**CE306/CE706 - Information Retrieval**

**Assignment 2**

Juntao Yu

March 2022

# Plagiarism

*You are reminded that this work is for credit towards the composite mark in CE306/CE706* ***and that the work you submit must therefore be your own.*** *Any material you make use of, whether it be from textbooks, the Web or any other source must be acknowledged as a comment in the program, and the extent of the reference indicated.*

# The context of your task

To properly evaluate a system, you usually need to construct test collections. Given information needs and documents, you need to collect relevance assessments. This is a time-consuming and expensive process involving human beings (in this case you). For tiny collections, exhaustive judgments of relevance for each query and document pair can be obtained. For large modern collections, it is usual for relevance to be assessed only for a subset of the documents for each query. The most standard approach is **pooling**, where relevance is assessed over a subset of the collection that is formed from the top *k* documents returned by many different IR systems.

**Your task**

This task comes in stages. Marks are given for each stage. The stages are as follows:

* **Assessing relevance (20%)** Suppose you are come to Colchester campus to attend you graduation with you parents, you travel by car and want to check out the car park policy for graduation. The search engine you used give you the following results, your task is to make a binary judgement for each document on their relevance (relevance/non-relevance) and explain why. (You need click each link and check the contents of the webpage)
  1. <https://www.essex.ac.uk/graduation/about-the-ceremony>
  2. <https://www1.essex.ac.uk/site/a-z.aspx>
  3. <https://www.essex.ac.uk/alumni/honorary/honorary-graduates>
  4. <https://www.essex.ac.uk/disclaimer/accessibility-hub/www1-applications>
  5. <https://www.essex.ac.uk/graduation/tickets-and-information>
  6. <https://www.essex.ac.uk/alumni/awards/alumnus-of-the-year/2008>
  7. <https://www.essex.ac.uk/information/travel-and-transport/car-parks-at-colchester-campus>
  8. <https://www.essex.ac.uk/life/colchester-campus/how-to-get-here>
  9. <https://www.essex.ac.uk/about/how-we-teach>
  10. <https://www.essex.ac.uk/arena/how-to-get-here>
* **Pooling (10%)** Now suppose we have three IR systems (IR1, IR2 and IR3) developed and we need to create a test collection by using the **Pooling.** In total we have 20 documents (1 - 20), a program f (x) = x%2 is served as human assessor in this task i.e., f (1) = 1 (relevant) and f (2) = 0 (non-relevant). For this task we construct the pool by putting together **top 10** retrieval results. What are the relevance labels we will get for each document? (**N.B.** Documents outside the pool are automatically considered to be irrelevant)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| **IR1** | 7 | 15 | 13 | 9 | 16 | 19 | 8 | 18 | 17 | 1 | 5 | 2 | 14 | 20 | 4 | 10 | 3 | 11 | 6 | 12 |
| **IR2** | 20 | 8 | 9 | 1 | 16 | 3 | 17 | 18 | 4 | 2 | 14 | 7 | 13 | 5 | 11 | 10 | 15 | 6 | 19 | 12 |
| **IR3** | 6 | 1 | 16 | 7 | 4 | 19 | 18 | 3 | 20 | 14 | 8 | 5 | 9 | 12 | 13 | 15 | 17 | 2 | 11 | 10 |

* **P/R@5 (15%)** Once you have your relevance labels from pooling, you can explore the effect of each IR system on the evaluation results. The first task you will need to compute the P@5 and R@5 for all three systems.
* **Average Precision (15%)** Next let’s compute the average precision for all three systems. Again, we will use the relevance labels from pooling.
* **Discounted Cumulative Gain (15%)** Next you need to compute the DCG for all systems, here we use the binary judgement, i.e., rel = 0/1 for non-relevant and relevant documents respectively the label is decided by the pooling step.
* **Precision-Recall Curves (15%)** Now you are required to draw the precision-recall curves for those systems. Suppose we need to select one system for a scholar search that require 80% of recall, which system you are going to choose?
* **Web Search (10%)** Finally you need to choose a system for web search, which metrics you will use to make your decision and why? According to the metric you choose which system you will use.

# Submission

You should submit:

* Report (**use the template below**)

The submission should be submitted as a single *pdf file* via the electronic submission system. Please check the details of the submission deadline with the CSEE School Office.

*The guidelines about late assignments are explained in the students’ handbook.*

**CE306 or CE706 - Information Retrieval 2022**

**Assignment 2**

1905876

# Assessing relevance (Task 1)

*Include here your judgments and the reason you make your decision.*

|  |  |  |
| --- | --- | --- |
| **Doc** | **Judgment** | **Reason** |
| 1. |  |  |
| 2. |  |  |
| 3. |  |  |
| 4. |  |  |
| 5. |  |  |
| 6. |  |  |
| 7. |  |  |
| 8. |  |  |
| 9. |  |  |
| 10. |  |  |

﻿ Pooling (Task 2)

*Include here the judgments for each document.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Doc** | **Judgment** | **Doc** | **Judgment** |
| 1. | Relevant | 11. | Irrelevant |
| 2. | Irrelevant | 12. | Irrelevant |
| 3. | Relevant | 13. | Relevant |
| 4. | Irrelevant | 14. | Irrelevant |
| 5. | Irrelevant | 15. | Relevant |
| 6. | Irrelevant | 16. | Irrelevant |
| 7. | Relevant | 17. | Relevant |
| 8. | Irrelevant | 18. | Irrelevant |
| 9. | Relevant | 19. | Relevant |
| 10. | Irrelevant | 20. | Irrelevant |

﻿ P/R@5 (Task 3)

*For each system, compute P/R@5, please include details of the computation.*

|  |  |  |
| --- | --- | --- |
| **Systems** | **P@5** | **R@5** |
| *IR1:* | *0.8* | *0.5* |
| *IR2:* | *0.4* | *0.25* |
| *IR3:* | *0.4* | *0.25* |

# Average Precision (Task 4)

*For each system, compute average precision, please include details of the computation.*

|  |  |
| --- | --- |
| **Systems** | **Average Precision** |
| *IR1:* | *0.83* |
| *IR2:* | *0.45* |
| *IR3:* | *0.46* |

# Discounted Cumulative Gain (Task 5)

*For each system, compute discounted cumulative gain, please include details of the computation.*

|  |  |
| --- | --- |
| **Systems** | **Discounted Cumulative Gain** |
| *IR1:* | *18.72* |
| *IR2:* | *24.56* |
| *IR3:* | *24.26* |

# Precision-Recall Curves (Task 6)

*For each system, draw the precision-recall curves, and choose your system for the scholar search and explain why.*

# Web search (Task 7)

*Which system you are going to use for web search? Choose one metric from above to support your decision and explain why you choose this metric.*