrels
/Users/Nithin/Desktop/outputfile
map                all    0.5718
Rprec              all    0.5755
P_2                all    0.9217
ndcg_cut_20        all    0.7342
set_P              all    0.4827
set_F              all    0.4639
```

Results for variant 1 – lnc.ltn

```
nithins-mbp:trec_eval.9.0 Nithin$ ./trec_eval -m "set_P" -m "set_F" -m "P_2" -m
"Rprec" -m "map" -m "ndcg_cut.20" -c
/Users/Nithin/Desktop/test200/train.test200.cbor.article.qrels
/Users/Nithin/Desktop/Outputs/lnc.ltn
map                all    0.5465
Rprec              all    0.5470
P_2                all    0.8965
ndcg_cut_20        all    0.7059
set_P              all    0.4789
set_F              all    0.4582
```

Results for variant 2 – bnn.bnn

```
nithins-mbp:trec_eval.9.0 Nithin$ ./trec_eval -m "set_P" -m "set_F" -m "P.2" -m  
"Rprec" -m "map" -m "ndcg_cut.20" -c  
/Users/Nithin/Desktop/test200/train.test200.cbor.article.qrels  
/Users/Nithin/Desktop/Outputs/bnn_bnn  
map                all    0.5175  
Rprec              all    0.5176  
P_2                all    0.8889  
ndcg_cut_20        all    0.6827  
set_P              all    0.4757  
set_F              all    0.4537
```

Results for variant 3 – anc.apc

```
nithins-mbp:trec_eval.9.0 Nithin$ ./trec_eval -m "set_P" -m "set_F" -m "P.2" -m  
"Rprec" -m "map" -m "ndcg_cut.10" -c  
/Users/Nithin/Desktop/test200/train.test200.cbor.article.qrels  
/Users/Nithin/Desktop/Outputs/anc_apc  
map                all    0.5070  
Rprec              all    0.5064  
P_2                all    0.8813  
ndcg_cut_20        all    0.6711  
set_P              all    0.4699  
set_F              all    0.4452
```

1. According to the scores, Inc.ltn is the best.

2. If we take a look for the scores, we can find lucene default ranking is the highest for each of those scores. So the variants perform worse than Lucene's default ranking model.

3.

Standard error(MAP) = 0.0293

Standard error(Rprec) = 0.03106

Standard error(ndcg_cut_20) = 0.02787

As we can see, the standard errors are so tiny. So the difference is not significant.

Q2)

$\text{SRCC}(\text{default}, \text{bnn.bnn}) = 0.636$

$\text{SRCC}(\text{default}, \text{Inc.ltn}) = 0.687$

$\text{SRCC}(\text{default}, \text{anc.apc}) = 0.523$

According to the results from SRCC, Inc.ltn is the closest to lucene's standard model.

Q3) Output files has been generated by tokenizing the keyword heading text. Check them here

<https://github.com/Team54ever/a3/tree/master/Outputs>