Reg. No. : | | | | | | | | | | |

Question Paper Code: 55802

B.E. / B.Tech. DEGREE EXAMINATION, MAY 2018

Fifth Semester

Information Technology

15UIT502 - DATA WAREHOUSING AND DATA MINING

(Regulation 2015)

Duration: Three hours Maximum: 100 Marks

Answer ALL Questions

PART A - (5 x 1 = 5 Marks)

1. Data modeling technique used for data mart is

   (a) Dimensional modeling (b) ER – model

   (c) Extended ER – model (d) Physical model

2. Which of the following process includes data cleaning, data integration, data selection, data transformation, data mining, pattern evolution and knowledge presentation?

   (a) KDD process (b) ETL process (c) KTL process (d) MDX process

3. A Business Intelligence system requires data from:

   (a) Data warehouse (b) Operational systems
   (c) All possible sources within the organization (d) Web servers

4. Which of the following is not an ETL tool?

   (a) Informatica (b) Oracle warehouse builder
   (c) Datastage (d) Visual studio

5. The generalization of multidimensional attributes of a complex object class can be performed by examining each attribute, generalizing each attribute to simple-value data and constructing a multidimensional data cube is called as

   (a) Object cube (b) Relational cube
   (c) Transactional cube (d) Tuple

PART - B (5 x 3 = 15 Marks)

6. Give nine decisions in the design of data warehouse.

7. Differentiate data mining and data warehousing.

8. Define outliers. List various outlier detection approaches.

9. What is meant by hierarchical clustering?

10. State the purpose of web content mining.

PART - C (5 x 16 = 80 Marks)

11. (a) Explain the design and construction of a data warehouse. (16)

Or

(b) List the components of data warehouse architecture and show how they are interconnected? (16)

12. (a) Explain DBMiner tool in data mining. (16)

Or

(b) Explain how data mining is used in banking industry. (16)

13. (a) Explain the constraint-based association mining in details. (16)

Or

(b) Explain the issues regarding classification and prediction methods. (16)

14. (a) Discuss the requirements of clustering in data mining. (16)

Or

(b) Explain the various methods for detecting outliers. (16)

15. (a) What are the salient features are of times series data mining? Explain in details. (16)

Or

(b) Analysis mining of spatial data bases and mining of text data bases. (16)