|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **National University of Computer and Emerging Sciences, Lahore Campus** | | | | |
| C:\Users\saif\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.Word\final design.jpg | **Course:** | **Information Retrieval** | **Course Code:** | **CS317** |
| **Program:** | **BS(Computer Science)** | **Semester:** | **Fall 2018** |
| **Duration:** | **25 Minutes** | **Total Marks:** | **13** |
| **Paper Date:** | **7-Nov-18** | **Weight** | **3.3%** |
| **Section:** | **A** | **Page(s):** | **2** |
| **Exam:** | **Quiz 2** | **Roll No:** |  |

**Question1:**

What are reasons for normalizing DCG score. [3 Marks]

**Solution:**

1. To get score between 0 and 1.
2. DCG can get good value for easier queries with large number of relevant documents and bad score for difficult queries. Normalized DCG is not influenced by number of relevant documents for a query. NDCG score tell us how well the search engine performed as compared to the best possible performance.

**Question2:**

Consider following collection of 3 documents. [5 Marks]

|  |  |
| --- | --- |
| Document | Words |
| D1 | |  | | --- | | a b b a b b c | |
| D2 | |  | | --- | | a a b a b a | |
| D3 | |  | | --- | | b b b b b b c c | |

Query = < a b >

Use Dirichelt smoothing to find similarity of document D1 with query. Mu = 5

**Solution:**

N = 7, mu = 5

Prob (a) = 7/12(2/7) + 5/12(6/21) = 0.29

Prob (b) = 7/12(4/7) + 5/12(12/21) = 0.57

Score f D1 = 0.29 \* 0.57 = 0.17

**Question3:**

1. Computer average precision of following list of documents. [3]

R N N R N R N R N

Leftmost document is top ranked document. Total relevant documents are 15

**Solution:**

(1 + 2/4 + 3/6 + 4/8 )/15 = 0.167

1. Does average precision also measure recall? Justify your answer. [2]

**Solution:**

Yes, because precision is divided by total number of relevant documents. If query has large number of relevant documents but only few are retrieved then average precision will be low since we will divide the sum of precisions by total number of relevant documents.