

Your Name:

Your Andrew ID:

Homework 2

0 Introduction

0.1 Collaboration and Originality

1. Did you receive help of any kind from anyone (other than the instructor or TAs) in developing your software for this assignment (Yes or No)? If you answered Yes, provide the name(s) of anyone who provided help, and describe the type of help that you received.
2. Did you give help of any kind to anyone in developing their software for this assignment (Yes or No)? If you answered Yes, provide the name(s) of anyone that you helped, and describe the type of help that you provided.
3. Did you examine anyone else's software for this assignment (Yes or No)? Do not describe software provided by the instructor.
4. Are you (or the course instructor) the author of every line of source code submitted for this assignment (Yes or No)? If you answered No:
 - a. identify the software that you did not write,
 - b. explain where it came from, and
 - c. explain why you used it.
5. Are you the author of every word of your report (Yes or No)? If you answered No:
 - a. identify the text that you did not write,
 - b. explain where it came from, and
 - c. explain why you used it.

0.2 Instructions

0.2.1 Content

Each experiment requires you to set parameters or weights. You must explain why you chose particular values, and how your choices relate to how the technique works. We look for good experimental design – parameters that explore interesting issues or hypotheses (even if the hypothesis turns out to be wrong).

You must analyze the experimental results. Explain what conclusions you can reach based on the experiment. You could discuss the stability of results across different parameter settings; effects on Precision or Recall; accuracy vs. computational effort; or other aspects of the experimental results that interest you. Usually a good analysis addresses several issues. Show that you understand what the results mean, based upon what we have discussed in class.

Advice from the TAs:

- Do not spend a lot of time discussing baseline results from HW1 unless there is something especially interesting about them for HW2.*
- It is difficult to write good discussions for unfocused parameter sweeps. It helps to have a strategy to guide parameter choices.*
- A discussion that just repeats what is obvious from the table (e.g., $A > B$) tends to receive weaker scores.*

0.2.2 Format

Instructions are shown in this red italic bold font. Do not include instructions in your report. For example, delete this page, and in the next section, delete the instruction paragraph.

Leave the page breaks between sections, as shown in this template. For example, Sections 1.1 and 1.2 must be on different pages.

There is a 2 point deduction for not following format instructions because it creates extra work for the TAs.

1 Experiment: Baselines

1.1 Baseline Experimental Results

Your .zip / .tgz file must include files named HW2-Exp-1.1a.qry, HW2-Exp-1.1a.param, etc., in the QryEval directory. The experimental results shown below must be reproducible by these files.

Short Queries				
	Ranked Boolean AND (Exp-1.1a)	BM25 BOW (Exp-1.1b)	Indri-D BOW (Exp-1.1c)	Indri-JM BOW (Exp-1.1d)
MRR	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000
Runtime	mm:ss	mm:ss	mm:ss	mm:ss

Long Queries				
	Ranked Boolean AND (Exp-1.2a)	BM25 BOW (Exp-1.2b)	Indri-D BOW (Exp-1.2c)	Indri-JM BOW (Exp-1.2d)
MRR	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000
Runtime	mm:ss	mm:ss	mm:ss	mm:ss

1.2 Parameter Choices for Baseline Experimental Results

Explain and justify your choice of parameters.

2 Sequential Dependency Models

2.1 Experimental Results for Sequential Dependency Models

Your .zip / .tgz file must include files named HW2-Exp-2.1a.qry, HW2-Exp-2.1a.param, etc., in the QryEval directory. The experimental results shown below must be reproducible by these files.

Short Queries					
	Indri BOW (body) (Exp-2.1a)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.1b)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.1c)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.1d)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.1e)
MRR	0.0000	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000	0.0000
Runtime	mm:ss	mm:ss	mm:ss	mm:ss	mm:ss

Long Queries					
	Indri BOW (body) (Exp-2.2a)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.2b)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.2c)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.2d)	0.00 AND 0.00 NEAR 0.00 WINDOW (Exp-2.2e)
MRR	0.0000	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000	0.0000
Runtime	mm:ss	mm:ss	mm:ss	mm:ss	mm:ss

2.2 Example SDM Query

Show the structured query for “Scottish Highland Games”.

2.3 SDM Weight Choices

Explain and justify your choice of weights.

2.4 Analysis of Sequential Dependency Model Results

Analyze the experimental results.

3 Indri Smoothing

3.1 Experimental Results for Indri Smoothing

Your .zip / .tgz file must include files named HW2-Exp-3.1a.qry, HW2-Exp-3.1a.param, etc., in the QryEval directory. The experimental results shown below must be reproducible by these files.

	Short Queries μ (Note: $\lambda=0$)				
	$\mu=1500$ (Exp-3.1a)	$\mu=?$ (Exp-3.1b)	$\mu=?$ (Exp-3.1c)	$\mu=?$ (Exp-3.1d)	$\mu=?$ (Exp-3.1e)
MRR	0.0000	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000	0.0000

	Long Queries μ (Note: $\lambda=0$)				
	1500 (Exp-3.2a)	$\mu=?$ (Exp-3.2b)	$\mu=?$ (Exp-3.2c)	$\mu=?$ (Exp-3.2d)	$\mu=?$ (Exp-3.2e)
MRR	0.0000	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000	0.0000

3.2 Parameter Choices for Indri Parameter Adjustment

Explain and justify your choice of parameters.

3.3 Analysis of Indri Parameter Adjustment Results

Analyze the experimental results.

4 Multiple Fields

4.1 Experimental Results for Multiple Fields

Your .zip / .tgz file must include files named HW2-Exp-4.1a.qry, HW2-Exp-4.1a.param, etc., in the QryEval directory. The experimental results shown below must be reproducible by these files.

	Short Queries			
	Ranked Boolean AND (Exp-4.1a)	BM25 BOW (Exp-4.1b)	Indri-D BOW (Exp-4.1c)	Indri-JM BOW (Exp-4.1d)
MRR	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000

	Long Queries			
	Ranked Boolean AND (Exp-4.2a)	BM25 BOW (Exp-4.2b)	Indri-D BOW (Exp-4.2c)	Indri-JM BOW (Exp-4.2d)
MRR	0.0000	0.0000	0.0000	0.0000
P@10	0.0000	0.0000	0.0000	0.0000
P@20	0.0000	0.0000	0.0000	0.0000
P@30	0.0000	0.0000	0.0000	0.0000
MAP	0.0000	0.0000	0.0000	0.0000
R@100	0.0000	0.0000	0.0000	0.0000
R@1000	0.0000	0.0000	0.0000	0.0000

4.2 Parameter Choices for Multiple Fields

Explain and justify your choice of parameters.

4.3 Analysis of Multiple Fields Results

Analyze the experimental results.