Total No. of Questions : 4]

SEAT No. : [ ]

P5245

[Total No. of Pages : 2

[6188] - 200
B.E. (I.T.) (Insem)
INFORMATION & STORAGE RETRIEVAL
(2019 Pattern) (Semester - VII) (414441)

Time : 1 Hour] ***[Max. Marks : 30***

Instructions to the candidates:

1) ***Answer Q1 or Q2, Q3 or Q4.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right side indicate full marks.***
4) ***Assume suitable data, if necessary.***

Q1) a) Analyse the concept behind the single pass algorithm and the single link algorithm for clustering. Compare and contrast these two techniques, highlighting their advantages, disadvantages, and situations where each approach would be more suitable. Provide a scenario for each algorithm to illustrate their application. **[8]**

b) Differentiate between data retrieval and information retrieval. **[4]**

c) Define the following terms. **[3]**

i) Precision

ii) Recall

OR

Q2) a) Explain Conflation algorithm to generate document representative of a document with a proper example. **[8]**

b) What are the different measures of association? Explain any three matching coefficients with suitable examples. **[7]**

P.T.O.

Q3) a) Demonstrate the application of the Boolean search technique using a set of documents and a complex Boolean query involving multiple operators (AND, OR, NOT). **[5]**

b) Explain the concept of Inverted index file. How it can be used in Information Retrieval. **[6]**

c) List and explain the types of queries. **[4]**

OR

Q4) a) Explain exhaustivity and specificity with respect to Index term weighting. **[4]**

b) Compare and contrast the basic concepts, strengths, and limitations of the Boolean Model, Vector Model, and Probabilistic Model. Provide insights into when and why each model would be chosen to optimize search results, considering factors like ranking accuracy and complexity. **[8]**

c) Elaborate cluster-based retrieval in brief. **[3]**

❄ ❄ ❄