P 5510

BE/INSEM/OCT-86

BE. (Computer Engineering)

DATA MINING TECHNIQUES AND APPLICATIONS

(Elective - I) (2012 Pattern) (Semester - I)

Time : 1 Hour]   *[Max. Marks :30*

Instructions to the candidates:

1) *Answer Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6*

2) *Neat diagrams must be drawn wherever necessary.*

3) *Figures to the right side indicate full marks.*

4) *Assume Suitable data if necessary.*

Q1) a) Consider the following price data in rupees :: 6,8,16,22,22,25,26,29,35. Use the following binning methods for data smoothing : **[6]**

   i) Bin boundaries

   ii) Bin means

   iii) Bin medians

   b) Compare OLTP with OLAP. **[4]**

<div align="center">OR</div>

Q2) a) What are the major tasks in data preprocessing? Explain them in brief.**[6]**

   b) What are missing values? Explain methods to handle missing values.**[4]**

Q3) a) A database has nine transactions. Let min. sup. = 22% and min. conf. = 70% **[8]**

TID	List of Items
T100	A1, A2, A5
T200	A2, A4
T300	A2, A3
T400	A1, A2, A4
T500	A1, A3
T600	A2, A3
T700	A1, A3
T800	A1, A2, A3, A5
T900	A1, A2, A3

P.T.O.

Find all frequent itemsets using Apriori algorithm.

   b) What are applications of Association Rule mining? **[2]**

<div align="center">OR</div>

Q4) a) Write and explain FP- Growth algorithm. Enlist advantages of FP - Growth algorithm? **[6]**

   b) What are frequent itemsets and closed itemsets? **[4]**

Q5) a) Write and Explain k-NN Algorithm for classification. **[6]**

   b) State various performance metrics that are used to evaluate various classifiers. **[4]**

<div align="center">OR</div>

Q6) a) What is need of Decision pruning? Explain any two methods in short.**[6]**

   b) Explain Bayes Theorem in short. **[4]**

<div align="center">ⓘ ⓘ ⓘ</div>