Information Retrieval Programming Assignment 4: Language Models

Shayan Amani
(993550898, sa1149)

Soheil Gharatappeh
(933639024, sg1147)

Source Code

Please refer to our GitHub repository for the source code.

Implement Unigram Language Models with Smoothing

1. Find the `run files` for all variants in here.

2. All the measures are compared in the following table and plot.

Evaluation	MAP	Precision@R	NCCG@20
U-L	?	?	?
U-JM	0.568	0.5734	0.7275
U-DS	0.564	0.5635	0.724

3. The following table shows mean and standard error for these two evaluations.

Evaluation	MAP	Precision@R	NCCG@20
Mean	0.566	0.56845	0.72575
Standard Error	0.002	0.00495	0.00175

- According to the table, U-DS performs better than U-JM.

- As we can see in the following figures, Unigram Language Models have a better performance in comparison to TF-IDF.

- No, the deference $(MAP, Prec@R, ndcg@20) = (0.002, 0.00495, 0.00175)$ is not significant.
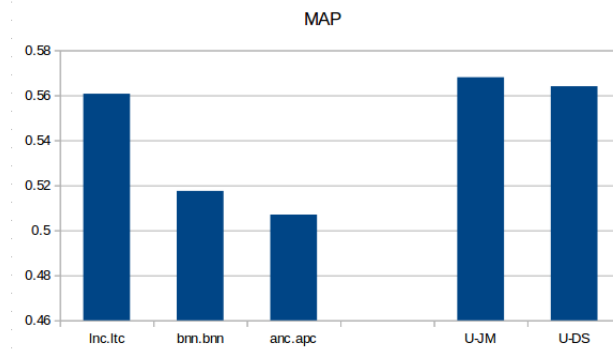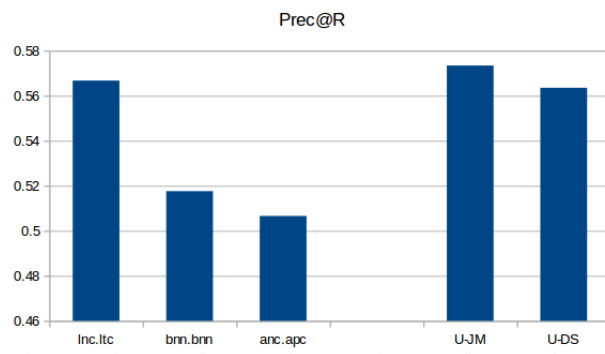
1

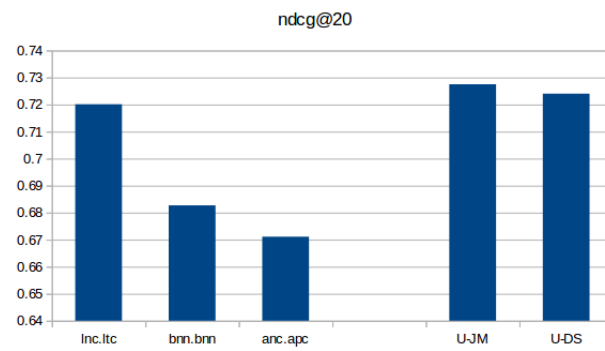Figure 1: Map comparison.

Figure 2: Prec@R comparison.

Figure 3: ndcg@20 comparison.

- No, they theoretically cannot agree on "the best" method. As mentioned, each of this measures, focuses on one thing. So, it cannot perform well in the other. The root cause of that is that we don't have a perfect model that can give us 100% precision and 100% recall. When we increase precision, the recall would decrease. Because we take a set of more strict criteria in our search engine. This is true the other way around; if we recall more docs, there is more chance to recall non-relevant docs too.