

R13**Code No: 815BJ****JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****MCA V Semester Examinations, August - 2017****INFORMATION RETRIEVAL SYSTEMS****Time: 3 Hours****Max. Marks: 60****Note:** This question paper contains two parts A and B.

Part A is compulsory which carries 20 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 8 marks and may have a, b, c as sub questions.

PART - A**5 × 4 Marks = 20**

- 1.a) What is IR and provide an example for IR problem? [4]
- b) Describe the procedure for 'pseudo relevance feedback'. [4]
- c) Provide some basic XML concepts. [4]
- d) What is "Low rank approximations" in matrix decompositions? [4]
- e) What is the role of Web graphs in link analysis? [4]

PART - B**5 × 8 Marks = 40**

2. How do we process a query using an inverted index and the basic Boolean retrieval model? Explain with example. [8]

OR

3. Explain the following spelling correction strategies in brief.
a) K-gram indexes b) context sensitive. [4+4]

4. Write short notes on the following "parametric and zone indexes".
a) Weighted zone scoring b) The optimal weighting. [4+4]

OR

5. Describe "Rocchio algorithm for relevance feedback". [8]

- 6.a) Compare text-centric and data-centric XML retrieval.
b) What is "The 1/0 loss case" probability ranking principle? [4+4]

OR

7. Describe about "Query likelihood models in language models" for Information retrieval. [8]

8. Write about the following models of extended support vector machines.
a) Soft margin classification b) Multi class SVMs. [4+4]

OR

9. Describe the following Hierarchical clustering methods.
a) Centroid clustering b) Divisive Clustering. [4+4]

10. Explain the role of "Index size and clustering" in approximation of web search comprehensiveness. [8]

OR

11. What is Web Crawling? Draw and explain the basic Web Crawler architecture. [8]