

[4]

- (c) Multidimensional Data model
बहुआयामी डेटा मॉडल
- (d) Neural network-based algorithms
न्यूरल नेटवर्क आधारित एल्गोरिदम

Total No. of Questions : 8]

[Total No. of Printed Pages : 4

Enrolment No. 013.....

CD 703 (D) / CS 703 (B) (GS)

**B. Tech., VII Semester
Examination, December 2025**

Grading System (GS)

Data Mining and Warehousing

Time : Three Hours

Maximum Marks : 70

- Note :** (i) Attempt any *five* questions.
किन्हीं पाँच प्रश्नों को हल कीजिए।
- (ii) All questions carry equal marks.
सभी प्रश्नों के समान अंक हैं।
- (iii) In case of any doubt or dispute the English version question should be treated as final.
किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

1. (a) Define data warehousing. Explain the architecture of a data warehouse with a suitable diagram.
डेटा वेयरहाउसिंग को परिभाषित करें। उपयुक्त चित्र के साथ डेटा वेयरहाउस की संरचना की व्याख्या करें।
- (b) Describe the techniques used for data preprocessing in data warehousing.
डेटा वेयरहाउसिंग में डेटा प्रीप्रोसेसिंग के लिए उपयोग की जाने वाली तकनीकों का वर्णन करें।

[2]

2. (a) What are the different types of data warehouse schemas? Explain with examples.
डेटा वेयरहाउस स्कीमाओं के विभिन्न प्रकार कौन-कौन से हैं? उदाहरणों के साथ समझाइए।
- (b) Define OLAP and Compare different types of OLAP servers.
OLAP को परिभाषित करें और विभिन्न प्रकार के OLAP सर्वरों की तुलना करें।
3. (a) Describe backup and recovery strategies in data warehouse environments.
डेटा वेयरहाउस वातावरण में बैकअप और रिकवरी रणनीतियों का वर्णन करें।
- (b) Explain the role of Hardware components in Operational Design of Data Warehouse.
डेटा वेयरहाउस के संचालन डिजाइन में हार्डवेयर घटकों की भूमिका को समझाएँ।
4. (a) Define data mining and differentiate it from knowledge discovery in databases.
डेटा माइनिंग को परिभाषित करें और इसे डेटाबेस में ज्ञान की खोज से अलग करें।
- (b) Describe similarity measures and their role in clustering and classification.
समानता माप और उनके क्लस्टरिंग और वर्गीकरण में भूमिका का वर्णन करें।
5. (a) Discuss the key differences between supervised and unsupervised learning with examples.
सुपरवाइज्ड और अनसुपरवाइज्ड लर्निंग के बीच मुख्य अंतर पर उदाहरणों के साथ चर्चा करें।

[3]

- (b) Describe the k-Nearest Neighbors (k-NN) algorithm and its role in classification.
k-नियरस्ट नेबर्स (k-NN) एल्गोरिदम का वर्णन करें और वर्गीकरण में इसकी भूमिका बताएं।
6. (a) What are rule-based classifiers? Explain how they derive classification rules from data.
नियम-आधारित वर्गीकर्ता क्या हैं? समझाएं कि वे डेटा से वर्गीकरण नियम कैसे निकालते हैं।
- (b) Define probabilistic classifiers and explain how they differ why are they useful in uncertain environments?
सांख्यिकीय (प्रॉबैबिलिस्टिक) वर्गीकर्ताओं को परिभाषित करें और असंदिग्ध (अनिश्चित) परिस्थितियों में वे उपयोगी क्यों हैं?
7. (a) Define clustering and explain its importance in unsupervised learning.
क्लस्टरिंग को परिभाषित करें और अनसुपरवाइज्ड लर्निंग में इसके महत्व को समझाएं।
- (b) What are association rules? Differentiate Apriori and FP-Growth algorithms.
एसोसिएशन रूल्स क्या हैं? Apriori और FP-Growth एल्गोरिदम में अंतर बताइए।
8. Write short notes on any two :
किसी भी दो पर संक्षिप्त नोट्स लिखें :
- (a) OLAP Queries and Operations
OLAP क्वेरी और संचालन
- (b) Summary statistics and Data distributions
सारांश सांख्यिकी और डेटा वितरण