

Serial No. of Booklet :

Code No. : 1297

B. C. A. (Fourth Semester) Examination, 2023-24

AFFIX PRESCRIBED  
RUBBER STAMP

Paper Third

DATA WAREHOUSING AND DATA MINING

Course Code—BCA-403T (Major)

In Figures (अंकों में) :

|  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|

Roll No.

In Words (शब्दों में) :

---

Date : . . . . .

Time : 2 Hrs.

Signature of Invigilator

कक्ष निरीक्षक के हस्ताक्षर

Max. Marks : 75

**Important Instructions :**

1. The candidate will write his/her Roll Number only at the places provided for, i. e., on the cover page and on the OMR answer sheet at the end and nowhere else.

2. Immediately on receipt of the question booklet, the candidate should check up the pages and that no question is missing. If the candidate finds any discrepancy in the question booklet, he/she should report the invigilator within 10 minutes of the issue of this booklet and a fresh question booklet without any discrepancy is obtained.

3. No second question booklet shall be given to a candidate under any circumstances after 10 minutes. The candidate should be careful in handling the question booklet and in filling the OMR answer sheet given separately with this booklet.

**महत्वपूर्ण निर्देश :**

1. अभ्यर्थी अपने अनुक्रमांक केवल उन्हीं स्थानों पर लिखेंगे जो इसके लिए दिये गये हैं, अर्थात् प्रश्न पुस्तिका के मुख्य पृष्ठ तथा साथ दिये गये ओ. एम. आर. उत्तर पत्र पर, तथा अन्यत्र कहीं नहीं लिखेंगे।

2. प्रश्न पुस्तिका मिलते ही अभ्यर्थी को जाँच करके सुनिश्चित कर लेना चाहिए कि पुस्तिका में पूरे पृष्ठ हैं और कोई प्रश्न छूटा तो नहीं है। यदि कोई विसंगति है तो प्रश्न पुस्तिका मिलने के 10 मिनट के भीतर ही कक्ष परिप्रेक्षक को सूचित करना चाहिए और बिना झुट्टि की दूसरी प्रश्न पुस्तिका प्राप्त कर लेना चाहिए।

3. किसी भी परिस्थिति में 10 मिनट बाद अभ्यर्थी को दूसरी प्रश्न पुस्तिका नहीं मिलेगी। अभ्यर्थी को प्रश्न पुस्तिका को उपयोग में लाने और ओ. एम. आर. उत्तर पत्र को भरने में सावधानी बरतनी चाहिए।



(Contd. on the last page / अन्य निर्देश अन्तिम पृ. पर)

1. What is a Data Warehouse primarily designed for ?
  - (A) Real-time transaction processing
  - (B) Historical data analysis and reporting
  - (C) Online gaming
  - (D) Social media interaction
2. What are the components of a Data Warehouse ?
  - (A) Web servers, databases and applications
  - (B) Input devices, processing units and output devices
  - (C) Tables, queries and forms
  - (D) Data sources, ETL processes and storage
3. What is the process of building a Data Warehouse ?
  - (A) Data Mining
  - (B) Data Analysis
  - (C) Data Integration
  - (D) Data Visualization
4. What is the difference between a Database System and a Data Warehouse ?
  - (A) A database is used to store and capture data, while a data warehouse is used for data analytics.
  - (B) A database is used for data analytics, while a data warehouse is used to store and capture data.
  - (C) A database is used for data analytics, while data warehouses are designed for transactional data storage.
  - (D) None of the above
5. What is a multi-dimensional data model in Data Warehousing ?
  - (A) A model that stores data in a single dimension
  - (B) A model that stores data in two dimensions
  - (C) A model that stores data in more than two dimensions
  - (D) A model that stores data in a non-relational format

