

**BCA Sem.-6 (Rep) Examination**  
**CC 308**

**Intro to Data Mining & Data Warehousing**

**Time : 2-30 Hours]**

**September-2024**

**[Max. Marks : 70**

- 
- Q. 1 (A) What are the different kinds of data that can be mined? Describe with an example. 07  
(B) Write a note on: "Mining Frequent Patterns, Associations and Correlations". 07
- OR**
- Q. 1 (A) Explain iterative sequence of Knowledge Discovery with a proper diagram. 07  
(B) Write a note on: "Major issues in Data Mining". 07
- Q. 2 (A) Write a data warehouse definition given by William H. Inmon. Explain its features. 07  
(B) Explain OLAP operations: Roll up, Slice and Dice, Drill Down and Pivot operation. 07
- OR**
- Q. 2 (A) Explain the Multitier Architecture used in Data Warehousing with proper diagram. 07  
(B) What is data cube? Explain a multidimensional data cube representation. 07
- Q. 3 (A) List out basic methods of data cleaning & explain "Handling of missing values". 07  
(B) Discuss Apriori algorithm with suitable example. 07
- OR**
- Q. 3 (A) Define Noise. Explain data smoothening technique – Binning in detail. 07  
(B) What is Apriori property? Explain its usefulness with an example. 07
- Q. 4 (A) Write a note on: Decision tree induction algorithm. 07  
(B) Write a note on: Data mining in Science and Engineering application area. 07
- OR**
- Q. 4 (A) Write a note on: K-Means partitioning method. 07  
(B) Write a note on: Data mining in Retail and Telecommunication Industries. 07
- Q. 5 Answer the following MCQ's: (Any 7) 14
- (i) \_\_\_\_\_ is a process where data relevant to the analysis task are retrieved from database.  
(A) Data cleaning (B) Data Integration  
(C) Data selection (D) Data transformation
- (ii) Data \_\_\_\_\_ is a process to remove noise and inconsistent data.  
(A) cleaning (B) selection  
(C) reduction (D) integration
- (iii) A task in a data mining pattern can be classified into two categories \_\_\_\_ and \_\_\_\_.  
(A) descriptive, predictive (B) frequent, non-frequent  
(C) characterization, discrimination (D) clustering, outlier
- (iv) Clustering can be used to generate \_\_\_\_\_ labels for the group of data.  
(A) class (B) object  
(C) instance (D) dataset
- (v) An OLAP system adopts either a star or a \_\_\_\_ schema for a subject oriented database design.  
(A) snowflake (B) fact constellation  
(C) galaxy (D) multidimensional
- (vi) Active learning and supervised learning are types of \_\_\_\_\_.  
(A) passive learning (B) machine learning  
(C) unsupervised Learning (D) semi-supervised Learning

N 495-2

Q. 5 Continue .....

- (vii) In a Data Cube each dimension may have a table associated with it, called \_\_\_\_\_.  
(A) Dimension table (B) Fact table  
(C) Association table (D) Cuboid table
- (viii) \_\_\_\_\_ kind of data warehouse provides information from a historical perspective.  
(A) Non volatile (B) Volatile  
(C) Time-variant (D) History-variant
- (ix) The 0-D cuboid, which holds the highest level of summarization, is called the \_\_\_\_\_.  
(A) Base cuboid (B) Side cuboid  
(C) Top cuboid (D) Apex cuboid
- (x) \_\_\_\_\_ factor of data quality refers to submit data on time.  
(A) interpretability (B) Timeliness  
(C) believability (D) consistency
- (xi) \_\_\_\_\_ rule of association reflects certainty to the rule.  
(A) Confidence (B) Support  
(C) Minimum-confidence (D) Minimum-support
- (xii) \_\_\_\_\_ operation removes the non-frequent subset of items.  
(A) Prune (B) Delete  
(C) Remove (D) Outcast
- (xiii) Models, called \_\_\_\_\_, predict categorical class labels.  
(A) Test sets (B) Classifiers  
(C) Tuple sets (D) Overfit sets
- (xiv) \_\_\_\_\_ as data mining application area helps in loan payment prediction.  
(A) Science and Engineering (B) Retail and Telecommunication  
(C) Intrusion Detection (D) Financial data Analysis

\*\*\*\*\* \* \*\*\*\*\*