

Title: Laboratory 2

Date: 02/11/21

Student names: Vincent Achukwu, Alan Hian Wu

4. Manual examination of a TREC test collection and sample search results

We decided to pick 4 topics - International Organized Crime, Poliomyelitis and Post-Polio, Hubble Telescope Achievements, and Endangered Species (Mammals). We decided to only use the BM25 evaluation algorithm.

Topic 1: International Organized Crime (ID 301)

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **title** the top 10 results are the following

Results	Relevant
FBIS4-40260	0
FBIS4-41991	0
FBIS3-23986	0
FBIS4-46734	0
FBIS4-39330	0
FBIS3-21961	1
FBIS4-7811	1
FBIS3-19646	1
FBIS4-41684	0
FBIS3-24145	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **description** the top 10 results are the following

Results	Relevant
FR940223-2-00242	0
FR940412-2-00131	0
FR940202-2-00150	0

FR940425-2-00078	0
FBIS3-23986	1
FBIS4-16951	0
FR940728-2-00151	0
FR940202-2-00151	0
FR941004-1-00089	0
FBIS4-40260	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **narrative** the top 10 results are the following

Results	Relevant
FR940503-0-00046	0
FBIS3-21722	0
FBIS3-10397	1
FBIS4-20966	1
FBIS3-41236	0
FBIS4-56513	1
FBIS4-67144	1
FBIS4-67492	1
FBIS4-9843	1
FBIS3-32453	1

Conclusion for Topic 1:

for topic 1 we choose to use (ID:301) International Organized Crime, before we started we concluded that just looking up the title will yield the least amount of relevant articles while looking up the narrative will yield the most relevant articles, this assumption is purely based of the fact that, the more information you provide to the search engine the more it has to work with and give us the articles we want. It is interesting to note that the first 5 results and the last 2 for the title search International Organized Crime which gave a result of 0 in terms of relevance were all about the Russian Federation.

For the results for the description part, we expected it to score in the middle of the ballpark between title and narrative. As expected the top 10 results had a shift, whereas the top 5 when is using the title has a relevant score of 0, only the top 4 of the description has a relevant score of 0, but unexpectedly the rest of the results are all 0.

Topic 2: Poliomyelitis and Post-Polio (ID: 302)

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **title** the top 10 results are the following

Results	Relevant
FBIS4-67701	1
LA043090-0036	1
LA013089-0022	0
FBIS4-30637	1
FBIS3-60403	1
FBIS3-60404	1
LA031489-0032	1
FR940126-2-00106	1
LA071590-0110	0
FBIS3-60405	1

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **description** the top 10 results are the following

Results	Relevant
FBIS4-67701	1
LA043090-0036	1
LA013089-0022	0
FR940603-2-00060	1
FBIS4-30637	1
LA082490-0065	1
LA072890-0066	1
FR940620-2-00117	1
FR940620-2-00118	0
FR940126-2-00107	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **narrative** the top 10 results are the following

Results	Relevant
FBIS3-60561	1
FR940620-2-00118	0
FBIS4-67701	1
FBIS3-60562	1
FR940126-2-00106	1
FR940620-2-00117	1
LA072890-0066	1
LA071590-0110	0
LA043090-0036	1
FBIS3-60403	1

Conclusion for Topic 2:

The results of topic 2 are based on Poliomyelitis and Post-Polio (ID: 302). We noticed that there were a lot more relevant documents from the results when searching the title, description, and narrative. We also observed how the trend of the relevancy of the top 10 results on the title, description, and narrative was similar. Again, we believe that the more terms are included, the more relevant results appear since it removes the ambiguity of the search term (i.e the title has fewer words than the description and narrative hence why we noticed less relevant results on the topic).

We observed that most of the results on this topic addressed the outbreak of polio disease. For instance, document [LA071590-0110](#) described “...*polio epidemics would sweep through a city or town; in 1916, about 27,000 people were paralyzed, 6,000 died...*”. Locations included India, China, and Southern Asian countries. Some of the results describe the plans involved for reducing disease outbreaks, for instance in document [FBIS4-67701](#) states “...*Specific goals set up by the government for other diseases include elimination of neonatal tetanus in all districts by 1995, reduction in measles cases by 90 per cent by 1995 compared to the 1985 level, reduction of deaths and cases due to diarrhoeal diseases by 30 per cent and ten per cent by 1985 and 70 per cent and 25 per cent by the end of this century....*”.

