B.Voc. (Software Development)   (Sem.–2)
CONCEPTS OF DATA MINING
Subject Code : 5.GV.06
M.Code : 77038
Date of Examination : 06-07-22

Time : 3 Hrs.   Max. Marks : 30

INSTRUCTIONS TO CANDIDATES :
1. **SECTION-A is COMPULSORY consisting of TEN questions carrying ONE mark each.**
2. **SECTION-B contains FIVE questions carrying 2$^1/_2$ (Two and Half) marks each and students have to attempt any FOUR questions.**
3. **SECTION-C contains THREE questions carrying FIVE marks each and students have to attempt any TWO questions.**

SECTION-A

1. **Write short notes on :**

   a) Define Data Mart.

   b) What is Data Cube?

   c) What is a Decision Tree?

   d) Define ROLAP.

   e) How is the accuracy of a classifier measured?

   f) Differentiate between OLAP and OLTP.

   g) Why do we use data visualization?

   h) What is the use of multi-dimensional view?

   i) What is association rule?

   j) What is cross validation?

SECTION-B

2. Compare the Snowflake Schema, Star Schema model with help of an example.

3. Differences between Operational Database Systems and Data Warehouses.

4. Write note on following methods :

    a) Decision Tree

    b) Naive Bayes

5. What are the major challenges of mining a huge amount of data (e.g., billions of tuples) in comparison with mining a small amount of data (e.g., data set of a few hundred tuple)?

6. Why is Data Pre-processing required? Explain the different steps involved in Data Pre-processing.

SECTION-C

7. What is Data Warehouse? Describe the components of Data Warehouse in detail.

8. Discuss different types of cluster analysis? Explain in brief each one with an example.

9. What is association rule mining? Explain Apriori algorithm in data mining.

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.