

GANPAT UNIVERSITY
M. TECH SEMESTER - II – Computer Engineering
3CE203: Data Mining & Data Warehousing
Regular Examination | April - June 2015

Time: 3Hour]

[Total Marks: 60

- Q-1**
- A** Explain roll-up, drill down, slice and dice operations in the multidimensional data model with the help of suitable examples. 5
- B** What is lattice of Cuboid? Discuss with an example 5
- OR
- A** Explain Fact Constellation schema with an example. 5
- B** Explain density based methods. Define terms core object, directly density reachable, density reachable and density connected with help of examples. 5

Q-2

	A1	A2	A3	A4	A5	A6	
A	O1	True	True	True	False	False	True
	O2	False	True	True	False	True	False

5

For the above given objects having asymmetric attributes, where True is more significant than False; find a) Dissimilarity measure and b) Jaccard coefficient.

- B** For the given distance matrix apply agglomerative hierarchical clustering using: 5
- a) Single-link b) complete-link
- c) Plot the dendro-grams for the solutions to part a) and b)

	A	B	C	D
A	0	4	5	1
B		0	6	2
C			0	3
D				0

OR

- Q-2**
- A** Use the k-medoid algorithm to cluster the following 8 objects into three clusters. 5
 $X_1=(2,5)$, $X_2=(2,10)$, $X_3=(8,4)$, $X_4=(5,8)$, $X_5=(7,5)$, $X_6=(6,4)$, $X_7=(4,9)$, $X_8=(1,2)$.
 Take initial clusters as X_2 , X_4 and X_8 and distance measure as Manhattan distance.
 1) Find final three clusters and their medoids formed after 2 iterations.
- B** Differentiate between OLTP and OLAP. 5
- Q-3**
- A** Use the k-means algorithm to cluster the following 8 objects into three clusters. 5
 $X_1=(2,5)$, $X_2=(2,10)$, $X_3=(8,4)$, $X_4=(5,8)$, $X_5=(7,5)$, $X_6=(6,4)$, $X_7=(4,9)$, $X_8=(1,2)$.
 Take initial clusters as X_2 , X_4 and X_8 and distance measure as Euclidean distance.
 1) Find final three clusters and their centroids formed after 3 iterations.
- B** What is sequential Pattern Mining? How it differs from Association Rule Mining? 5
 Discuss with an example.

Section II

- Q-4 A Given minimum support as 60% and minimum confidence as 80%. Find out the frequent itemsets and strong association rules for the example given below Using Vertical Data Format (ECLAT). 5

TID	ITEMS
100	{A,B,C,D,E,F}
200	{G, B,C,D,E,F}
300	{A,H,D,E}
400	{A,I,J,D,F}
500	{J,B,B,D,K,E}

- B Explain the various sampling techniques used in data mining. 5

OR

- Q-4 A Explain ROCK algorithm for clustering with the help of an example. 5

- B Explain Inter transaction Association Rule Mining using example. 5

- Q-5 A Explain Interval-Scaled, Categorical, Ordinal, and Ratio-Scaled Variables used in data clustering analysis. 5

- B Explain KDD process with an example. 5

OR

- Q-5 A Draw a diagram of influence of data mining with multiple discipline and discuss it. 5

- B Given a data set X with 3-dimensional categorical samples: 5

Construct a decision tree using the computation steps given in the C4.5 algorithm

Attribute1	Attribute2	Class
T	1	C2
T	2	C1
F	1	C2
F	2	C2

- Q-6 A Suppose that the data for analysis include the attribute age. The age values for the data tuples are (in increasing order): 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. 5

a) Use min-max normalization to transform the value 35 for age onto the range [0.0, 1.0]

b) Use z-score normalization to transform the value 35 for age, where the standard

c) Use normalization by decimal scaling to transform the value 35 for age.

deviation of age is 12.94 years

- B Predict a class label of an unknown tuple X= {age = '<=20', Income= 'Medium', Student='Yes', Credit_Rating='Fair'} using Naïve Bayesian Classification. 5

Age	Income	Student	Credit_rating	Class: Buys Laptop
>30	Medium	No	Excellent	No
<=20	High	No	Fair	No
21..30	High	Yes	Fair	Yes
<=20	High	No	Excellent	No
21..30	Medium	No	Excellent	Yes
21..30	High	No	Fair	Yes
<=20	Medium	Yes	Excellent	Yes
>30	Medium	No	Fair	Yes
>30	Medium	Yes	Fair	Yes
>30	Low	Yes	Fair	Yes
<=20	Low	Yes	Fair	Yes
>30	Low	Yes	Excellent	No
21..30	Low	Yes	Excellent	Yes
<=20	Medium	No	Fair	No

====End====