Code No. 6298 / O

FACULTY OF ENGINEERING

B.E. 4/4 (CSE) II – Semester (Old) Examination, April / May 2014

Subject: Data Mining

Time: 3 Hours **Max.Marks: 75**

Note: Answer all questions from Part A. Answer any five questions from Part B.

PART – A (25 Marks)

1	Define Data Warehouse.	2
2	Differentiate between classification and clustering.	2
3	What is binning? Smooth the following data using by bin means and by bin boundaries 4, 8, 15, 21, 24, 25, 28, 34.	3
4	What is concept hierarchy? Give few examples.	3
5	Define iceberg query.	2
6	Define multilayer feed forward neural network.	3
7	Define support and confidence.	2
8	Define Bayes theorem.	3
9	What is the role of meta data repository in a data warehouse?	2
10	How do you handle missing values?	3

PART – B (50 Marks)

11	(a) Explain data mining as a step in the process of knowledge discovery.	5
	(b) Differentiate DLAP and OLTP.	5
12	(a) Explain about analysis of attribute relevance.	5
	(b) How is the analytical characterization performed? Explain with an example.	5
13	Write and explain Apriori algorithm to find all frequent item sets and strong association rules for the following database, where min_sup = 60% and min_conf = 80%	10

Tid	Items
T100	{K, A, D, B}
T200	{D, A, C, E, B}
T300	{C, A, B, E}
T400	{B, A, D}

14	Explain classification problem using decision trees.	10
15	Explain how classifier accuracy can be estimated. Discuss the general techniques for improving classifier accuracy.	10
16	(a) Discuss distance based outlier detection.	5
	(b) Explain OPTICS algorithm for clustering.	5
17	Explain the following:	
	a) The knowledge to be mined	5
	b) The construction of FP-tree with example.	5

****

Code No. 6561 / N

FACULTY OF ENGINEERING

B.E. 4/4 (CSE) II – Semester (New) (Main) Examination, April / May 2014

Subject: Data Mining

Time: 3 Hours **Max.Marks: 75**

Note: Answer all questions from Part A. Answer any five questions from Part B.

PART – A (25 Marks)

1	What is data mining?	2
2	Define pre-processing. Why do we need pre-processing?	3
3	What is Apriori property?	3
4	Define clustering with an example.	2
5	Why do we require pruning in decision tree?	3
6	What is linear-regression?	2
7	Explain Bayes theorem.	3
8	Define: a) Agglomerative hierarchical clustering b) Divisive hierarchical clustering	3
9	How K - means algorithm differ from K - mediods?	2
10	What is meant by data reduction?	2

PART – B (50 Marks)

11	(a) Explain various data mining functionalities.	5
	(b) Discuss various issues in data mining.	5
12	(a) Explain the architecture of data warehouse.	5
	(b) Differentiate OLAP and OLTP.	5
13	Explain Apriori algorithm with a suitable example.	10
14	Explain the naïve Bayes classification and give an example.	10
15	Describe the working of the DBSCAN algorithm, and also explain the concept of a cluster as used in DBSCAN.	10
16	Explain various grid based methods for clustering.	10
17	Write short notes on any two of the following: a) Text mining b) Web mining c) Mining multimedia databases.	10

****