[BE SEM VII Cchime bad KT Summer 2026] CMPN

Duration: - 3 Hours **Marks: 80 Marks**

NB: - Question 1 is compulsory
Solve any four questions from Question no. 1 .
Solve any three questions from the remaining.

1 a. Explain the difference between information and data with suitable examples. How does this distinction influence the design of IR systems? 20 (4x5)

b. List steps of the IR retrieval process. Identify difference between digital and traditional libraries.

c. Identify and explain the main advantages of the vector space model over Boolean models in IR

d. What are different types of queries in Information Retrieval?

e. Define Multimedia IR and Text IR. State one key difference and one challenge.

2 a. What are the various categories of information retrieval models? Provide a detailed explanation of probabilistic models. 10

b. Compare the advantages of Algebraic models over Classical IR models. Can algebraic models be more suitable in certain contexts? 10

3 a. Explain the process of user relevance feedback and how it improves retrieval performance. Illustrate with a sample interaction loop. 10

b. Explain the role of a suffix tree in the indexing of Information Retrieval, Pattern matching, Structured queries with example 10

4 a. Compare and contrast the use of parametric indexes and zone indexes in optimizing retrieval in structured text databases. Identify one scenario where each is preferred. 10

b. Explain how Tf-idf weighting helps in ranking documents. Include formula and example. 10

5 a. How do precision and recall metrics differ when evaluating the effectiveness of ranked and unranked retrieval results? 10

b. Explain the architecture of a Distributed Information Retrieval system. What are its advantages?

6 Write short notes on any two 20

a. Inexact top K document retrieval

b. Automatic local analysis vs Automatic global analysis

c. The strengths and weaknesses of Probabilistic Latent Semantic Indexing

d. Flat browsing vs hypertext browsing model.

****