

## **B3.E4-R5 : DATA WAREHOUSING AND DATA MINING**

**DURATION : 03 Hours**

**MAXIMUM MARKS : 100**

**Roll No. :**

**Answer Sheet No. :**

**Name of Candidate :** \_\_\_\_\_ ; **Signature of Candidate :** \_\_\_\_\_

### **INSTRUCTIONS FOR CANDIDATES :**

- Carefully read the instructions given on Question Paper, Answer Sheet.
- Question Paper is in English language. Candidate has to answer in English Language only.
- Question paper contains Seven questions. The Question No. 1 is compulsory. Attempt any FOUR Questions from Question No. 2 to 7.
- Parts of the same question should be answered together and in the same sequence.
- Questions are to be answered in the ANSWER SHEET only, supplied with the Question Paper.
- Candidate cannot leave the examination hall/ room without signing on the attendance sheet and handing over his/her Answer Sheet to the Invigilator. Failing in doing so, will amount to disqualification of Candidate in this Module/Paper.
- After receiving the instruction to open the booklet and before answering the questions, the candidate should ensure that the Question Booklet is complete in all respects.

**DO NOT OPEN THE QUESTION BOOKLET UNTIL YOU ARE TOLD TO DO SO.**

- What is Data Mining ? How does it relate to Data Warehousing ?
  - What are the major issues in Data Warehousing ?
  - What do you mean by Partitioned clustering ?
  - What is Web usage mining ?
  - Discuss in short about Data warehouse architecture.
  - What is outlier analysis ? Why is outlier mining important ?
  - Calculate the two regression equations of X on Y from the data given below, taking deviations from the actual means of X and Y.

<b>Price (Rs.)</b>	10	12	13	12	16	15
<b>Amount demanded</b>	40	38	43	45	37	43

(7x4)

- Elaborate the term Data Warehouse and why it is important ?
  - Discuss data model and its types in detail.
  - What is the purpose of data mart ? Explain.

(8+6+4)

- What is ETL in data warehouse ? Explain the types of ETL tools in detail.
  - Differentiate between ETL and ELT processes.
  - What is Change Data Capture Technique ? Explain its benefits.

(10+4+4)

- Why data cleaning is necessary ? Discuss with the help of example.
  - What is Data Mining Query Language ? Explain in detail.
  - Explain KDD process in detail.

(6+6+6)

- Compare MOLAP Model and ROLAP Model.
  - Differentiate between OnLine Transactional Processing (OLTP) and OnLine Analytical Processing (OLAP).
  - What is the relationship between Data Warehousing and Data Mining ?

(8+4+6)

6. (a) Explain decision tree classification with the help of suitable example.  
(b) Explain k nearest neighbor classifier with an illustration.  
(c) What is Nonlinear Regression ? Give its expression.

**(6+8+4)**

7. (a) Explain the Apriori Algorithm and its role in Association Rule Mining.  
(b) What is Hierarchical Clustering ? Explain in detail.  
(c) Elaborate the steps involved in data preprocessing.

**(6+8+4)**

**- o O o -**

**SPACE FOR ROUGH WORK**