P3

Patrick Verga

November 29, 2014

1 Setup and Background

1.1 Search Engine

For this project I used the Galago search engine. My queries were run using SDM. I found Galago relatively easy to use and worked well for this task.

1.2 Query Creation

Query creation on the book data set proved to be very difficult. This is mainly because it was very hard to create a query with any relevant documents which could be due to the books being hundreds of years old. It was hard to know what types of information would or would not be included in the book set. The robust set was significantly easier as it was early 1990's news stories. The difficulty for robust was coming up with a query that did not return all relevant documents.

1.3 Judgments

The judgment process wasn't too bad. Sometimes it was hard to decide where on the 1-4 spectrum a document should fall. It also was occasionally a lengthy process when longer documents appeared.

2 Evaluations

For each of our three data sets we performed three evaluations : mean average precision, normalized discounted cumulative gain, and precision at 10. Table 2.1 shows the individual

Table 2.1: Books evaluations separated by query.

Query ID	MAP	NDCG@20	P@10
2	0.278	0.389	0.700
5	0.254	0.387	0.300
6	0.647	0.324	0.900
7	0.000	0.000	0.000
11	0.227	0.234	0.400
16	0.863	0.928	1.000
17	0.429	0.361	0.700
20	0.388	0.480	0.700
22	0.477	0.304	0.400
28	0.354	0.358	0.500
29	0.418	0.819	0.400

Table 2.2: Class robust evaluations separated by query.

Query ID	MAP	NDCG@20	P@10
1	0.106	0.070	0.300
5	0.102	0.388	0.100
10	0.601	0.391	0.700
12	0.483	0.521	0.600
15	0.569	0.366	0.500
17	0.491	0.900	0.400
18	0.697	0.911	1.000
20	0.128	0.117	0.200
24	0.521	0.593	0.400
27	0.572	0.680	0.800
30	0.702	0.384	0.800
33	0.761	0.907	1.000
34	0.136	0.243	0.100
35	0.500	0.631	0.250

Table 2.3: Overall scored for each data set.

Query ID	MAP	NDCG@20	P@10
Books	0.394	0.417	0.545
Class Robust	0.455	0.507	0.511
Community Robust	0.259	0.418	0.434

results for each of the book queries. Table 2.2 shows the individual results for each of the class robust queries. Table 2 shows the cumulative results for each data set.

3 PERFORMANCE

The system seems to have performed fairly well based on each of the evaluations. There seems to be quite a wide range in query results. On the book data, query 7 returned nothing relevant while query 16 was almost perfect. The range on the robust set was not quite as large ranging from query 5 with a MAP of .102 to query 33 with a map of .761.

3.1 ROBUST VS BOOKS

The class robust queries performed slightly better than the book queries. This could be due to the book dataset itself which has additional difficulties including spanning time and OCR errors.

3.2 Class vs Community Robust

The class robust queries performed significantly better than the community robust queries. This could be due to either differences in the way judgments were created or it could be the case that the class robust queries were simply easier.