113

Reg. No. : | | | | | | | | | | | | |

Question Paper Code : 90397

B.E./B.Tech. DEGREE EXAMINATIONS, NOVEMBER/DECEMBER 2022.

Seventh/Eighth Semester

Computer Science and Engineering

CS 8080 – INFORMATION RETRIEVAL TECHNIQUES

(Common to: Computer and Communication Engineering/ Information Technology)

(Regulations 2017)

Time : Three hours Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Give the components of Search Engine and the Performance measures.
2. State the purpose of Query Interface.
3. What is meant by Zone Index?
4. Construct the Vector space model representation.
5. Define meta crawler. List any two major meta crawlers.
6. Mention the use of implementing own DNS resolver as a component of the web crawler.
7. State the discrete-time stochastic process.
8. What is the use of blockrank?
9. Differentiate supervised learning and unsupervised learning.
10. Mention the use of dendrogram.

PART B — (5 × 13 = 65 marks)

11. (a) (i) Demonstrate the working of IR architecture with a diagram. (6)

(ii) Discuss about how N-gram data structure can be used for stemming. (7)

Or

(b) (i) Compare in detail Information Retrieval and Web Search with examples. (8)

(ii) Describe the various components of a Search Engine. (5)

12. (a) (i) Explain in detail about binary independence model for Probability Ranking Principle (PRP). (7)

(ii) Analyze how the query generation probability for query likelihood model can be estimated. (6)

Or

(b) (i) Explain Latent Semantic Indexing and latent semantic space with an illustration. (8)

(ii) Analyze the use of LSI in Information Retrieval. What is its need in synonyms and semantic relatedness? (5)

13. (a) (i) Evaluate the Agglomerative Clustering and HAC in detail. (7)

(ii) Discuss the Types of data and evaluate it using any one clustering techniques. (6)

Or

(b) (i) Analyze the working of Nearest Neighbor algorithm along with one representation. (7)

(ii) Analyze the K-Means Clustering method and the problems in it. (6)

14. (a) Develop a web search structure for searching a newly hosted web domain by the naïve user with step by step procedure.

Or

(b) Describe the following with suitable examples:

(i) Bag of Words and Shingling (7)

(ii) Hashing Min Hash and Sim Hash (6)

115

15. (a) Illustrate the advantages and disadvantages of Content based and collaborative filtering recommendation system.

Or

(b) (i) Describe the rules of High Level Architecture. (7)

(ii) Difference between Hybrid and Collaborative Recommendation. (6)

PART C — (1 × 15 = 15 marks)

16. (a) (i) Grade the optimization techniques available for search engine and rank them by your justification. (9)

(ii) Present and anlyse Web Crawler Taxonomy in detail. (6)

Or

(b) Demonstrate Recommendation based on User Ratings using appropriate example. Narrate in detail about a model for Recommendation system. (15)