

(3 Hours)

(Total Marks : 80)

**N.B.** 1. Question 1 is compulsory

2. Attempt any three questions out of the remaining five questions

- Q.1 (a) Give any two techniques of data preprocessing (5)  
(b) Explain Confusion matrix with one example (5)  
(c) Give the steps of K means clustering algorithm (5)  
(d) Explain concept hierarchy with example (5)
- Q.2 (a) Write a short note on Naïve Bayesian classification (10)  
(b) Explain bagging technique (10)
- Q.3 (a) Explain density based clustering (10)  
(b) What do you mean by outlier? Give the types of it. (10)
- Q.4 (a) Explain market basket analysis (10)  
(b) Explain how the efficiency of Apriori algorithm is improved (10)
- Q.5 (a) Suppose a group of 12 sales price records has been stored as follows: (10)  
30, 36, 47, 50, 52, 52, 56, 60, 63, 70, 70, 110  
Find mean, median, mode, inter quartile range (IQR)  
(b) What is BI? Define decision support system. (10)
- Q.6 (a) What is an attribute? Explain its types. (10)  
(b) Explain the knowledge discovery process with diagram (10)

\*\*\*\*\*