**Homework 4**

# Introduction

## 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.

No.

1. Did you give help of any kind to anyone in developing their software for this assignment (Yes or No)?

No.

1. Did you examine anyone else’s software for this assignment (Yes or No)? It is not necessary to mention software provided by the instructor.

No.

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.

Yes.

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

Yes.

# Experiment 1: Baselines

## Experimental Results

|  |  |  |  |
| --- | --- | --- | --- |
|  | **BM25**  **(Exp-1a)** | **Indri**  **BOW**  **(Exp-1b)** | **Indri**  **SDM**  **(Exp-1c)** |
| **P@10** | 0.436 | 0.432 | 0.46 |
| **P@20** | 0.422 | 0.432 | 0.456 |
| **P@30** | 0.4173 | 0.4373 | 0.4547 |
| **NDCG@10** | 0.3481 | 0.3276 | 0.3475 |
| **NDCG@20** | 0.3509 | 0.3379 | 0.3559 |
| **NDCG@30** | 0.3538 | 0.3515 | 0.3649 |
| **MAP** | 0.2412 | 0.2576 | 0.2718 |

## Parameters

retrievalAlgorithm=Indri

BM25:k\_1=1.2

BM25:b=0.75

BM25:k\_3=0

Indri:mu=2500

Indri:lambda=0.4

For sequence dependency, I choose 0.8 for #AND operator, 0.1 for #NEAR operator and 0.1 for #WINDOW operator.

# Custom Features

Feature 17, the date of the document is collected. This feature could be very helpful, because people might be more interested in new document than old document. The computation complexity is very low.

Feature 18 finds out all the positions of terms in the query and calculates its standard deviation. The hypothesis here is that if all terms appear together rather than diversely in the document, the document could be a better match. Because terms could be use together to form some concept appearing in the query. I use the standard deviation of all the positions to represent how closely terms are used in the document. The time complexity of creating this feature is in the order of the length of the document. Because a pass of scan of all positions is needed.

# Experiment 2: Learning to Rank

## Experiment Table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **BM25** | **IR**  **Fusion**  **(Exp-3a)** | **Content-**  **Based**  **(Exp-3b)** | **Base**  **(Exp-3c)** | **All**  **(Exp-3d)** |
| **P@10** | 0.436 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@20** | 0.422 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **P@30** | 0.4173 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@10** | 0.3481 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@20** | 0.3509 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@30** | 0.3538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.2412 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## Parameters

***Briefly describe the parameter settings that were used to obtain these results.***

## Discussion

***Discuss the trends that you observe; whether the learned retrieval models behaved as you expected; how the learned retrieval models compare to the baseline methods; and any other observations that you may have.***

# Experiment 3: Feature Combinations

***Experiment with different combinations of features.***

## Experiment Table

***Your .zip / .tgz file must include files named HW4-Exp-4a.qry, HW4-Exp-4a.param, etc., in the QryEval directory. The experimental results shown above must be reproducible by these files and the parameter values shown in the table.***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All**  **(Baseline)**  **(Exp-4a)** | **Comb1**  **(Exp-4b)** | **Comb2**  **(Exp-4c)** | **Comb3**  **(Exp-4d)** | **Comb4**  **(Exp-4e)** |
| **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 |
| **NDCG@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

***11-742 students should add a second table with the additional experiments.***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All**  **(Baseline)**  **(Exp-4a)** | **Comb5**  **(Exp-4f)** | **Comb6**  **(Exp-4g)** | **Comb7**  **(Exp-4h)** | **Comb8**  **(Exp-4ji** |
| **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 |
| **NDCG@10** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@20** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **NDCG@30** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **MAP** | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

## Parameters

***Briefly describe the parameter settings that were used to obtain these results.***

## Discussion

***Describe each of your feature combinations, including its computational complexity. Explain the intuitions behind your choices. This does not need to be a lengthy discussion, but you need to convince us that your combinations are investigating interesting hypotheses about what delivers good search accuracy. Were you able to get good effectiveness from a smaller set of features, or is the best result obtained by using all of the features? Why?***

***Be sure to discuss the effectiveness of your custom features. This should be more insightful than “They improved P@10 by 5%”. Discuss the effect on your retrieval experiments, and if there is variation in the metrics that are affected (e.g., P@k, MAP), how those variations compared to your expectations.***

# Analysis

***Examine the model files produced by SVMrank. Discuss which features appear to be more useful and which features appear to be less useful. Support your observations with evidence from your experiments. Keep in mind that some of the features are highly correlated, which may affect the weights that were learned for those features.***

***Some of this discussion may overlap with your discussion of your experiments. However, in this section we are primarily interested in what information, if anything, you can get from the SVMrank model files.***