

SI-138**September-2020****BCA., Sem.-VI****CC-308 : Data Mining and Data Warehousing
(New Course)****Time : 2 Hours]****[Max. Marks : 50**

- Instructions :**
- (1) All questions in Section-I carry equal marks.
 - (2) Attempt any **two** questions from Section-I.
 - (3) Question No-5 in Section-II is compulsory.

SECTION - I

1. (A) Discuss Data Mining. Also discuss major issues in data mining. **10**
(B) Discuss the technologies used for data mining. **10**
2. (A) Explain the concept of Data Warehouse. Also discuss its key features. **10**
(B) Discuss : The typical OLAP operations in detail. **10**
3. (A) What is meant by Data cleaning ? Discuss the methods of data cleaning. **10**
(B) What is meant by data preprocessing ? What are the major tasks in data preprocessing. **10**
4. (A) What is classification ? Discuss the two step process of data classification in detail. **10**
(B) What is cluster analysis ? Discuss the basic clustering methods. **10**

SECTION - II

10

Answer any 5 :

(1) _____ is a summarization of general characteristics or features of a target class of data.

- (a) Data warehousing
- (b) Data Discrimination
- (c) Data characterization
- (d) Data Patterns

(2) Data mining is

- (a) The process of discovering interesting patterns from large data sets.
- (b) The stage of selecting right data for a KDD process.
- (c) A subject-oriented integrated time variant non-volatile collection of data in support of management.
- (d) None of these

(3) _____ is the science of searching for documents or information in documents.

- (a) Regression Analysis
- (b) Cluster Analysis
- (c) Classification
- (d) Information Retrieval

(4) _____ is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management decisions.

- (a) Data Mining
- (b) Data Warehousing
- (c) Web Mining
- (d) Text Mining

(5) An _____ system is customer oriented and is used for transaction and query processing by clerks and clients.

(a) OLAP

(b) OLTP

(c) Both (a) and (b)

(d) None of the above

(6) Data warehouse can be built using

(a) Top-down approach

(b) Bottom-up approach

(c) Combination of top-down and bottom-up approaches

(d) All of the above

(7) The _____ view includes fact tables and dimensions tables for designing data warehouse.

(a) Top-down view

(b) Data source view

(c) Data warehouse view

(d) Business query view

(8) Data quality includes

(a) Accuracy and completeness

(b) Consistency and timeliness

(c) Believability and interpretability

(d) All of the above

(9) Which of the following is a data smoothing technique ?

- (a) Binning
- (b) Regression
- (c) Outlier analysis
- (d) All of above

(10) _____ quantize the object space into a finite number of cells that form a grid structure.

- (a) Partitioning method
- (b) Hierarchical method
- (c) Grid-based method
- (d) Density based method