



PRESIDENCY UNIVERSITY

BENGALURU

Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Mid - Term Examinations - MARCH 2026

Date: 10 - 03- 2026

Time: 02:00pm - 03:30pm

School: SOCSE	Program: B.TECH		
Course Code : ISE2502	Course Name: Information Retrieval		
Semester: VI	Max Marks: 50	Weightage: 25%	

CO - Levels	C01	C02	C03	C04	C05
Marks	26	24	0	0	NA

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks.

5Q x 2M=10M

1	Define Information Retrieval.	2 Marks	L1	C01
2	Differentiate Stemming and Lemmatization	2 Marks	L2	C01
3	Explain the IDF of term t.	2 Marks	L2	C02
4	What is the advantage of TF-IDF weighting scheme over TF-weighting scheme.	2 Marks	L2	C02
5	Explain dictionary based lemmatization.	2 Marks	L2	C01

Part B

Answer the Questions.

Total Marks 40M

6.	a.	Explain working of Boolean Retrieval model with suitable example.	10 Marks	L2	CO1
	b.	Explain Information Retrieval model with suitable diagram.	10 Marks	L2	CO1
Or					
7.	a.	Describe the steps involved to clean text data.	10 Marks	L2	CO1
	b.	Describe the types of automatic Indexing.	10 Marks	L2	CO1

8.	a.	Explain BIM (Binary Independence Model) in detail	10 Marks	L2	CO2
	b.	D1: "data mining techniques" D2: "data mining mining" D3: "mining techniques" Consider these documents and create a document term matrix with TF-IDF weighting scheme with TF raw count.	10 Marks	L3	CO2
Or					
9.	a.	Explain types of Term Frequencies with examples.	10 Marks	L2	CO2
	b.	Consider the following documents and a query: D1: "information retrieval system" D2: "information system system" D3: "retrieval system" Query (Q): "information retrieval" Find the cosine similarity of all these documents w.r.t. Query Q. Use TF-IDF weighting scheme with TF raw count.	10 Marks	L3	CO2