Topic 3: Hubble Telescope Achievements (ID: 303)

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **title** the top 10 results are the following

Results	Relevant
LA071090-0047	0
FT921-7107	1
LA041990-0151	0
LA050390-0109	0
LA041490-0064	0
LA070890-0154	0
LA072590-0006	0
LA042590-0152	0
LA070590-0031	0
LA033090-0082	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **description** the top 10 results are the following

Results	Relevant
LA041490-0064	0
LA033090-0082	0
LA040190-0178	0
LA041990-0151	0
LA041090-0148	0
LA042590-0135	0
LA070990-0052	0
LA042590-0152	0
FT934-5418	0
LA041690-0035	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **narrative** the top 10 results are the following

Results	Relevant
LA052890-0021	0
FT921-7107	1
FBIS3-42547	0
LA110590-0076	1
LA042790-0070	0
FBIS4-46650	0
LA041990-0151	0
LA120590-0110	0
LA071390-0122	0
FT944-128	1

Conclusion for Topic 3:

There were many irrelevant documents from this topic. We know the relevant documents are relevant because they describe how the Hubble telescope has produced new data, better quality data than previously available, data that has increased human knowledge of the universe, or data that has led to disproving previously existing theories or hypotheses. For example, in document [LA042790-0070](#) "...Today, the telescope will pass another milestone when ground controllers order it to open its "aperture door," which serves as a lens cap to prevent sunlight from accidentally entering the telescope and burning out some of the instruments....". This suggests that this operation can lead to an increase in human knowledge. Also in document [FT944-128](#) "...The research team found 20 individual stars of a type called Cepheid variables in the distant M100 galaxy...".

Topic 4: Endangered Species (Mammals) (ID: 304)

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **title** the top 10 results are the following

Results	Relevant
FR940419-2-00009	0
FR940228-2-00026	0
FR941216-2-00020	0
FR940203-0-00059	0
FR940617-0-00103	0

FR941205-0-00114	0
FR941227-2-00012	0
FR940617-0-00104	1
FR940706-2-00012	1
FR940930-2-00144	1

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **description** the top 10 results are the following

Results	Relevant
FR940713-1-00064	1
FR940203-0-00093	0
FR940616-0-00132	1
FR940204-0-00063	0
FR940816-1-00069	1
FR940816-1-00084	1
FR941122-0-00067	0
FR941122-0-00092	1
FR941026-2-00030	0
FR940328-1-00053	0

Method, when using BM25 $k=1.2$ $b=0.75$

By just using the **narrative** the top 10 results are the following

Results	Relevant
FR940617-0-00103	0
FR940104-1-00049	0
FR940617-0-00104	1
FR940419-2-00009	1
FR941216-2-00020	0
FR940228-2-00026	0

FR940627-1-00012	0
FR941227-2-00087	1
FR940804-1-00036	0
FR940203-0-00059	1

Conclusion for Topic 4:

The main narrative for this topic to determine document relevancy is with regards to how endangered mammals are. We found only 12 relevant documents for this topic, very few are based on the title. Again, we found that using fewer words on the search query resulted in fewer relevant results in comparison to using the description and narrative which have more words. Relevant documents such as [FR940930-2-00144](#) mentions mammals as well as how they are being endangered. *“...polar bear, sea and marine otters, walrus, manatees (three species) and dugong. The report reviews the Service's marine mammal-related activities during the report period. Administrative actions discussed include appropriations, marine mammals in Alaska, endangered and threatened marine mammal species...”*.

5. Exploring Evaluation Metrics

After setting up the trec_eval command upon installing the latest package, we compiled and ran the program with the sample run data “trec678.res” along with the qrel data in “qrels.trec678.adhoc”. The image below shows the results obtained from running the program.

```
vinny@LAPTOP-JNEHHUSV:~/trec_eval-9.0.7$ ./trec_eval ../trec/trec_lab/qrels.trec678.adhoc ../trec/trec_lab/trec678.res
runid          all      1m
num_q           all      150
num_ret         all     142395
num_rel         all     14013
num_rel_ret     all      7282
map             all      0.2145
gm_map          all      0.0991
Rprec           all      0.2644
bpref           all      0.2367
recip_rank      all      0.5778
iprec_at_recall_0.00 all     0.6393
iprec_at_recall_0.10 all     0.4549
iprec_at_recall_0.20 all     0.3590
iprec_at_recall_0.30 all     0.3014
iprec_at_recall_0.40 all     0.2377
iprec_at_recall_0.50 all     0.1925
iprec_at_recall_0.60 all     0.1510
iprec_at_recall_0.70 all     0.1125
iprec_at_recall_0.80 all     0.0728
iprec_at_recall_0.90 all     0.0528
iprec_at_recall_1.00 all     0.0272
P_5             all      0.4240
P_10            all      0.4027
P_15            all      0.3738
P_20            all      0.3487
P_30            all      0.3111
P_100           all      0.1940
P_200           all      0.1369
P_500           all      0.0785
P_1000          all      0.0485
```


The results are averaged based on all of the topics from “topics.301-450.xml”. We noticed that adding the “-q” flag to the command would yield a larger result set of what is shown above.

5.1 Examine the values of Precision at rank cutoffs, Mean Average Precision (MAP) and Recall

Based on the results, it seems that as the precision “n” increases, the number of relevant documents decreases. The smaller “n” is, the more relevant documents we have. We may conclude that the most relevant documents would be listed in the first few results. This makes sense because usually, this is how we want search results to be displayed to avoid users having to search and scroll through results associated with their search queries. This can also mean that users would typically look through the first few results rather than the first 100 or more resulting documents.

