Code No: 5258BP

R15

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
M. Tech II Semester Examinations, June/July - 2019
INFORMATION RETRIEVAL SYSTEMS
(Computer Science and Engineering)

Time: 3hrs   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)   Define precision and recall with suitable examples.   [5]
   b)   Explain inverse document frequency with an example.   [5]
   c)   Explain the differences between text centric xml retrieval and data centric xml retrieval.
                                                                            [5]
   d)   Explain K-means clustering algorithm.   [5]
   e)   Explain the URL frontier.   [5]

PART - B

5 × 10 Marks = 50

2.a)   Define wildcard queries. Explain k-gram indexes for wildcard queries.
  b)   Explain blocked sort-based indexing with an example.   [5+5]
                                    OR
3.a)   Write the advantages and limitations of distributed indexing.
  b)   Explain Dynamic Indexing with an example.   [5+5]

4.   Explain the following with an example.
       a) Precision recall curve
       b) Interpolated precision
       c) Eleven point interpolated average precision
       d) R-precision
       e) ROC curve.   [10]
                                    OR
5.   Explain Rocchio algorithm for relevance feedback.   [10]

6.   Explain Bayesian network approaches to information retrieval.   [10]
                                    OR
7.a)   Explain the following
       i) X^2 feature selection
       ii) Frequency based feature selection
       iii) Feature selection for multiple classifiers
                                                                            [6+4]
  b)   Explain about SVMs.

8.a) Explain centroid clustering algorithm.
  b) Explain group average agglomerative clustering algorithm. [5+5]

OR

9.a) Explain machine learning methods in ad hoc information retrieval.
  b) Explain hierarchical agglomerative clustering algorithm. [5+5]

10. Draw and explain the crawler architecture. [10]

OR

11. Explain about distributed indexes. [10]

---ooOoo---