Total No. of Questions : 10]

SEAT No. :

P3647

[Total No. of Pages :2

[4959]-1138

B.E.(Information Technology)

c: INFORMATION STORAGE AND RETRIEVAL

(2012 Course) (Semester-II)(Elective-III) (414463)

Time :2½Hours] ***[Max. Marks : 70***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right indicate full marks.***
4) ***Assume suitable data if necessary.***

Q1) a) Explain Luhn's idea in details. **[5]**

b) What is clustering? Explain the use of clustering in IR. **[5]**

OR

Q2) Clusters the documents using single pass clustering algorithm for the following example. Threshold value is 10 **[10]**

	Terms in document				
	T1	T2	T3	T4	T5
Doc 1	1	2	0	0	1
Doc 2	3	1	2	3	0
Doc 3	3	0	0	0	1
Doc 4	2	1	0	3	0
Doc 5	2	2	1	5	1

Q3) a) Explain Boolean model in detail **[5]**

b) Give the difference between suffix array and suffix tree. **[5]**

OR

P.T.O.

Q4) a) Explain the term precision and recall and calculate the same for the following example **[5]**

A set of relevant documents for query

q = {d3,d7,d8,d11,d14,d19,d23,d25}

A new retrieval algorithm returns following answer set

= {d1,d2,d3,d7,d9,d10,d14,d20,d23,d24,d25}

b) Explain the terms Harmonic mean, E measure, R precision, Precision histogram **[5]**

Q5) a) Explain GEMINI approach for multimedia IR. **[9]**

b) How are queries processed in distributed IR. **[9]**

OR

Q6) a) Write a note on MULTOS **[9]**

b) Describe multimedia data support in commercial DBMS. **[9]**

Q7) a) Discuss challenges involved in web searching. **[8]**

b) Explain crawler indexer architecture in details. **[8]**

OR

Q8) a) Write a note on characterizing the web **[8]**

b) What are meta crawlers? Explain with suitable example **[8]**

Q9) a) What is collaborative filtering? Discuss its advantages and disadvantages **[8]**

b) Explain semantic web in details **[8]**

OR

Q10)a) Explain the method for extracting data from text **[8]**

b) Explain Collecting and Integrating Specialized Information on the web. **[8]**