Information Retrieval 8130

Assignment 8

Submitted by: Hasan Mashrique

**Due Date:** 11/28/2022

**Problem 1 [30 points]**.

Prepare a 2-page report describing the results of evaluating your search engine. For measuring performance use the following 10 queries to report

Precision, Recall, and F-measure. Do your best to compute recall for

the queries.

The QUERIES are.

q1: Computer Science

q2: President of the university

q3: Information Retrieval

q4: Learner Data Institute

q5: International student office

q6: Graduate school admissions

q7: What is the mascot of the University of Memphis?

q8: College of Arts and Sciences Dean

q9: Dunn Hall

q10: to be or not to be

**Solution:**

The formula for precision and recall are:





For calculating precision, we are taking the top 100 retrieved documents from the list.

For this prediction, we are assuming the number of relevant documents for each query= 90

**Q1: “Computer Science”**

Retrieved documents: 100

Relevant documents: 72

**Precision:** 72/100= .72 ~ 72%

**Recall:** 72/90 = .80 ~ 80%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.72\*.8/ (.72+.8)

= 0.76

**Q2: “President of the university”**

Retrieved documents: 100

Relevant documents: 50

**Precision:** 50/100= .5 ~ 50%

**Recall:** 50/90= .56 ~ 56%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.5\*.56/ (.5+.56)

= 0.52

**Analysis:** a key reason for low precision would be there are many documents that contains the word “university”. Also, most of the university website pages, at the bottom, has a banner that contains “president’s office” text in it.

**Q3: “Information Retrieval”**

Retrieved documents: 100

Relevant documents: 70

**Precision:** 70/100= .7 ~ 70%

**Recall:** 70/90 = .78 ~ 78%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.7\*.78/ (.7+.78)

= 0.74

**Q4:** “**Learner Data Institute”**

Retrieved documents: 100

Relevant documents: 85

**Precision:**  85 /100= .85 ~ 85%

**Recall:** = 85/90 = .94 ~ 94%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.85\*.94/ (.85+.94)

= 0. 89

**Q5:** “**International student office”**

Retrieved documents: 100

Relevant documents: 75

**Precision:** 75 /100= .75 ~ 75%

**Recall:** = 75/90= .83 ~ 83%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.75\*.83/ (.75+.83)

= 0. 78

**Q6:** “**Graduate school admissions”**

Retrieved documents: 100

Relevant documents: 80

**Precision:** 80 /100= .80 ~ 80%

**Recall:** = 80/90= .89 ~ 89%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.8\*.89/ (.8+.89)

= 0. 84

**Q7:** “**What is the mascot of the University of Memphis?”**

Retrieved documents: 100

Relevant documents: 5

**Precision:** 5/100= .05 ~ 5%

**Recall:** =5/90=.06 ~ 6%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.05\*.06/ (.06+.05)

= 0.05

**Q8:** “**College of Arts and Sciences Dean”**

Retrieved documents: 100

Relevant documents: 85

**Precision:**  85 /100= .85 ~ 85%

**Recall:** = 85/90= .94 ~ 94%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.8\*.89/ (.8+.89)

= 0. 89

**Q9:** “**Dunn Hall”**

Retrieved documents: 100

Relevant documents: 73

**Precision:** 73 /100= .73 ~ 73%

**Recall:** = 73/90= .82 ~ 82%

**F1-measure =** 2\* (recall \*precision)/ (precision +recall)

= 2\*.73\*.82/ (.73+.82)

= 0. 77

**Q10:** **“to be or not to be”**

Retrieved documents: 42

Relevant documents: 0

**Precision:** 0/42= 0

**Recall:** 0

**F1-measure:** 0