BACHELOR OF TECHNOLOGY (CBCS) (2020 COURSE)
B.Tech.Sem - VI INFORMATION TECHNOLOGY : SUMMER : 2024
SUBJECT: DATA WAREHOUSING & DATA MINING

Day : Friday
Date : 24/05/2024

S-24745-2024

Time : 02:30 PM-05:30 PM
Max. Marks : 60

N.B :

1) All questions are **COMPULSORY**.
2) Figures to the right indicate **FULL** marks.
3) Draw neat and labeled diagrams **WHEREVER** necessary.
4) Assume suitable data if necessary.

Q.1 Illustrate the generic architecture of a data warehouse. Discuss all the components in details. **(10)**

OR

Q.1 Define multidimensional data. Identify any one real time case study of multidimensional data and discuss how is it put to use in data warehouse. **(10)**

Q.2 What are data objects and types of attributes in data mining? Discuss in detail. Discuss the importance of statistical data in data mining. **(10)**

OR

Q.2 What are data similarity and dissimilarity measures in data mining? What are the different methods of calculating similarity and dissimilarity? Discuss in detail. **(10)**

Q.3 What is a frequent pattern? How is frequent pattern mining significant? Explain the categories of frequent pattern mining. **(10)**

OR

Q.3 How to measure if the association rule is qualified or not, for a particular data set? Explain with a suitable example. **(10)**

Q.4 What is overfitting in classification? How does it affect decision tree? Does pruning reduce overfitting? Justify your answer. **(10)**

OR

Q.4 How are IF-THEN rules applicable in classification? With a suitable example, explain the rule extraction from a decision tree. **(10)**

Q.5 Define clustering. Categorize the clustering techniques and explain any one of them in detail. **(10)**

OR

Q.5 What is an outlier? What is the purpose of outlier analysis? Discuss any one technique of outlier detection. **(10)**

Q.6 What are the different categories of learning algorithms? Enlist and specify which of these algorithms can be implemented in WEKA? Explain the implementation of any one of them in detail. **(10)**

OR

Q.6 What is WEKA? What are the features of WEKA? What are the data mining operations that can be performed using WEKA? **(10)**

* * * * *