

Code No: 155EN

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD**B. Tech III Year I Semester Examinations, March - 2024****INFORMATION RETRIEVAL SYSTEMS****(Computer Science and Engineering – Artificial Intelligence and Machine Learning)****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 precision. [2]
- b) What do you mean by proximity search? [3]
- c) What is meant by private index? [2]
- d) What is the scope of indexing? [3]
- e) Give the term frequency weighting formula used in TREC-4. [2]
- f) What are the classes of automatic indexing? [3]
- g) How to find similarity among clusters of items? [2]
- h) What are the different levels of bindings? [3]
- i) What are the components of hardware text search unit? [2]
- j) Give text streaming architecture. [3]

PART – B**(50 Marks)**

2. Explain the objectives and architecture of an information retrieval system. [10]
OR
3. Discuss the significance of Thesaurus and its expansion for searching activity in IRS. [10]
4. Describe the use of Bayesian networks in indexing with suitable examples. [10]
OR
- 5.a) Illustrate the functionality of Porter stemmer.
b) How a trigram data structure can be used in searching a term in a document? Explain with an example. [5+5]
6. Explain the concept of inverse document frequency for automatic indexing with necessary illustrations. [10]
OR
7. Describe any two techniques for the automatic generation of term clusters to create statistical thesauri. [10]

- 8.a) How is HMM considered a user search technique?
b) For the following query compute the Cosine, Jaccard and Dice similarity measures with the documents.
QUERY = (2, 2, 0, 0, 4)
DOC1 = (0, 2, 6, 4, 0)
DOC2 = (2, 6, 0, 0, 4) [5+5]

OR

- 9.a) Discuss the six characteristics of intelligent agents.
b) Illustrate the following visualization technologies:
Envision interface, tree map, cone tree. [5+5]

10. Demonstrate Knuth-Morris-Pratt Algorithm for the following scenario, explain each step.
String to be searched : abcac
Input String : ababdcabedacabcac [10]

OR

- 11.a) What kind of features can be used in indexing of audio data?
b) Discuss the challenges in video retrieval system. [5+5]

---ooOoo---

March-2024