The mean average precision is calculated based on getting the mean of all the AP scores which would result in a single value, and this determines how good the program can perform the query.

The recall is the fraction of the overall relevant documents retrieved and is a good measure of determining how relevant a document is.

6. Exploring Consistency of Relevance Assessment for Topics

The three topics we have selected for this topic are:

- 310 - Radio Waves and Brain Cancer
- 311 - Industrial Espionage
- 312 - Hydroponics

From these topic titles, we can see how there’s a variation with regards to how many words are in the topic. It will be interesting to see how many accurate and relevant results we’ll find for each topic.

For each topic, we will use BM25 as before with $k = 1.2$ and $b = 0.75$ as well as document the first 10 results.

(comments: seems like it's the same as the first section, except we have to determine if it's relevant ourselves then check the qrel file to see if it says it's relevant)

Topic 1: Radio Waves and Brain Cancer (ID: 310)

Results	Our Relevance Judgement	qrel	Summary
FT931-11958	1	1	Relationship between the use of mobile phones to brain cancer
LA100889-0041	0	0	Detecting diseases and monitoring their response to therapy without X-rays
LA053090-0120	0	0	Study of brain cancer (no mention of radio waves)
FT922-289	0	0	Study of brain cancer (no mention of radio waves)
FT922-15435	1	1	Cancer risk related to living near power lines
LA050190-0024	0	0	Bishop leaves hospital determined to fight cancer
LA100389-0004	0	0	Killing cancer cells with superconductivity
LA010190-0022	0	0	How bone marrow can be removed and treated chemically to kill cancer cells
LA031690-0091	0	N/A	Fred Lebow returns for a second biopsy on his brain
FT941-5519	0	0	Bodily cell reactions to cancer cells

Topic 2: Industrial Espionage (ID: 311)

Results	Our Relevance Judgement	qrel	Summary
FT944-15440	1	1	Company rivalry espionage
FBIS3-58751	1	1	Industrial espionage, and the measures taken to prevent it
FT931-3928	1	1	Founder of Britain's Special Air Service - carried out a three-year espionage operation against Europarks to acquire confidential information
FBIS4-19393	1	1	Measures against industrial spying
FT934-7093	0	0	Counter-espionage. No mention of the theft of assets
FBIS3-31719	1	1	Tamara (state-of-the-art air defence system) failed attempt at industrial espionage

FBIS4-51593	1	0	Aerial espionage on the strategic in-depth areas of the northern half of Korea in the skies above Yangpyong
FBIS4-27829	1	0	The U.S conducted photographing and electronic espionage on the whole area and coastal area of the Northern half of Korea, in advance of invading the Northern half of Korea
FBIS4-29890	1	0	The U.S Nuclear submarines extend espionage against DPRK
LA081789-0098	1	1	German computer hackers were suspected of giving the Soviet Union information from military

Topic 3: Hydroponics (ID: 312)

Results	Our Relevance Judgement	qrel	Summary
LA040289-0050	1	1	Discussions on agricultural practices in space via hydroponics
LA051090-0097	1	1	Growing marijuana with hydroponic equipment
LA052490-0029	1	1	Marijuana forest with a hydroponic feeding system
LA102689-0029	0	0	Mentions hydroponics once as an example, but the overall topic revolves around how tomatoes are a perennial choice for cooks
LA102490-0077	1	1	Mentions how tomatoes are grown hydroponically - <i>"the tomatoes are planted in ceramic pebbles and watered four times a day"</i>
FR941007-0-00016	0	0	Simply gave the definition of the term <i>"hydroponic tomatoes"</i> , but overall doesn't cover the science of hydroponics - <i>"Hydroponic tomatoes means tomatoes grown in solution without soil"</i>
FBIS4-67023	1	1	How drug marketers grow drugs via hydroponics and how they can grow more quickly this way
FR940830-2-00026	0	0	No mention of the science/history/advantages of hydroponics or nutrients needed for it. The document seems

			to list out random scientific terms, including hydroponics
FR940602-0-00005	0	0	Only mentions hydroponics as part of an eligible type of business activity
FT932-13105	1	1	Growing chicory with hydroponics

Conclusions regarding topic relevancy:

We observed that one of the topics (Topic 1 - Radio Waves and Brain Cancer (ID: 310)) had a document that was not present in the qrel file and was deemed to be irrelevant as stated in the lab sheet - *“Any documents that have been retrieved by the search system and are not present in the qrel file for a topic were not judged for relevance when the qrel file was created for TREC, and were considered to be non-relevant by the TREC organisers”*. We indeed found it not to be relevant anyway.

Interestingly, topic 2 (Industrial Espionage (ID: 311)) had the most relevant documents according to our observations, though this may be due to the uncertainty with whether or not this document was relevant in terms of military action. The qrel file deemed it to be irrelevant with how the U.S used espionage techniques to surveil North Korea, but we believed that it was considered espionage. Maybe we’re wrong, or the qrel was wrong.