

II B. Tech II Semester Supplementary Examinations, December - 2023

DATA WAREHOUSING AND MINING

(Common to CSE(AIML),CSE(AI),CSE(DS),CSE(AIDS),AIDS,AIML& CSD)

Time: 3 hours

Max. Marks: 70

Answer any **FIVE** Questions each Question from each unitAll Questions carry **Equal** Marks

~~~~~

## UNIT-I

- 1 a) Explain the steps in designing the data warehouse. [7M]  
 b) Enumerate the concept of multidimensional data model. [7M]

**Or**

- 2 a) Explain the following in OLAP [7M]  
 i) Roll up operation      ii) Drill Down operation  
 iii) Slice operation      iv) Dice operation  
 b) Write in brief about schemas in multidimensional data model. [7M]

## UNIT-II

- 3 a) What is KDD? Explain about data mining as a step in the process of knowledge discovery. [7M]  
 b) Illustrate on Feature Subset Selection. [7M]

**Or**

- 4 a) Discuss about Variable Transformation with example. [7M]  
 b) Discuss Data reduction in detail. [7M]

## UNIT-III

- 5 a) What do you mean Classification? Discuss in detail general approach to solve a classification problem. [7M]  
 b) Explain decision tree induction algorithm for classifying data tuples and discuss suitable Example. [7M]

**Or**

- 6 a) Explain the role of cross-validation in data mining process. [7M]  
 b) Elaborate the Bayes Theorem with a suitable example. [7M]

## UNIT-IV

- 7 a) Explain the Apriori Algorithm and also mention its drawbacks. [9M]  
 b) Discuss the various applications of association analysis. [5M]

**Or**

- 8 a) Explain the FP Growth Algorithm. [9M]  
 b) Can we design a method that mines the complete set of frequent item sets without candidate generation? If yes, explain with an example. [5M]

## UNIT-V

- 9 a) Discuss K-means clustering algorithm with step by step. [7M]  
 b) Illustrate density based and Grid based clustering. [7M]

**Or**

- 10 a) Discuss various Data mining application. [7M]  
 b) Explain DB scan strength and weakness. [7M]