

नोट : दोनों खण्डों से निर्देशानुसार उत्तर दीजिए। प्रश्नों के अंक उनके दाहिनी ओर अंकित हैं।

Note: Answer from Both the Section as Directed. The Figures in the right-hand margin indicate marks.

**Section - A**

1. Choose the correct Answer: (1x10)

(a) Among the following which is not a characteristic of data ware house.  
 (i) Integrated      (ii) Volatile  
 (iii) Time – Variant (iv) Subject Oriented.

(b) “Convert data and queries from one data model to another”, this statement is described about?  
 (i) Data Transformation      (ii) Wrapper  
 (iii) Data Extraction      (iv) Data cleaning

(c) In terms of data warehouse, metadata can be define as :  
 (A) Metadata is a road-map of data warehouse.  
 (B) Metadata in Datawarehouse defines the warehouse objects.  
 (C) Metadata acts as a directory  
 (i) A & B is correct      (ii) A & C is correct  
 (iii) B & C is correct      (iv) All are correct.

(d) Which one of the following is not an action of data integration?  
 (i) Resolve inconsistencies  
 (ii) summarized Data  
 (iii) Eliminate duplications  
 (iv) Compare Query Results.

(e) Which of the following is not correct regarding data cleansing.  
 (A) find & Remove duplicate tuples  
 (B) Detect inconsistent, wrong data  
 (C) Patch missing, unreadable data  
 (D) Convert data to uniform format  
 (i) A & B      (ii) A & C  
 (iii) Only C      (iv) Only D

(f) In the OLAP model, the ..... provides the multidimensional view.  
 (i) Data layer      (ii) Application layer  
 (iii) Data link layer (iv) Presentation layer

(g) How to describe the data contained in the data warehouse.  
 (i) Relational Data      (ii) Operational Data  
 (iii) Meta data      (iv) Informational Data

(h) What is a Data Mart?  
 (i) It is the process of extracting Data.  
 (ii) It is the collection of Data in the data warehouse.  
 (iii) It is the access layer of the data warehouse environment that is used to get data out to the uses.  
 (iv) It is the process of extracting patterns from large data sets.

(i) Bayesian classifiers is:  
 (i) A class of learning algorithm that tries to find an optimum classification of a

set of examples using the probabilistic theory.

(ii) Any mechanism employed by a learning system to contain the search space of a hypothesis.

(iii) Both A & B

(iv) None of them

(j) Background Knowledge referred to

(i) Additional acquaintance used by a learning algorithm to facilitate the learning process.

(ii) A Neural network that makes use of a hidden layer.

(iii) It is a form of automatic learning

(iv) None of these.

2. Answer the following short – Answer type questions: 2X5

(a) Define cluster Analysis?

(b) Explain Association rule in Mathematical notations?

(c) Define the concept of classification?

(d) Define OLTP & OLAP?

(e) What are the steps involved in KDD process.

**Section - B**

Answer the following long Answer type questions: 12X5

**Unit-I**

3. What is Data Mining? Explain its functionalities and classification.

**OR**

Explain various steps in Data pre-processing?

**Unit-II**

4. Explain about various OLAP Operations? Differentiated between multidimensional OLAP & Multi operational OLAP.

**OR**

Explain 3-tier Datawarehouse Architecture? Differentiate operational database system and data warehouse.

**Unit-III**

5. Discuss Bayesian classification & predication.

**OR**

Explain different Association rules in detail?

**Unit-IV**

6. Define clustering and describe the categorization of major clustering method.

**OR**

Explain Apriori Algorithm with an example?

**Unit-V**

7. Explain about different methods used for mining text databases? Discuss about time-series analysis?

**OR**

Discuss about web mining? What is sequential pattern mining?