Reg. No.															

Question Paper Code	13393

B.E. / B.Tech. - DEGREE EXAMINATIONS, APRIL / MAY 2025

Seventh Semester

Computer Science and Engineering

(Common to Eighth Semester - Computer and Communication Engineering)

20CSEL703 - INFORMATION RETRIEVAL TECHNIQUES

Regulations - 2020

Duration: 3 Hours Max. Marks: 100

PART - A (MCQ) (10 × 1 = 10 Marks)

Answer ALL Questions

		Marks	*K – Level*	*CO*
1.	What is the primary problem of IR? (a) Storing data in memory (b) Retrieving exact copies of documents (c) Finding relevant documents for a user query (d) Compressing large databases	*1*	*K1*	*CO1*
2.	Which component of IR software handles indexing of documents? (a) Parser (b) Indexer (c) Query Processor (d) Ranker	*1*	*K1*	*CO1*
3.	State the assumption of probabilistic model. (a) Users always know what they want (b) Document relevance can be predicted using probability (c) All documents are equally important (d) Relevance is binary and fixed	*1*	*K1*	*CO2*
4.	List the main use of Neural network models in IR. (a) Static document storage (b) Predicting document structure (c) Learning patterns in relevance from training data (d) Indexing only image data	*1*	*K1*	*CO2*
5.	Which of the following is a common application of text classification? (a) Image segmentation (b) Audio filtering (c) Spam detection in emails (d) Memory allocation	*1*	*K1*	*CO3*
6.	What determines the splitting of nodes in a decision tree? (a) Graph structure (b) Entropy or Gini Index (c) Euclidean distance (d) Hyperplane separation	*1*	*K1*	*CO3*
7.	Outline the goal of the algorithm in K-means clustering. (a) Maximize distances between clusters (b)Minimize intra-cluster distances (c) Maximize cluster sizes (d) Eliminate all features	*1*	*K2*	*CO4*
8.	Define Accuracy in a classification model. (a) Number of false predictions / Total predictions (b) Correct predictions / Total predictions (c) True positives / Total negatives (d) False positives / True negatives	*1*	*K1*	*CO4*
9.	Summarize about data in a cluster-based architecture. (a) Stored in one central server (b) Ignored during retrieval (c)Distributed across multiple machines (d) Always encrypted	*1*	*K2*	*CO5*
10.	What types of data do recommender systems primarily rely on? (a) User behavior and preferences (b) Climate data (c) Programming syntax (d) Network topology	*1*	*K1*	*CO6*

PART - B (12 × 2 = 24 Marks)

Answer ALL Questions

11.	Differentiate between data retrieval and information retrieval.	*2*	*K2*	*CO1*
12.	Outline about the type of natural language technology used in information retrieval.	*2*	*K2*	*CO1*
13.	What is a Boolean model?	*2*	*K1*	*CO2*
14.	Summarize about sparse vector.	*2*	*K2*	*CO2*

15.	Compare Supervised and Unsupervised algorithm.		*2*	*K2*	*CO3*
16.	List the advantages and disadvantages of Decision Tree algorithm.		*2*	*K1*	*CO3*
17.	Interpret about Hash-based Dictionary in Indexing.		*2*	*K2*	*CO4*
18.	Describe Brute Force in Sequential Search.		*2*	*K2*	*CO4*
19.	Illustrate the challenges in data traversing by search engines.		*2*	*K2*	*CO5*
20.	Identify the applications of web crawlers.		*2*	*K2*	*CO5*
21.	Define knowledge based Recommendation.		*2*	*K1*	*CO6*
22.	List some advantages of Mobile Recommendation system.		*2*	*K1*	*CO6*

PART - C (6 × 11 = 66 Marks)

Answer ALL Questions

23.	a)	Describe the working of IR architecture with a diagram.	*11*	*K2*	*CO1*
		OR			
	b)	Explain the role of Artificial Intelligence in Information Retrieval Systems.	*11*	*K2*	*CO1*
24.	a)	Develop a method for Relevance Feedback and query expansion to improve search results in an IR system.	*11*	*K3*	*CO2*
		OR			
	b)	Develop a simple Latent Semantic Indexing (LSI) model for a sample set of documents.	*11*	*K3*	*CO2*
25.	a)	Develop a methodology for implementing a K-Nearest Neighbors (KNN) classifier for text data.	*11*	*K3*	*CO3*
		OR			
	b)	Construct a Decision Tree classifier for a sample text classification task. Develop the tree structure based on a set of text features, and explain the process of training and classifying new text samples.	*11*	*K3*	*CO3*
26.	a)	Develop a hierarchical clustering approach using agglomerative clustering for a sample dataset.	*11*	*K3*	*CO4*
		OR			
	b)	Construct a K-means clustering model and then apply dimensionality reduction before clustering.	*11*	*K3*	*CO4*
27.	a)	Design and develop a Web search Architecture and the components of search engine and its issues.	*11*	*K3*	*CO5*
		OR			
	b)	Identify about the working of WEB CRAWLER with its diagram.	*11*	*K3*	*CO5*
28.	a)	Explain Collaborative Recommendation System in detail.	*11*	*K2*	*CO6*
		OR			
	b)	Discuss the high-level architecture of a content-based recommender system.	*11*	*K2*	*CO6*