

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Jan/Feb 2025

~~Nov-Dec 2024~~

(B. Tech) Program: Computer Engineering Scheme: II **B**
Supplementary Regular Examination: LY Semester: VII
Course Code: CEDLC7044 and Course Name: Information Retrieval
Date of Exam: ~~23-11-2024~~ Duration: 02.5 Hours Max. Marks: 60

31-01-2025

Instructions:

- (1) All questions are compulsory.
- (2) Draw neat diagrams wherever applicable.
- (3) Assume suitable data, if necessary.

Q. No.	Question	Max. Marks	CO	BT level
Q 1	Solve any two questions out of three: (05 marks each)	10		
a)	Illustrate any one of the query protocol in detail.		CO3	U
b)	What is TF-IDF (term frequency-inverse document frequency)? How is TF-IDF calculated? illustrate this with a good example.		CO2	U
c)	Explain the following concepts w.r.t. searching the data I) Single Word Queries II) Multiple Word Queries III) Phrase Queries IV) Proximity Queries V) Boolean Queries		CO5	U
Q 2	Solve any two questions out of three: (05 marks each)	10		
a)	Build encoding using the 'Huffman Code' technique for the following file "Eerie eyes seen near lake."		CO4	Ap
b)	Differentiate and analyze the centralized crawler-indexer architecture with Harvest architecture. Justify your analysis.		CO5	An
c)	Discuss the 'Global Query Refinement' technique in detail with the help of a suitable example.		CO3	U
Q.3	Solve any two questions out of three. (10 marks each)	20		
a)	Draw and explain the Retrieval Process in detail.		CO1	U
b)	Discuss the conditions on multimedia data and explain the multimedia IR system		CO2	U
c)	i) Explain in detail the 'Information access process'. 5M ii) Related to the Search Process, illustrate the following concepts. 5M		CO6	U

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Jan / Feb 2023

~~Nov - Dec 2024~~

Supplementary (B. Tech) Program: Computer Engineering Scheme: II
Regular Examination: LY Semester: VII
Course Code: CEDLC7044 and Course Name: Information Retrieval
Date of Exam: 31-01-25 Duration: 02.5 Hours Max. Marks: 60

	x) Starting points y) Relevance feedback z) Interface Support			
Q.4	Solve any two questions out of three. (10 marks each)	20		
a)	What is Rocchio's relevance feedback algorithm? Apply this to any suitable data and find the modified query. (Assume $\alpha = 1$, $\beta = 0.75$, $\gamma = 0.25$.)		CO3	Ap
b)	With the use of the Boyer-Moore matching algorithm, find the pattern position in the given string below. String :- bacbabababacaca Pattern:- ababaca		CO5	Ap
c)	Create and analyze the 'Elias Delta' and 'Elias Gamma' coding and decoding forms separately on the following given numbers: a)2 b)8 c)9 d)10		CO4	An
