

(3 Hours)

[Total Marks:80]

**NB: 1. Question no. 1 is compulsory.**

2. Answer any **three** out of the **remaining** questions.
3. Assume data, if missing, with justification.

**Q.1.(a)** Describe different types of attributes with example. [05]

**(b)** Explain KDD process with diagram. [05]

**(c)** Define and explain: **i) Support ii) Confidence iii) Information Gain iv) Entropy v) Gini index** [05]

**(d)** Apply K-means Algorithm to divide the given set of values {2,3,6,8,9,12,15,18,22} into 3 clusters. [05]

**Q.2.(a)** Explain DBSCAN clustering algorithm with an example [10]

**(b)** Explain Regression. Explain linear regression with example. [10]

**Q.3.(a)** Suppose we have five objects with name A, B, C, D and E. Apply single linkage clustering and draw dendrogram for the given data. [10]

	X	Y
<b>A</b>	1	1
<b>B</b>	1.5	1.5
<b>C</b>	5	5
<b>D</b>	3	4
<b>E</b>	4	4
<b>F</b>	3	3.5

**(b)** What is an outlier? Describe methods that are used for outlier analysis. [10]

**Q4.(a)** Using the given training dataset classify the following tuple using Naïve Bayes Algorithm: [10]  
 <Homeowner: No, Marital Status: Married, Job experience:3>

Homeowner	Marital Status	Job experience (in years)	Defaulted
Yes	Single	3	No
No	Married	4	No
No	Single	5	No
Yes	Married	4	No
No	Divorced	2	Yes
No	Married	4	No
Yes	Divorced	2	No
No	Married	3	Yes
No	Married	3	No
Yes	Single	2	Yes

(b) What are multiple level and multidimensional association rules? Explain with suitable examples for each. [10]

Q5. (a) Explain Business Intelligence issues [10]

(b) Explain Market-Basket analysis with example. [10]

Q6. (a) What is data visualization? Explain any 3 visualization techniques with example. [10]

(b) Suppose that data for analysis includes the attribute age. The age values for data tuples are (in increasing order): [10]

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

i) What is mean of data? What is median of data?

ii) What is mode of data? Comment on data's modality.

iii) What is mid-range of data?

iv) Give the five- point summary of the data.

v) Show box plot of the data.

-----