

Reg. No. :

--	--	--	--	--	--	--	--	--	--

**Question Paper Code: 41881**

B.E. / B.Tech. DEGREE EXAMINATION, MAY 2017

Elective

Information Technology

14UIT921 - DATA MINING CONCEPTS AND TECHNIQUES

(Regulation 2014)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

- \_\_\_\_\_ Is a subject oriented Integrated, Time Variant, and Non-Volatile collection of data in support of management decision.
  - Data mining
  - Data Warehousing
  - Web Mining
  - Text mining
- The Data Warehouse is
  - Read Only
  - Write Only
  - Read and Write only
  - None of these
- \_\_\_\_\_ describes the data contained in the data warehouse.
  - Relational Data
  - Operational Data
  - Meta Data
  - Informational Data
- Data can be updated in \_\_\_\_\_ Environment.
  - Data mining
  - Data Warehousing
  - Operational
  - Informational
- Data scrubbing is which of the following?
  - A process to reject data from the data warehouse and to create the necessary indexes
  - A process to load the data in the data warehouse and to create the necessary indexes

- (c) A process to upgrade the quality of data after it is moved into a data warehouse
  - (d) A process to upgrade the quality of data before it is moved into a data warehouse
6. A goal of data mining includes which of the following?
- (a) To explain some observed event or condition
  - (b) To confirm that data exists
  - (c) To analyze data for expected relationships
  - (d) To create a new data warehouse
7. A data warehouse is which of the following?
- (a) Can be updated by end users
  - (b) Contains numerous naming conventions and formats
  - (c) Organized around important subject areas
  - (d) Contains only current data
8. Which of the following is a true statement about classification method in data mining?
- (a) The objective is to provide an overall description of data, either in itself or in each class or concept
  - (b) The main approaches in obtaining data description are data characterization and data Discrimination
  - (c) The purpose is to search for the most significant relationship across a large number of variables or attributes
  - (d) It is the process of finding models or functions that map records into one of several discrete prescribed classes
9. Which of the following is not a data mining method?
- (a) Classification and prediction
  - (b) A priori algorithms
  - (c) Dependency analysis
  - (d) Data description
10. Management turnover, loss of market or license or supplier, shortages and labour problems are indications of what kind of problems?
- (a) Fraud
  - (b) Control environment
  - (c) Going concern
  - (d) Cash flow

PART - B (5 x 2 = 10 Marks)

11. Define data warehouse.
12. Write the steps of KDD process.
13. Describe the different classification of association rule mining.
14. Define smoothing.
15. Define spatial data mining.

PART - C (5 x 16 = 80 Marks)

16. (a) Diagrammatically illustrate and discuss the data warehousing architecture with briefly explain components of data warehouse. (16)

Or

- (b) Explain OLAP and discuss about the following:

- (i) Categories of OLAP tools (10)
- (ii) OLAP Operation (6)

17. (a) (i) Discuss the architecture of data mining system with a neat sketch. (8)
- (ii) Explain data mining task primitives with example. (8)

Or

- (b) Explain preprocessing method in data mining. (16)

18. (a) Explain bayesian classification in detail. (16)

Or

- (b) Write short notes on support vector machine and Lazy learners. (16)

19. (a) (i) Explain the different types of data used in cluster Analysis. (10)
- (ii) Discuss the use of outlier analysis. (6)

Or

- (b) Explain how data mining is used for retail Industry. (16)

20. (a) Discuss about spatial mining and its algorithm with real time example. (16)

Or

(b) Explain the following mining in detail with real time example: (i) Web Mining  
(ii) Temporal mining. (16)

---