(Time: 3 Hours)   **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) List different data mining techniques. Explain KDD process in detail? **[10]**

B) Elucidate, with examples, some of the ethical issues pertaining to adoption of business intelligence methodologies and data mining techniques for organizational decision making. **[10]**

Q 2 A) Discuss Star and snow flake schema with a suitable example **[10]**

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

Q 3A) Define classification. Explain decision tree with suitable example **[10]**

B) Define Regression. Given the following data where x is the no. of years of experience of students and y is the salary data, use linear regression to predict the salary (y) of a student if his years of experience is 10 (x). **[10]**

x years of experience	y salary in thousands
3	30
8	57
9	64
13	72
3	36
6	43
11	59
21	90
1	20
16	83

Q 4A) Explain the concept of a data-cube as a multi-dimensional data model. What is the role of concept hierarchies in defining the dimensions of a data-cube? Illustrate, with an example. **[10]**

B) What are the characteristics and benefits of data marts? **[10]**

Q 5A) Explain the Page Ranking Algorithm with respect to web mining. **[10]**

B) Apply Naïve Bayes algorithm and predict that if a fruit has the following properties then which type of fruit it is. Fruit{Yellow,Sweet,Long} **[10]**
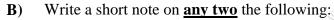
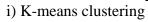| Fruit	Yellow	Sweet	Long	Total
Mango	350	450	0	650
Banana	400	300	350	400
Others	50	100	50	150
Total	800	850	400	1200

1685EEF251B9899113CE47ECAC40C5C7

Q 6A) What is Market Basket Analysis? Find out strong association rule form the given example using apriori algorithm with the support of 50% and confidence of 70% **[10]**

Trans_Id	Item
1.	Laptop, Mouse, Headphones, Pendrive, Speakers
2.	Laptop, Headphones
3.	Laptop, Mouse, Pendrive
4.	Mouse, Speakers
5.	Laptop, Pendrive

B) Write a short note on **any two** the following: **[10]**
   i) K-means clustering
   ii) text mining
   iii) OLAP operations

_____