6. What are data cubes, stars, snowflakes and fact constellations in Data Warehousing ?
- (A) Architectural elements of a Data Warehouse
  - (B) Types of data storage devices
  - (C) Data encryption techniques
  - (D) Data visualization tools
7. What is the first step in the Data Warehouse process and technology ?
- (A) Data Analysis
  - (B) Data Integration
  - (C) Data Mining
  - (D) Data Visualization
8. What is the strategy behind Data Warehousing ?
- (A) Storing data in a flat file format
  - (B) Deleting historical data
  - (C) Performing real-time data analysis
  - (D) Extracting data from multiple sources and integrating it into a single repository
9. What are the management and support processes involved in Data Warehousing ?
- (A) Hardware maintenance
  - (B) Data backup and recovery
  - (C) Network security
  - (D) Web development
10. What is involved in the planning and implementation of a Data Warehouse ?
- (A) Setting up social media accounts
  - (B) Creating data visualizations
  - (C) Designing data models and schemas
  - (D) Building mobile applications
11. \_\_\_\_\_ is a subject-oriented integrated, time-variant, nonvolatile collection of data in support of management decisions.
- (A) Data Mining
  - (B) Data Warehousing
  - (C) Web Mining
  - (D) Text Mining

12. What is the primary purpose of a Data Warehouse ?
- (A) Historical data analysis and reporting
  - (B) Real-time transaction processing
  - (C) Online gaming
  - (D) Social media interaction
13. The data is stored, retrieved and updated in \_\_\_\_\_.
- (A) OLAP
  - (B) OLTP
  - (C) SMTP
  - (D) FTP
14. What is the process of selecting only relevant data from a larger dataset called :
- (A) Sampling
  - (B) Filtering
  - (C) Data reduction
  - (D) Data integration
15. Which of the following is not a commonly used data warehouse architecture ?
- (A) Inmon
  - (B) Kimball
  - (C) Snowflake
  - (D) Bayesian
16. Which of the following is not a common data mining application ?
- (A) Fraud detection
  - (B) Customer segmentation
  - (C) Sales forecasting
  - (D) Inventory management
17. The star schema is composed of \_\_\_\_\_ fact table.
- (A) one
  - (B) two
  - (C) three
  - (D) four

8. The modern CASE tools belong to \_\_\_\_\_ category.

- (A) analysis
- (B) development
- (C) coding
- (D) delivery

19. Data scrubbing is :

- (A) a process to reject data from the data warehouse and to create the necessary indexes.
- (B) a process to load the data in the data warehouse and to create the necessary indexes.
- (C) a process to upgrade the quality of data after it is moved into a data warehouse.
- (D) a process to upgrade the quality of data before it is moved into a data warehouse.

20. What is the primary goal of data mining ?

- (A) Real-time data processing
- (B) Storing historical data for reporting
- (C) Conducting online transactions
- (D) Extracting meaningful patterns from large datasets

21. What is the main motivation behind data mining ?

- (A) Generating random data
- (B) Predictive modeling
- (C) Data visualization
- (D) Network security

22. What is data pre-processing in the context of data mining ?

- (A) The process of cleaning and transforming raw data
- (B) Encrypting data for security purposes
- (C) Storing data in a database
- (D) Analyzing data patterns

23. How is noisy data typically addressed in data cleaning ?

- (A) By using clustering techniques
- (B) Through human inspection
- (C) By ignoring it
- (D) None of the above

24. What does data reduction involve in data mining ?
- (A) Increasing the volume of data for analysis
  - (B) Decreasing the amount of data while preserving meaningful information
  - (C) Randomly sampling data points
  - (D) Ignoring data outliers
25. What technique is used for data cube aggregation ?
- (A) Clustering
  - (B) Regression analysis
  - (C) Dimensionality reduction
  - (D) Aggregating data along multiple dimensions
26. What is dimensionality reduction in data mining ?
- (A) Increasing the number of dimensions in a dataset
  - (B) Decreasing the number of dimensions in a dataset while preserving meaningful information
  - (C) Storing data in a hierarchical structure
  - (D) Grouping similar data points together
27. What is data compression in data mining ?
- (A) Reducing the size of the dataset without losing information
  - (B) Increasing the size of the dataset for better analysis
  - (C) Encrypting data for security purposes
  - (D) Storing data in a compressed format
28. Which technique is used for numerosity reduction ?
- (A) Clustering
  - (B) Regression analysis
  - (C) Discretization
  - (D) Random sampling
29. What is concept hierarchy generation in data mining ?
- (A) Creating a hierarchical structure for organizing data attributes
  - (B) Removing irrelevant data from the dataset
  - (C) Applying statistical techniques to analyze data patterns
  - (D) Predicting future trends based on historical data

