Subject- DWM TE SEM V(Academic year 2025-26)

QUESTIONS BANK -UNIT 1, UNIT 2

Unit 1

Q1 .Write short note on: Operational support system.

Q2. Explain about datawarehouse and characteristics of datawarehouse

Q3. Describe datawarehouse architecture.Also differentiate between Data Warehouse and Data Mart

Q4.Differentiate top-down and bottom-up approaches for building data warehouses. Discuss the merits and limitations of each approach.

Q5. What is meant by Metadata in the context of a Data warehouse? Explain the different types of metadata stored in a data warehouse. Illustrate with a suitable example.

Q6. What is metadata?Why do we need metadata when search engines like google seem so effective?

Q7 Explain the steps in dimensional modeling and its advantages of dimensional modeling

Q8. Explain with examples Factless Fact Table and Fact constellation.

Q9. Describe the process of Extraction, Transformation and Loading (ETL)with a neat and labeled diagram.

Q10. Discuss various OLAP models and their architecture.

Q11. Describe about the following OLAP operations on a cube:

i) Rollup ii) Drill down iii) Slice iv) Dice v) Pivot.

Q12. All Design Star Schema case study problems

Unit 2

Q1. Describe the steps involved in knowledge Discovery in Databases (KDD)

Q2. Explain types of attributes with example

Q3 In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem .

Q4. Explain different normalization techniques such as min-max normalization, z-score and decimal scaling with example

Q5. Discuss different steps involved in data preprocessing.

Q6. Numerical on- i) Binning Method ii) 5 steps summary(mean, mode, medium, boxplot, histogram)

Q7. Explain different Data visualization techniques

Q8. Explain data discretization and concept hierarchy generation

Q9. Describe any five issues in data mining.

Q10. Describe any five applications of data mining.

Q11.For the given data apply smoothing by a) bin- mean, b) bin- boundaries

8,9,15,30,16,24,26,27,21,21,30,34

Q12. For the given data values 5,10,11,12,12,12,13,14,15,16,18,22 find,

Mean, Median, Mode, Midrange,Q1,Q2,Q3, five number summary, also draw box plot

Q13.explain the following data preprocessing methods

i) data dimensionality reduction ii) data transformation and discretization

Unit 3

1) Explain Decision Tree Induction algorithm
2) Explain how Naive Bayes classification Algorithm
3) Explain how Naive Bayes makes predictions and discuss the "naive" assumption in Naive Bayes. Provide an example to illustrate the application of Naive Bayes in a real-world scenario.
4) Describe in detail about how to evaluate accuracy of the classifier
5) Explain attribute selection measures of Decision tree classifier i) Information gain ii) Gini Index with an example

6)Explain Holdout and random subsampling method for evaluating a classifier

7) Numericals on

i)Naive Bayes Classification algorithm

ii) Decision Tree algorithm