## Your Name:

**Your Andrew ID:**

**Homework 2**

# Collaboration and Originality

Your report must include answers to the following questions:

1. Did you receive help of any kind from anyone in developing your software for this assignment (Yes or No)? It is not necessary to describe discussions with the instructor or TAs.

If you answered Yes, provide the name(s) of anyone who provided help, and describe the type of help that you received.

1. 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.

1. Are you the author of every line of source code submitted for this assignment (Yes or No)? It is not necessary to mention software provided by the instructor.

If you answered No:

* 1. identify the software that you did not write,
  2. explain where it came from, and
  3. explain why you used it.

1. Are you the author of every word of your report (Yes or No)?

If you answered No:

* 1. identify the text that you did not write,
  2. explain where it came from, and
  3. explain why you used it.

## Your Name:

**Your Andrew ID:**

**Homework 2**

# Experiment 1: Baselines

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Ranked**  **Boolean** | **BM25**  **BOW** | **Indri**  **BOW** |
| **P@10** | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 |

# Experiment 2: Queries with Synonyms and Phrases

## Queries

List your queries.

## Query descriptions

For each query, provide a brief (1-2 sentences) description that identifies which strategy was used for that query, any important deviations from your default strategies, and your intent, i.e., why you thought that particular structure was a good choice.

## Experimental Results

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Ranked**  **Boolean** | **BM25**  **BOW** | **Indri**  **BOW** | **Ranked**  **Boolean**  **Syn/Phr** | **BM25**  **Syn/Phr** | **Indri**  **Syn/Phr** |
| **P@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## Discussion

Discuss any trends that you observe; whether the use of synonyms and phrases behaved as you expected; and any other observations that you may have.

# Experiment 3: BM25 Parameter Adjustment

## k1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **k1** | | | | | | | |
| 1.2 | Value 1 | Value 2 | Value 3 | Value 4 | Value 5 | Value 6 | Value 7 |
| **P@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

.

## b

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **b** | | | | | | | |
| 0.75 | Value 1 | Value 2 | Value 3 | Value 4 | Value 5 | Value 6 | Value 7 |
| **P@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## Discussion

Explain your reasons for choosing the values that you tested, and how those reasons are related to how BM25 works. Discuss any changes in retrieval performance that you observed, and the significance of any trends that you observed.

# Indri Parameter Adjustment

## µ

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **µ** | | | | | | | |
| 2500 | Value 1 | Value 2 | Value 3 | Value 4 | Value 5 | Value 6 | Value 7 |
| **P@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## 𝜆

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **𝜆** | | | | | | | |
| 0.4 | Value 1 | Value 2 | Value 3 | Value 4 | Value 5 | Value 6 | Value 7 |
| **P@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## Discussion

Explain your reasons for choosing the values that you tested, and how those reasons are related to how Indri works. Discuss any changes in retrieval performance that you observed, and the significance of any trends that you observed.