Roll No…… Total Pages: 3

10013

MCA/M-17

DATA WAREHOUSING AND MINING

Paper: MCA-14-43

Time: Three Hours Maximum Marks: 80

Note: Attempt five questions including No. 1 which is compulsory. All questions carry equal marks.

Compulsory Question

1. (a) What do you mean by rollup and dice OLAP operation? Give examples.
   (b) Write a note on lattices of cuboid.
   (c) What do you understand by discrimination?
   (d) Briefly explain data visualization.
   (e) Write a note on similarity model.
   (f) "Birds of a feather flock together." Comment.
   (g) What do you mean by support and confidence?
   (h) Briefly explain Information gain.

Unit-I

2. What are the different steps for designing and construction of data warehouse? How a data warehouse system will be implemented?

3. (a) What do you mean by data warehouse schema? Discuss different types of data warehouse schema.

   (b) What are the different methods for the computation of data cube? Explain.

Unit-II

4. (a) Define data mining? Write and explain the different phases of knowledge discovery process.

   (b) What do you mean by outlier? Discuss different type of outliers. What are the challenges of outlier detection?

4. What is the role of data preprocessing in KDD process? How data integration and transformation is performed? Discuss.

Unit-III

6. (a) What are the steps for designing similarity models? Discuss.

   (b) What do you mean by memory based reasoning? Explain the challenges of memory reasoning.

7. What do you mean by partition based clustering? Calculate k-means clustering for the following dataset using Euclidean distance function. When K=2 and initial data clusters are T3 and T4.

Data Point	X	Y
T1	2	5
T2	5	7
T3	6	6
T4	4	2
T5	5	4
T6	5	3
T7	1	1
T8	3	8

Unit-IV

8. (a) Draw a neat sketch of neural network. How classification by back propagation Is performed?

   (b) How association rules are generated by using candidate key and without using Candidate key ? Discuss.

9. What do you mean by decision tree? How will you finalize the root for an appropriate decision tree? Draw a decision tree for the following dataset Assume that information gain for attribute : Course = 0.3421, Game=0.1213 and Dress=0.4312.

Training		Attributes		Class
Pattern	**Course**	**Game**	**Dress**	
T1	M.Tech.	Cricket	Casuals	Indian
T2	M.Tech.	Football	Formals	Foreign
T3	M.Tech	Cricket	Formals	Indian
T4	B.Tech	Hockey	Formals	Foreign
T5	M.Tech.	Hockey	Casuals	Indian
T6	B.Tech.	Cricket	Formals	Indian
T7.	M.Tech.	Hockey	Formals	Foreign
T8	B.Tech.	Hockey	Casuals	Indian

Generate classification rules from the decision tree for the following recall Pattern.

Recall Pattern	Attributes			Class
	Course	**Game**	**Dress**	
R1	B.Tech.	Cricket	Casuals	?
R2	B.Tech	Football	Formals	?
R3	M.Tech	Football	Formals	?
R4	B.Tech.	Football	Casuals	?