30. What is one form of data cleaning technique used for handling inconsistent data ?
- (A) Clustering
  - (B) Regression analysis
  - (C) Computer inspection
  - (D) Random sampling
31. What is the main objective of data mining ?
- (A) Real-time data processing
  - (B) Extracting meaningful patterns from large datasets
  - (C) Storing historical data for reporting
  - (D) Conducting online transactions
32. What drives the motivation behind data mining ?
- (A) Generating random data
  - (B) Data visualization
  - (C) Network security
  - (D) Predictive modeling
33. To represent any n-Dimension data we need a series of \_\_\_\_\_ Dimension cubes.
- (A)  $(n - 1)$
  - (B)  $n$
  - (C)  $n + 1$
  - (D)  $n + 2$
34. Which technique is commonly used to handle missing values in data cleaning ?
- (A) Clustering
  - (B) Regression
  - (C) Binning
  - (D) Computer inspection
35. What is data integration in data mining ?
- (A) Combining data from multiple sources into a single repository
  - (B) Separating data into different categories
  - (C) Encrypting data for security
  - (D) Deleting irrelevant data

36. The \_\_\_\_\_ operation performs a selection on one dimension of the given cube, resulting in a subcube.
- (A) Pivot
  - (B) Slice
  - (C) Roll-up
  - (D) Drill-down
37. What is classification in the context of data mining ?
- (A) Sorting data into different categories
  - (B) Combining multiple datasets into a single repository
  - (C) Storing historical data for analysis
  - (D) Randomly sampling data points
38. What is data generalization ?
- (A) Adding noise to data for privacy protection
  - (B) Removing irrelevant attributes from the dataset
  - (C) Analyzing the statistical significance of data
  - (D) Summarizing data at a higher level of abstraction
39. What is analytical characterization in data mining ?
- (A) Sorting data into different categories
  - (B) Removing duplicates from the dataset
  - (C) Examining the distribution of data values
  - (D) Adding noise to data for privacy protection
40. What does analysis of attribute relevance involve ?
- (A) Sorting data into different categories
  - (B) Removing outliers from the dataset
  - (C) Generating association rules between attributes
  - (D) Identifying the significance of each attribute in relation to the classification task



41. Which type of algorithms use statistical measures in large databases for classification ?
- (A) Statistical-based algorithms
  - (B) Distance-based algorithms
  - (C) Decision tree-based algorithms
  - (D) Clustering algorithms
42. What do distance-based algorithms rely on for classification ?
- (A) Statistical measures
  - (B) Decision trees
  - (C) Association rules
  - (D) Similarity and distance measures
43. Which algorithm is a decision tree-based algorithm commonly used for classification ?
- (A) K-means
  - (B) C4.5
  - (C) CURE
  - (D) DBSCAN
44. What is clustering in the context of data mining ?
- (A) Sorting data into different categories
  - (B) Grouping similar data points together
  - (C) Storing historical data for analysis
  - (D) Randomly sampling data points
45. What are similarity and distance measures used for in clustering ?
- (A) Analyzing the distribution of data values
  - (B) Sorting data into different categories
  - (C) Determining the similarity between data points
  - (D) Adding noise to data for privacy protection
46. Which algorithm is an example of a hierarchical clustering algorithm ?
- (A) K-means
  - (B) DBSCAN
  - (C) OPTICS
  - (D) CURE

47. Which clustering method is density-based ?
- (A) DBSCAN
  - (B) Hierarchical clustering
  - (C) Partitional clustering
  - (D) STING
48. What is the goal of grid-based clustering methods ?
- (A) Grouping data points based on their density
  - (B) Separating data into different categories
  - (C) Analyzing the statistical significance of data
  - (D) Summarizing data at a higher level of abstraction
49. What is the model-based method for clustering based on ?
- (A) Similarity measures
  - (B) Decision trees
  - (C) Statistical approach
  - (D) Distance measures
50. What is association rules mining ?
- (A) Sorting data into different categories
  - (B) Finding interesting relationships between variables in large datasets
  - (C) Analyzing the distribution of data values
  - (D) Randomly sampling data points
51. What do large item sets represent in association rules mining ?
- (A) Groups of similar data points
  - (B) Outliers in the data
  - (C) Missing values
  - (D) Frequent patterns in the dataset
52. Which algorithm is commonly used for generating association rules ?
- (A) K-means
  - (B) Apriori
  - (C) CURE
  - (D) DBSCAN

