

Roll No

CS-703 (B) (GS)**B.Tech., VII Semester**

Examination, June 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) Give detailed information about Meta data in data warehousing. 7

डाटा वेयरहाउसिंग में मेटा डाटा के बारे में विस्तृत जानकारी दें।

- b) Explain the role played by sourcing, acquisition, clean up and transformation tools in data warehousing. 7

डाटा वेयरहाउसिंग में सोर्सिंग, अधिग्रहण, सफाई और परिवर्तन उपकरणों द्वारा निभाई गई भूमिका की व्याख्या करें।

2. a) Write in brief about schemas in multidimensional data model. 7

बहुआयामी डाटा मॉडल में स्कीमा के बारे में संक्षेप में लिखें।

- b) Describe in brief about data warehouse implementation. 7

डाटा वेयरहाउस कार्यान्वयन के बारे में संक्षेप में वर्णन करें।

3. a) Write the difference between multi-dimensional OLAP (MOLAP) and Multi-relational OLAP (ROLAP). 7

मल्टी-डायमेंशनल OLAP (MOLAP) और मल्टी-रिलेशनल OLAP (ROLAP) के बीच अंतर लिखें।

- b) Discuss in detail about the case study of Data warehouse for the Government of Madhya Pradesh. 7

मध्य प्रदेश सरकार के लिए डाटा वेयरहाउस की केस स्टडी के बारे में विस्तार से चर्चा करें।

4. a) What is KDD? Explain about data mining as a step in the process of knowledge discovery. 7

KDD क्या है? ज्ञान खोज की प्रक्रिया में एक कदम के रूप में डाटा माइनिंग के बारे में बताइए।

- b) Discuss about data mining task primitives with examples. 7

डाटा माइनिंग टास्क प्रिमिटिव के बारे में उदाहरण सहित चर्चा करें।

5. a) Explain major requirements and challenges in data mining. 7

डाटा माइनिंग में प्रमुख आवश्यकताओं और चुनौतियों की व्याख्या करें।

- b) Explain the data mining functionalities. 7

डाटा माइनिंग कार्यप्रणाली की व्याख्या करें।

6. a) Explain whether association rule mining is supervised or unsupervised type of learning. 7
 बताएं कि क्या एसोसिएशन नियम mining पर्यवेक्षित या अपर्यवेक्षित प्रकार का शिक्षण है।
- b) Explain briefly about statistical distribution-based outlier detection. 7
 सांख्यिकीय वितरण-आधारित बाह्य पहचान के बारे में संक्षेप में बताइए।
7. a) What is data cleaning? What are the different techniques for handling missing values? 7
 डाटा सफाई क्या है? Missing मूल्यों को संभालने के लिए विभिन्न तकनीकें क्या हैं?
- b) Briefly explain the differences between "classification" and "Clustering" and give an informal example of an application that would benefit from each technique. 7
 "वर्गीकरण" और "क्लस्टरिंग" के बीच अंतर को संक्षेप में समझाइए और एक ऐसे एप्लिकेशन का अनौपचारिक उदाहरण दें जो प्रत्येक तकनीक से लाभान्वित होगा।
8. a) Explain the following clustering methods in detail. 7
 निम्नलिखित क्लस्टरिंग विधियों को विस्तार से समझाइए।
 i) BIRCH
 ii) CURE
- b) Explain different data types used in clustering. 7
 क्लस्टरिंग में प्रयुक्त विभिन्न डाटा प्रकारों की व्याख्या करें।
