USN | | | | | | | | | | |

07MCA542

Fifth Semester MCA Degree Examination, December 2010
Data Mining

Time: 3 hrs. Max. Marks:100

Note: ***Answer any FIVE full questions.***

1 a. What is data mining? With a neat sketch, explain the process of knowledge discovery in databases. **(10 Marks)**

b. What is an attribute? Explain different types of attributes, with examples. **(10 Marks)**

2 a. What is sampling? List and explain different types of sampling techniques, with examples. **(10 Marks)**

b. Find the distance matrix using city block and Euclidean distance for the following data points. **(10 Marks)**

	x	y
P_1	4	4
P_2	8	4
P_3	15	8
P_4	24	4
P_5	24	12

3 a. List the different characteristics of decision tree induction. **(10 Marks)**

b. Explain rule – based classification technique and list the characteristics. **(10 Marks)**

4 a. Discuss the apriory algorithm for frequent itemset generation. **(10 Marks)**

b. Explain briefly the alternative methods for generating frequent itemsets. **(10 Marks)**

5 a. Explain the different steps involved in subsequent extension of FP – tree. **(10 Marks)**

b. Explain how association patterns are evaluated. **(10 Marks)**

6 a. What is clustering? What are the types of clusters? Explain the basic K – means cluster algorithm. **(10 Marks)**

b. Perform hierarchical clustering of five samples, using single linkage or minimum method. Draw the dendogram. **(10 Marks)**

	x	y
1	4	4
2	8	4
3	15	8
4	24	4
5	24	12

7 a. Explain the phases involved in multidimensional analysis of complex data objects. **(10 Marks)**

b. Write a note on mining multimedia database. **(10 Marks)**

8 Write note on any two of the following :

a. Data mining applications.

b. Additional themes in data mining.

c. Trends in data mining. **(20 Marks)**

*****