**Your Name**

**Your Andrew ID**

**Homework 3**

# Collaboration and Originality

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 3**

# 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 |

Document the parameter settings that were used to obtain these results.

# Experiment 2: Different representations

**Example Query:** Provide your structured query for query “sherwood regional library”.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Indri**  **BOW**  **(body)** | **0.00 url**  **0.00 keywords**  **0.00 title**  **0.00 body**  **0.00 inlink** | **0.00 url**  **0.00 keywords**  **0.00 title**  **0.00 body**  **0.00 inlink** | **0.00 url**  **0.00 keywords**  **0.00 title**  **0.00 body**  **0.00 inlink** | **0.00 url**  **0.00 keywords**  **0.00 title**  **0.00 body**  **0.00 inlink** | **0.00 url**  **0.00 keywords**  **0.00 title**  **0.00 body**  **0.00 inlink** |
| **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 |

Describe your strategy for setting the weights on the different representations. Describe how you expected the different weight combinations to perform (before the experiment), and why.

Discuss the trends and stability that you observed in Experiment 2; whether the different weight combinations behaved as you expected; how Precision and Recall tended to vary as the weights were varied; how the differences in accuracy (if any) relate to different computational cost; and your other observations, interpretations, or conclusions from running this experiment.

# Experiment 3: Sequential dependency models

**Example Query:** Provide your structured query for query “sherwood regional library”.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Indri**  **BOW**  **(body)** | **0.00 AND**  **0.00 NEAR**  **0.00 WINDOW** | **0.00 AND**  **0.00 NEAR**  **0.00 WINDOW** | **0.00 AND**  **0.00 NEAR**  **0.00 WINDOW** | **0.00 AND**  **0.00 NEAR**  **0.00 WINDOW** | **0.00 AND**  **0.00 NEAR**  **0.00 WINDOW** |
| **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 |

Describe your strategy for setting the weights for the different components of the sequential dependency model. Describe how you expected the different weight combinations to perform (before the experiment), and why.

Discuss the trends and stability that you observed in Experiment 3; how Precision and Recall tended to vary as the weights were varied; whether the more complex query behaved as you expected; whether the improvement in accuracy (if any) is worth the increased computational cost; and your other observations, interpretations, or conclusions from running this experiment.