

Total No. of Questions : 5]

SEAT No. :

PC-2959

[Total No. of Pages : 2

[6380]-32

M.B.A.

SC-BA-02: DATA MINING
(2019 Pattern) (Semester - II) (206 BA)

Time : 2½ Hours]

[Max. Marks : 50

Instructions to the candidates:

- 1) *Attempt All questions.*
- 2) *Figures to right indicates full marks.*
- 3) *State your assumptions clearly.*

Q1) Solve any Five questions :

[10]

- a) Define the term Big Data.
- b) Explain classification.
- c) What are the different types of attributes?
- d) Explain data cleaning concept in data preprocessing.
- e) Define partitional clustering method.
- f) Explain Frequent Pattern Mining (FPM).
- g) What is hyper plane.
- h) Describe the descriptive modeling.

Q2) Solve any Two :

[10]

- a) Explain the characteristics of data mining in Business environment.
- b) Explain the data preprocessing techniques in detail.
- c) Classification is supervised learning. Justify.

Q3) a) What is feature selection/Instance selection? Explain the importance of feature selection with reference to Big Data Scenario. [10]

OR

- b) Prathamesh distributors is in FMCG retail business from many years. The management wants to analyse the transaction data of customers to classify its outlet products as per need, preferences of customer. As a data analyst consultant suggest a suitable algorithm or model to management of prathamesh distributors

P.T.O.

- Q4) a)** What is hierarchical clustering? Discuss the two types of Algorithms and write different steps to draw the dendrogram? **[10]**

OR

- b)** Draw a Dendrogram for the following Data Set by using Agglomerative clustering.

	P1	P2	P3	P4	P5
P1	0				
P2	9	0			
P3	3	7	0		
P4	6	5	9	0	
P5	11	10	2	8	0

- Q5) a)** Discuss the Decision-Tree Based Approach. Explain the different concepts and strategies to draw decision tree by using ID3 Algorithm. **[10]**

OR

- b)** State and Explain apriori algorithm and apply it for the following data set to find the list of frequent item sets, if the minimum support = 60%, confidence = 80%.

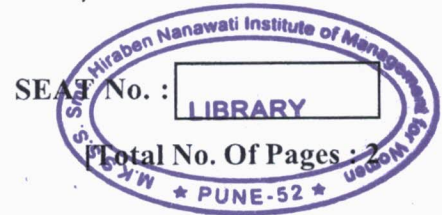
Transaction ID	List of Items
1	a, b, c
2	b, c, d
3	c, d
4	b, d
5	a, c
6	a, c, d



25 feb 23

Total No. Of Questions : 5]

PA-4176



[5946]-212

M.B.A.

(SC-BA-02): Data Mining (Business Analytics)
(2019 Pattern) (206BA) (Semester-II)

Time : 2½Hours]

[Max. Marks : 50

Instructions to the candidates :

- 1) Attempt all questions.
- 2) Figures to right indicate marks.
- 3) State your assumptions clearly.

Q1) Solve any Five questions :

[5 × 2 = 10]

- a) Define the term Data Mining.
- b) Define clustering with example.
- c) Explain Data Normalization.
- d) Explain the concept of predictive modeling.
- e) What is outlier in mining algorithm?
- f) What is association rule?
- g) Write the importance of feature selection.
- h) Explain the term customer profiling.

Q2) Solve any Two questions :

[2 × 5 = 10]

- a) What is Big data? Write it's characteristics.
- b) Explain the data preprocessing process with suitable example.
- c) Elaborate market segmentation in product distribution with suitable example.

P.T.O.

Q3) a) Discuss Decision-Tree Based approach with suitable example. **[10]**

OR

b) Explain any two applications of data mining.

Q4) a) Discuss clustering w.r.t. partition and Hierarchical clustering methods. **[10]**

OR

b) Write detail note on Density-based clustering in data mining with example.

Q5) a) Discuss Apriori Algorithm. **[10]**

OR

b) Write short notes (any Two) : **[2 × 5 = 10]**

- i) B₂B customer buying path analysis
- ii) Data cleaning
- iii) Big data analytics in business environment.



19 Aug 22 (35)

Total No. of Questions : 5]

SEAT No. :

LIBRARY

[Total No. of Pages : 2

P7265

[5860]-212

F.Y.M.B.A

SC-BA-02: DATA MINING
(2019 Pattern) (Semester - II) (206 BA)

Time : 2½ Hours]

[Max. Marks : 50

Instructions to the candidates:

- 1) All questions are compulsory.
- 2) Figure to the right indicate marks for questions/sub questions.

Q1) Solve Any Five :

[10]

- a) What is Data Mining?
- b) What is Data Preprocessing?
- c) What is Association Analysis? Give an Example.
- d) What is Clustering? List the methods of clustering.
- e) What is Classification? Name any two Algorithms used for it.
- f) What is big data Analysis?
- g) What is ratio data? Write any two characteristics of ratio data.
- h) What is the role of Business intelligence in decision making?

Q2) Solve Any Two :

[10]

- a) Why data cleaning is needed before data analysis?
- b) Explain Hierarchical clustering giving a suitable example.
- c) Explain Decision - tree Approach of data classification.

P.T.O.

- Q3)** Apply Apriori Algorithm to the given dataset to find frequent itemsets. (Given support value = 40%) [10]

Tid	Items Purchased
100	Bread, Milk, Cake
101	Bread, Diaper, Beer
102	Milk, Diaper, Beer, Eggs
103	Bread, Milk, Diaper, Beer
104	Bread, Milk, Diaper, Cake

OR

- Consider the dataset given below and cluster the dataset by using Hierarchical clustering and plot the dendrogram for it. [10]

Item	A	B	C	D	E
A	0				
B	7	0			
C	2	5	0		
D	6	4	8	0	
E	10	4	3	7	0

- Q4)** Explain the use of Association Analysis in purchasing behaviour of the customers. [10]

OR

- Explain the Density - based Clustering method giving a suitable example. [10]

- Q5)** Elaborate the use of data mining in target Marketing. [10]

OR

- Elaborate the use of data mining for customer profiling. [10]
