Indian Institute of Information Technology Sri City, Chittoor

(An Institute of National Importance under Act of Parliament)

End Semester Examination - Monsoon 2021

[30 November 2021 - 08:30 AM - 11:00 AM]

Course: **Information Retrieval**   Total Marks: **20 + 20 (10 x 1 + 5 x 2 = 20 - MCQs) Marks**
Category: **Set-IR01**   Time: **60 + 20 (MCQs) Minutes = 80 MInutes**

Instructions:

a) The actual examination is scheduled for 60 + 20 minutes and another 15 minutes would be given for organizing the answer scripts, scanning and submitting over online.

b) Precisely answer the questions with relevant details. Avoid writing unnecessary explanations.

c) The file to be uploaded should be named as follows: **ABCD-YYYY-endsem-IR2021-AYCOY.pdf** where **ABCD** is the last 4 digits of your roll number; **YYYY** - year of admission (probably either 2018 or 2019); **AYCOY** - any 5 characters in CAPITAL letters (this may act as a secret key - Do not share with others). Please use hyphen (-) and NOT the "underscore" (_).

d) Submission portal: **http://smartmiss.iiits.ac.in/upload** and choose "**Information Retrieval**" as the course name.

Descriptive Questions: Answer ALL Questions [Please state your assumptions, if any]

1) [6 Marks]
   a) What is Relevance Feedback and how is it different from Pseudo Relevance Feedback?
   b) Illustrate RF approach using an example from any online store specifically for identifying the trending likes / dislikes of mobile phones

2) [8 Marks] Consider the following term - document matrix where N = 10 documents

	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
Antony	1	3	2	4	1	1	3	2	1	0
Brutus	2	3	1	1	0	0	2	1	2	1
Caesar	1	2	1	0	1	0	0	1	1	0
Calpurnia	1	3	2	1	1	3	1	1	0	1
Cleopatra	2	1	3	2	1	1	2	0	2	0
Rome	3	2	1	0	2	1	2	0	1	2
banquet	2	1	2	1	0	2	1	1	2	0
infinite	1	1	3	0	1	1	1	1	0	1
garlands	3	2	2	2	3	0	0	1	1	0
Octavius	0	3	1	3	1	2	3	2	0	1

Now show the ranked list of documents for the following queries

   a) Antony calls Cleopatra
   b) Octavius garlands Calpurnia

3) [6 Marks] Draw a Precision-Recall (PR) curve for the following ranked list of documents:
   a) Query - 1

   | A | B | C | D | E | F | G | H | I | J |

   b) Query - 2

   | A | B | C | D | E | F | G | H | I | J |

   Here the shaded documents are relevant documents and others are non-relevant.