

Seat No. 6309

OCT_NOV_2024 WINTER EXAMINATION**Bachelor of Computer Application(BCA) NEP 2.0****Sub. Name: Data Warehousing and Data Mining****Sub. Code: 108795/96643****Day and Date: NOVEMBER ,27-11-2024****Total Marks: 80****Time: 10:30 AM To 01:30 PM****Instructions: 1. Figures to the right indicate full marks****Special Inst.: Que. 1 and Que. 8 are compulsory and attempt any 3 questions from Que. No. 2 to Que. No. 7****Q1) Solve following MCQ.****[12]**

- i. Which of the following is not the type of OLAP
 - A. Relational OLAP
 - B. Multidimensional OLAP
 - C. Desktop OLAP
 - D. Object Oriented OALP
- ii. The key function of OLAP is to _____
 - A. Support real-time transaction processing
 - B. Perform multi-dimensional data analysis
 - C. Store historical data
 - D. Retrieve unstructured data
- iii. In a 3-tier data warehouse architecture, the middle tier is responsible for _____
 - A. Data transformation
 - B. Data retrieval and analysis
 - C. Data storage
 - D. Data entry
- iv. Data Mart can be best described as _____
 - A. A small subset of a data warehouse
 - B. A volatile data store
 - C. An unorganized data source
 - D. A single large database
- v. From where are classification rules extracted?
 - A. Branches
 - B. Decision tree
 - C. Siblings
 - D. Root node

[1]**P.T.O.**

- vi.refers to the overall process of discovering useful knowledge from data.
- A. KCC
 - B. KLL
 - C. KDD
 - D. None of above
- vii. Which one of the following can be considered as the correct application of the data mining?
- A. Fraud detection
 - B. Corporate Analysis & Risk management
 - C. Management and market analysis.
 - D. All of the above
- viii.is defined as the removal of noisy and irrelevant data from the collection.
- A. Data extraction
 - B. Data cleaning
 - C. Data Modeling
 - D. None of above
- ix. Which of the following is a classification algorithm?
- A. K-Means
 - B. Decision Tree
 - C. Apriori
 - D. Association rule mining
- x. In a star schema, which component connects to the fact table?
- A. Dimension table
 - B. Factless table
 - C. OLTP table
 - D. Metadata table
- xi. Which of the following best describes a data warehouse?
- A. A collection of integrated and subject-oriented databases
 - B. A real-time data processing system
 - C. A decentralized storage system for large datasets
 - D. An online transaction processing system (OLTP)
- xii. Which algorithm is used for finding frequent itemsets?
- A. K-Means
 - B. Apriori algorithm
 - C. Decision tree
 - D. Naïve Bayes

Q2) [16]

[2]

P.T.O.



- a. Define data mining? Explain kdd process in detail. [8]
- b. Discuss applications of data mining. [8]

Q3) What is multi-dimensional data model? Explain data cube and star model in detail. [16]

Q4) [16]

- a. What is the significance of ETL (Extract, Transform, Load) processes in a Data Warehouse? [8]
- b. What is a Data Warehouse? Explain main characteristics of Data Warehouse. [8]

Q5) [16]

- a. Describe the Bayesian classification method. What are its strengths and limitations? [8]
- b. Why is Data pre-processing essential in Data mining? Describe the key steps involved in Data pre-processing [8]

Q6) Discuss in detail frequent sub graph mining, tree mining and sequence mining. [16]

Q7) [16]

- a. Explain in detail partitioning methods like K-mean and K-medoids. [8]
- b. Discuss different clustering methods in detail. [8]

Q8) Write short note (Any Four) [20]

- a. Functionalities issues in data mining [5]
- b. Database system Vs Data Warehouse [5]
- c. Online analytical processing (OLAP) [5]
- d. Apriori Algorithm [5]
- e. Snowflake Schema [5]
- f. Advantages of Data Warehouse [5]