

Total number of printed pages-3

3 (Sem-6/CBCS) CSC HE 4

2023

COMPUTER SCIENCE

(Honours Elective)

Paper : CSC-HE-6046

(Data Mining)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following : 1×7=7
- (a) What is knowledge discovery in Database ?
 - (b) Define meta data.
 - (c) What is data visualization ?
 - (d) Define a frequent set.
 - (e) What is clustering ?

Contd.

- (f) Define K-clusters.
- (g) What is strongly connected attributes?
2. Answer the following questions : $2 \times 4 = 8$
- (a) What are the advantages of data mining?
- (b) Discuss the importance of discovering association rules.
- (c) What do you mean CART ?
- (d) What is a classification problem ? What is supervised classification ?
3. Answer **any three** of the following questions : $5 \times 3 = 15$
- (a) Illustrate the use of ID3 algorithm with an example.
- (b) Write down the CLARANS algorithm.
- (c) What are the disadvantage of the decision tree over other classification techniques ?
- (d) Write a computer program to implement the BIRCH algorithm.
- (e) Describe the salient features of CURE clustering techniques.

4. Answer **any three** of the following questions : $10 \times 3 = 30$
- (a) Develop the Apriori algorithm for generating frequent item set.
- (b) Define support and confidence in transaction. What is upward and downward closure property of item set?
- (c) Discuss the usefulness of data mining in e-commerce.
- (d) What are the problems of association rule with item-constraints ? Propose a method for this.
- (e) How is CLARANS different from CLARA ? Illustrate this using a small example.
- (f) Describe the working of the DBSCAN algorithm. Explain the concept of a cluster as used in DBSCAN.
-