

Total No. of Questions : 6]

SEAT No. :

P5185

[Total No. of Pages : 2

B.E./Insem.-587

B.E. (Computer Engineering) (Semester-I)
DATA MINING TECHNIQUES AND APPLICATIONS
(2012 Pattern) (Elective -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 indicate full marks.*
- 4) *Assume suitable data if necessary.*

- Q1)** a) Explain the knowledge discovery in databases (KDD) with diagram. What is the role of data mining steps in KDD? [6]
- b) Write a short note on OLAP. [4]

OR

- Q2)** 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) Explain forward selection and backward elimination method of attribute subset selection with example. [4]

- Q3)** a) Explain the Apriori algorithm for generation of association rule. How candidate keys are generated using Apriori algorithm? [6]
- b) What are frequent itemsets and closed itemsets? [4]

P.T.O.

OR

- Q4)** a) A database has five transactions. Let min. support = 40% and min. conf. = 80%. [6]

TID	Items
T100	{I1, I3, I4}
T200	{I2, I3, I5}
T300	{I1, I2, I3, I5}
T400	{I2, I5}
T500	{I1, I5}

Calculate all frequent item sets using Apriori algorithm.

- b) Explain multilevel association rules. What is uniform support? [4]

- Q5)** a) Explain decision tree classification algorithm ID3 with entropy formula. [6]

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

OR

- Q6)** a) Write and explain K-Nearest-Neighbor Classification Algorithm. [6]

- b) Differentiate between Supervised learning & Unsupervised learning. [4]