3. What is the objective of parallel and distributed algorithms in association rules mining ?
- (A) Analyzing the distribution of data values
  - (B) Sorting data into different categories
  - (C) Speeding up the mining process for large datasets
  - (D) Randomly sampling data points
54. What approach does the neural network approach take in association rules mining ?
- (A) Finding patterns in the data based on neuron activations
  - (B) Analyzing the distribution of data values
  - (C) Sorting data into different categories
  - (D) Randomly sampling data points
55. What is the primary purpose of data visualization in data warehousing ?
- (A) Presenting data in a visual format for easy understanding
  - (B) Generating random data
  - (C) Summarizing data at a higher level of abstraction
  - (D) Analyzing the statistical significance of data
56. What does aggregation involve in data visualization ?
- (A) Sorting data into different categories
  - (B) Summarizing data at a higher level of abstraction
  - (C) Removing duplicates from the dataset
  - (D) Adding noise to data for privacy protection
57. What type of information does historical information provide in data visualization ?
- (A) Real-time data processing
  - (B) Current trends and patterns in the data
  - (C) Past trends and patterns in the data
  - (D) Future predictions based on the data
58. What function does OLAP serve in data warehousing ?
- (A) Online Transaction Processing
  - (B) Online Data Mining
  - (C) Online Data Backup
  - (D) Online Analytical Processing

59. Which type of OLAP server stores data in a relational database ?
- (A) MOLAP
  - (B) HOLAP
  - (C) DOLAP
  - (D) ROLAP
60. What does the acronym MOLAP stand for ?
- (A) Massive Online Analytical Processing
  - (B) Multidimensional Online Analytical Processing
  - (C) Mixed Online Analytical Processing
  - (D) Mobile Online Analytical Processing
61. What is the purpose of a data mining interface ?
- (A) Analyzing the distribution of data values
  - (B) Sorting data into different categories
  - (C) Adding noise to data for privacy protection
  - (D) Presenting data mining results to users
62. What aspect of data warehousing does security address ?
- (A) Ensuring data confidentiality, integrity and availability
  - (B) Sorting data into different categories
  - (C) Presenting data in a visual format
  - (D) Adding noise to data for privacy protection
63. What is the purpose of backup and recovery in data warehousing ?
- (A) Analyzing the distribution of data values
  - (B) Storing data in a visual format
  - (C) Protecting against data loss
  - (D) Adding noise to data for privacy protection
64. What does tuning data warehouse involve ?
- (A) Fine-tuning database performance for better efficiency
  - (B) Sorting data into different categories
  - (C) Presenting data in a visual format
  - (D) Adding noise to data for privacy protection

65. What are some types of warehousing applications ?
- (A) Web mining, spatial mining and temporal mining
  - (B) Online gaming and social media interaction
  - (C) Data encryption and decryption
  - (D) Statistical analysis and regression modeling
66. What does web mining focus on ?
- (A) Sorting data into different categories
  - (B) Presenting data in a visual format
  - (C) Analyzing web server logs for patterns and trends
  - (D) Adding noise to data for privacy protection
67. What is the goal of spatial mining ?
- (A) Sorting data into different categories
  - (B) Analyzing geographical data for patterns and trends
  - (C) Presenting data in a visual format
  - (D) Adding noise to data for privacy protection
68. What does temporal mining focus on ?
- (A) Analyzing historical data for patterns and trends over time
  - (B) Sorting data into different categories
  - (C) Presenting data in a visual format
  - (D) Adding noise to data for privacy protection
69. What type of OLAP server stores data in a multidimensional cube structure ?
- (A) ROLAP
  - (B) HOLAP
  - (C) DOLAP
  - (D) MOLAP
70. What is the main purpose of data warehousing ?
- (A) Real-time data processing
  - (B) Long-term storage and analysis of data
  - (C) Generating random data
  - (D) Conducting online transactions

71. Which component is essential for building a data warehouse ?
- (A) Operating system
  - (B) Web browser
  - (C) Database Management System (DBMS)
  - (D) Programming language
72. What is the primary function of OLAP servers in data warehousing ?
- (A) Online analytical processing
  - (B) Data encryption
  - (C) Data visualization
  - (D) Database backup
73. Which type of data processing involves summarizing data at a higher level of abstraction ?
- (A) