R15

Code No: 825AD

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
MCA V Semester Examinations, June/July - 2018
INFORMATION RETRIEVAL SYSTEMS

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART - A

5 × 5 Marks = 25

1.a) What are the major steps in inverted index construction? [5]
b) Write a short note on term frequency and inverse document frequency. [5]
c) Give an example to show the tree representation of XML documents and queries. [5]
d) What is the minimum number of support vectors required for a data set? Explain. [5]
e) State and explain the concept of web graph. [5]

PART - B

5 × 10 Marks = 50

2. How do we process a query using an inverted index and the basic Boolean retrieval model? Explain with example. [10]

OR

3. Explain various tasks involved in determining the vocabulary of terms. [10]

4. State and explain the basic algorithm for computing vector space scores. [10]

OR

5. Describe how to measure effectiveness of adhoc information retrieval systems. [10]

6. Write an algorithm for scoring documents with SimNoMERGE. [10]

OR

7. Explain in detail about the query likelihood models in information retrieval. [10]

8. List and explain the issues involved in the classification of text documents. [10]

OR

9. Give the applications of clustering in information retrieval system. [10]

10. With an example explain the concept of near duplicates and shingling's. [10]

OR

11. With a neat diagram explain the basic architecture of a crawler. [10]

---oo0oo---