CSC849 HW4-Report
Chohee Kim 915409942

1. The search results' quality in terms of P@5, P@10, Recall@5, Recall@10, MRR (recip_rank) and MAP for each of the 3 parameter settings runs.

   a. X = 1

Evaluation	Query Id	Value
MAP	1	0.3333
Recip_rank	1	1.0000
P5	1	0.2000
P10	1	0.1000
Recall5	1	0.3333
Recall10	1	0.3333

Evaluation	Query Id	Value
MAP	2	1.0000
Recip_rank	2	1.0000
P5	2	0.6000
P10	2	0.3000
Recall5	2	1.0000
Recall10	2	1.0000

   b. X = 3

Evaluation	Query Id	Value
MAP	1	0.3333
Recip_rank	1	1.0000
P5	1	0.2000
P10	1	0.1000
Recall5	1	0.3333
Recall10	1	0.3333

Evaluation	Query Id	Value
MAP	2	1.0000
Recip_rank	2	1.0000
P5	2	0.6000
P10	2	0.3000
Recall5	2	1.0000
Recall10	2	1.0000

c. X = 5

Evaluation	Query Id	Value
MAP	1	0.3333
Recip_rank	1	1.0000
P5	1	0.2000
P10	1	0.1000
Recall5	1	0.3333
Recall10	1	0.3333

Evaluation	Query Id	Value
MAP	2	0.8056
Recip_rank	2	1.0000
P5	2	0.6000
P10	2	0.3000
Recall5	2	1.0000
Recall10	2	1.0000

2. Summarizes your observations about PRF and its parameter x.

Recall = (# of relevant items retrieved / # relevant items)
Precision = (# of relevant items retrieved)/ (# of retrieved items)

With different three parameters, x=1, 3, 5, the search program returns the same exact documents.

For the first query term, battery, returns document #5 which is the only document that contains the word "battery". However, my relevance judgment file doesn't only rely on the fact which whether the document contains the given word but also context of documents, which in this case document #1, 5, 8. In the result recall for the first query, it shows the relatively lower number compared to the second query. In Precision side, since I'm returning only one document out of three, its probability is small.
MAP value for the first query term is always small because it is returning one document out of three document.

For the second query term, screen, return a couple number of documents, document #1, 2, 3, 6, 8, and my relevance judgment file tells that only three documents, document #1, 3, and 6, are relevant document. It tells that my program only indicate its relevancy by checking appearance of word and same as the first query term, I was checking its relevancy considering context of document as well. As it all returns relevant documents, its precision is high but since I also returns non-relevant documents its recall is lower.
MAP value for the second query term is relatively high because it is returning all document out of relevant documents.

3. Reports all the kappa statistic scores that you computed in Part 1

0.9519038076152304
0.9540481400437637
0.9571734475374732
0.9521829521829522
0.9571734475374732
0.9558541266794626
0.9443254817987152
0.9540481400437637
0.9561243144424132
0.9559118236472945
0.9521829521829522
0.963531669865643
0.961456102783726