9. (a) What is clustering ? Explain how nearest neighbour clustering is different from K-nearest neighbour clustering technique. 8

(b) Explain the concept of Neural Networks. How is it implemented in data mining ? 6

10. (a) Differentiate between :

OLAP and Datamining. 5

(b) What is KDD ? Explain Knowledge discovery process with respect to the relationships or pattern it discovers with help of several stages. 8

VKR/KS/13/3357

Faculty of Engineering & Technology

Eighth Semester B.E. (Computer Science) (New) Examination

DATA WAREHOUSING AND MINING

Elective—IV

Sections—A&B

Time—Three Hours] [Maximum Marks—80

INSTRUCTIONS TO CANDIDATES

(1) All questions carry marks as indicated.

(2) Answer **THREE** questions from Section A and **THREE** questions from Section B.

(3) Assume suitable data wherever necessary.

(4) Illustrate your answers wherever necessary with the help of neat sketches/examples.

SECTION—A

1. (a) Why are operational systems not suitable for providing strategic information ? Give three specific reasons and explain. 4

(b) Name the characteristic feature of the computing environment needed to provide strategic information. 6

(c) A data warehouse is an environment, not a product. Discuss. 3

2. (a) Differentiate between :
   (i) Operational and Informational System
   (ii) Data warehouse and Data mart. 8

   (b) What is Metadata in DW ? Give its various types and significance while operation in DW. 5

3. (a) Why do you need a separate data staging component ? Explain the functions carried out by each of following components :
   (i) Data Extraction
   (ii) Data Transformation
   (iii) Data Loading. 7

   (b) Explain any six different methods for information delivery. 6

4. (a) Explain the methods of Gathering Information with respect to Interview and JAD session. 7

   (b) What is STAR Schema ? How it is useful for data design in relational model consisting of fact and dimension table. Also give its advantages. 6

5. (a) Describe slowly changing dimensions. What are the three types ? Explain in brief. 6

   (b) Write in short (any **TWO**) :
   (i) Snowflake Schema
   (ii) Aggregate fact table
   (iii) Family of STARS. 8

SECTION—B

6. (a) Describe the need for multidimensional analyses. 3

   (b) Differentiate between :
   (i) OLTP and DW
   (ii) MOLAP and ROLAP. 10

7. (a) State any five rule proposed by Dr. Coddo for an OLAP system. 5

   (b) Explain the following terms with neat diagram and example :
   (i) Drill down and Roll-up
   (ii) Slice and dicing or Rotation
   (iii) Hyper Cube. 9

8. (a) Explain any two reasons in details, why feeding data into the OLAP system directly from operational system is not recommended. 4

   (b) What is meant by Conforming Dimension ? Why this is important in a Data Warehouse ? 3

   (c) Write in short about :
   (i) Snapshot and transaction fact table 2
   (ii) Junk dimension 2
   (iii) Rapidly changing dimension. 2