

Total No. of Questions : 6]

SEAT No. :

P5822

[Total No. of Pages : 2

BE/Insem./Oct.-589

B.E. (Computer Engineering)

DATA MINING AND WAREHOUSING

(2015 Pattern) (Elective - I) (Semester - I)

Time : 1 Hour]

[Max. Marks : 30

Instructions to the candidates:

- 1) *Solve Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6.*
- 2) *Assume suitable data, if necessary.*
- 3) *Neat diagrams must be drawn wherever necessary.*
- 4) *Figures to the right indicate full marks.*

Q1) a) Suppose that the minimum and maximum values for the attribute income are \$12,000 and \$98,000 respectively. Normalize income value \$73,600 to the range [0.0, 1.0] using min-max normalization method. [4]

b) Explain various data cleaning techniques. [4]

c) What is correlation analysis? [2]

OR

Q2) a) Explain different methods for attribute subset selection (any 2). [4]

b) For the given attribute marks values : [4]

35, 45, 50, 55, 60, 65, 75

Compute mean, median, mode.

Also compute Five number summary of above data.

c) Enlist different methods of sampling. [2]

Q3) a) From the architectural point of view, explain different data warehouse models. [4]

b) Differentiate between ROLAP, MOLAP and HOLAP. [4]

c) What is Concept Hierarchy? Explain. [2]

OR

Q4) a) Draw and Explain a data warehouse architecture. [4]

b) Explain following OLAP operations with example. [4]

i) Drill Up

ii) Slice & Dice

c) What is fact table and dimension table. [2]

P.T.O.

Q5) a) Calculate Euclidean and Manhattan distance between following two objects. [4]

$$A = \{2, 4, 8, 6, \}, \quad B = \{3, 4, 6, 7\}$$

b) How to compute dissimilarity between categorical variables. Explain with suitable example. [4]

c) What is cosine similarity? [2]

OR

Q6) a) Compute cosine similarity among following documents using term frequency vector [4]

d_1 : "The sun in the sky is bright"

d^2 : "We can see the shining sun, the bright sun"

b) How to compute dissimilarity between ordinal variables. Explain with suitable example. [4]

c) Explain Data matrix and Dissimilarity matrix. [2]

