

Code No: 155EN**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****B. Tech III Year I Semester Examinations, August/September - 2024****INFORMATION RETRIEVAL SYSTEMS****(Common to CE(SE), CSE(AI&ML), CSE(DS), AI&DS, AI&ML)****Time: 3 Hours****Max. Marks: 75**

- Note:** i) Question paper consists of Part A, Part B.
ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.
iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A**(25 Marks)**

- 1.a) Define Information Retrieval Systems. [2]
- b) List out the objectives of IRS. [3]
- c) Explain Information Extraction. [2]
- d) Briefly explain N- Gram data structure. [3]
- e) What is statistical indexing? [2]
- f) List out the advantages of automatic indexing. [3]
- g) Define negative feedback. [2]
- h) Discuss on similarity measures. [3]
- i) What is multi-attribute retrieval? [2]
- j) What are the five elements of finite state automata used in text searching algorithms? [3]

PART – B**(50 Marks)**

- 2.a) Explain Miscellaneous Capabilities of IRS.
- b) Explain in detail about the components of Information Retrieval System with neat diagram. [5+5]

OR

- 3.a) Discuss the relationship between information retrieval systems and database management systems.
- b) Discuss the Search Capabilities of IRS in detail. [5+5]

- 4.a) What are the objectives of Indexing? Explain the process of Indexing.
- b) List and explain the various types of automatic indexing. [5+5]

OR

- 5.a) Explain the stemming algorithms in detail.
- b) Explain in detail about Inverted File Structure. [5+5]

- 6.a) Explain in detail about the classes of Automatic Indexing.
b) Discuss the hypertext linkages. [5+5]

OR

- 7.a) Explain the need and importance of weighting scheme for automatic indexing and the problems associated with the weighting scheme.
b) Consider the following term-term matrix:

	T1	T2	T3	T4	T5	T6
T1		15	6	8	12	14
T2	15		12	10	6	8
T3	6	12		16	4	10
T4	8	10	16		9	4
T5	12	6	4	9		13
T6	14	8	10	4	13	

- i) Determine the Term Relationship matrix using a threshold of 10 or higher
ii) Determine the clusters using the clique technique. [5+5]
- 8.a) What are the technologies used in information visualization? Discuss.
b) Discuss the significance of relevance feedback in ranking the documents. [5+5]

OR

- 9.a) Explain the Selective Dissemination of Information Search.
b) Briefly describe the aspects of the visualization process. [5+5]

- 10.a) Describe the features of Sage book for graph retrieval.
b) Explain the software text search algorithms in detail. [5+5]

OR

- 11.a) What are hardware text search algorithms? Explain them in detail?
b) Demonstrate Boyre-Moore Algorithm for the following scenario, explain each step.
String to be searched: abcac
Input String: ababdcabdacabcac. [5+5]

---ooOoo---