(3 Hours)   **Total Marks: 80**

N.B.: (1) Q.1 is compulsory.
    (2) Attempt any three out of remaining five.
    (3) Figures to the right indicate full marks.

Q 1A) Define data mining. State different data mining techniques. Explain KDD process in   **[10]**
detail?

B) Describe the major functional components of a data warehouse and illustrate its   **[10]**
consequent multi-tier architecture with a diagram.

Q 2 A) Define classification. Explain decision tree with suitable example   **[10]**

B) Explain Iterative Deterministic Hill Climber and Stochastic Hill Climber Algorithm in
detail?

    [10]

Q 3A) Highlight and illustrate,usingappropriate examples, some of the ethical issues   **[10]**
associated with adoption of business intelligence methodologies and data mining
methods in the making of organizational decisions.

B) What is Market Basket Analysis?Illustrate with examples of its real-world applications   **[10]**

Q 4A) Illustrate with an example OLAPoperations :Roll-Up,Drill-Down,Slice,Dice.   **[10]**

Q 4 B) Define Clustering.For the given dataset, apply Naïve Bayes algorithm and predict the   **[10]**
outcome for a car{RED,DOMESTIC,SUV}.

COLOR	TYPE	ORIGIN	STOLEN
RED	SPORTS	DOMESTIC	YES
RED	SPORTS	DOMESTIC	NO
RED	SPORTS	DOMESTIC	YES
YELLOW	SPORTS	DOMESTIC	NO
YELLOW	SPORTS	IMPORTED	YES
YELLOW	SUV	IMPORTED	NO
YELLOW	SUV	IMPORTED	YES
YELLOW	SUV	DOMESTIC	NO
RED	SUV	IMPORTED	NO
RED	SUV	IMPORTED	YES

Q 5A) Describe mathematical methods, distance methods, and logic methods as broad   **[10]**
categories of prediction methods. Illustrate one or two popular prediction techniques
within each of these categories.

B) Generate the clusters for the following dataset using K-means clustering (k = 2).   **[10]**

Person	Weight (W)	Height (H)
P1	102	147
P2	130	162
P3	111	147
P4	170	182
P5	175	180
P6	132	157

Q   6A) What is A-Priori algorithm? Find out strong association rules for the given dataset   **[10]**
where support = 50% and confidence = 50%

Transaction Id	Item Sets
T1	A,B,C
T2	A,C
T3	A,D
T4	B,E,F

B) Write a short note on the following: [5 marks each]   **[10]**
i) Data Mart
ii) ETL Process

_____

18FED2810B9E8F1D566A2D25F1A41CFE