

M. Sc. DEGREE END SEMESTER EXAMINATION - OCTOBER 2025**SEMESTER 3 - COMPUTER SCIENCE (ARTIFICIAL INTELLIGENCE)****COURSE: 24P3CAIT12: DATA MINING AND ANALYTICS***(For Regular 2024 Admission)*

Time: Three Hours

Max. Weightage: 30

PART - A**Answer any 8 Questions.****Weight: 1**

1. What is an axis-parallel decision tree? (An, CO2)
2. What is K-Medoids clustering? (E, CO3)
3. What is single-document summarization? (An, CO4)
4. Explain the concept of F-Measure in classification evaluation. (An, CO2)
5. What is the Semantic Web? (An, CO5)
6. What is the primary purpose of data preparation in data analytics? (An, CO1)
7. Define density-based clustering. (An, CO3)
8. State one application of link analysis in web mining. (Cr, CO5)
9. Define cosine similarity in text processing. (An, CO4)
10. Define data cleaning and illustrate with an example. (An, CO1)

(1 x 8 = 8)**PART – B****Answer any 6 Questions.****Weights: 2**

11. Describe different types of recommendation systems. (Cr, CO5)
12. a) Apply a confusion matrix to compute precision, recall and F1-score for a classifier. (An, CO2)
b) Calculate precision, recall, and F1-score from the following confusion matrix:
True Positive = 45, False Positive = 5, False Negative = 15, True Negative = 35.
13. Discuss the concept of Boolean retrieval with an example. (An, CO4)
14. Explain how PageRank algorithm ranks web pages (Cr, CO5)
15. Explain the architecture of a data warehouse. (An, CO1)
16. Describe the steps of hierarchical clustering with an example. (E, CO3)
17. Compare DBSCAN and BIRCH algorithms with respect to noise handling. (E, CO3)
18. Explain the importance of text summarization in modern information retrieval. (An, CO4)

(2 x 6 = 12)

PART - C**Answer any 2 Questions.****Weights: 5**

19. Design a Naïve Bayes based sentiment analysis model for product reviews. Explain preprocessing, training and evaluation steps. (An, CO2)
20. Evaluate the effectiveness of extractive vs. abstractive text summarization approaches. Provide examples and discuss their practical applications. (An, CO4)
21. Design a collaborative filtering recommendation system for an e-commerce platform. Explain user-based vs. item-based approaches. (Cr, CO5)
22. Design and implement a clustering strategy for customer segmentation in retail data. Choose an appropriate algorithm, justify your choice, and explain expected outcomes. (E, CO3)

(5 x 2 = 10)**OBE: Questions to Course Outcome Mapping**

CO	Course Outcome Description	CL	Questions	Total Wt.
CO1	Explain the fundamental paradigms of machine learning and the principles of density estimation techniques.	An	Q6,Q10,Q15	4
CO2	Apply dimensionality reduction methods like PCA and SVD and implement classification algorithms such as Perceptron, Feed Forward Network, and SVM.	An	Q1,Q4,Q12,Q19	9
CO3	Analyze clustering techniques and regression models to discover patterns and predict outcomes from data.	E	Q2,Q7,Q16,Q17,Q22	11
CO4	Evaluate the efficiency of probabilistic models like Bayesian Networks, HMMs, and CRFs, and apply TensorFlow for machine learning tasks.	An	Q3,Q9,Q13,Q18,Q20	11
CO5	Evaluate the performance and applicability of deep learning architectures like CNNs, RNNs, and LSTMs in solving real-world problems.	Cr	Q5,Q8,Q11,Q14,Q21	11

Cognitive Level (CL): Cr - CREATE; E - EVALUATE; An - ANALYZE; A - APPLY; U - UNDERSTAND; R - REMEMBER;