62239

III Semester M.B.A. (Day and Evening) Degree Examination, April/May - 2026

MANAGEMENT

Data Warehousing and Data Mining

(CBCS Scheme 2019 Onwards)

ACHARYA INSTITUTE OF GRADUATE STUDIES
LIBRARY
Date:
BANGALORE-107

Time : 3 Hours Maximum Marks : 70

SECTION - A

Answer any Five of the following questions. Each question carries 5 marks.(5×5=25)

1. Define a Data Warehouse. Explain its key characteristics and advantages.
2. Distinguish between OLTP and OLAP systems with suitable examples.
3. Explain the Kimball Lifecycle approach in Data Warehouse development.
4. Describe the ETL process and explain any two data transformation techniques.
5. Define Data Mining. Discuss its major functionalities.
6. Differentiate between Classification and Clustering techniques.
7. Explain the difference between Data Warehouse and Data Lake.

SECTION - B

Answer any Three of the following questions. Each question carries 10 marks.

(3×10=30)

8. Discuss the planning and project management issues involved in constructing a Data Warehouse. Explain the Data Warehouse Development Life Cycle.
9. Compare MOLAP, ROLAP, and HOLAP architectures. Design a Star Schema for a retail sales system and explain the role of Fact and Dimension tables.
10. Explain the complete ETL process in detail. Discuss methods of handling missing data, outlier detection, and data loading techniques.
11. Explain major Data Mining techniques such as Association Rules, Classification, Clustering, Neural Networks, and Decision Trees. Discuss their business applications.

[P.T.O.

SECTION-C

12. Compulsory Case Study. **(1×15=15)**

A multinational retail company, GlobalMart International, operates physical stores and an e-commerce platform across multiple countries. Over the years, the company has accumulated large volumes of transactional data, customer information, inventory records, and logistics data. However, the organisation is facing increasing difficulty in managing and analysing this data effectively.

Each regional branch uses different operational systems. Some countries use modern ERP- based OLTP systems, while others rely on legacy databases and spreadsheets. The online platform generates real-time data such as customer browsing behaviour, online purchases, payment records, and product reviews. In addition, the company maintains a customer loyalty program that collects demographic information, purchase frequency, and reward points data. Logistics partners also provide shipment and warehouse tracking information.

Due to the diversity of systems and formats, the company experiences several challenges. There are inconsistencies in product codes and customer IDs across regions. Duplicate records and missing values are common. Management reports often take a long time to generate, and executives lack a unified view of global sales performance. The marketing department finds it difficult to identify high-value customers or design targeted promotional campaigns. Similarly, the supply chain team struggles with inaccurate demand forecasting, resulting in stock shortages in some locations and excess inventory in others.

The senior management has decided to implement an enterprise Data Warehouse to centralise and integrate data from all sources. The objective is to support faster reporting, multidimensional analysis using OLAP tools, and advanced Data Mining techniques. The IT department is also considering cloud-based infrastructure to ensure scalability, flexibility, and cost efficiency. In addition to integration, the company aims to use Data Mining to identify buying patterns, improve customer retention, optimise inventory levels, and enhance decision-making. However, the organisation must carefully design the warehouse architecture, select appropriate ETL processes, and ensure data quality, security, and governance standards. GlobalMart's transformation into a data-driven enterprise depends on the successful implementation of Data Warehousing and Data Mining solutions.

Questions:

a) Suggest a suitable Data Warehouse architecture for GlobalMart and justify your recommendation.

b) Outline an appropriate ETL strategy to address data inconsistency, duplication, and missing values.

c) Recommend Data Mining techniques that can improve customer retention and demand forecasting.
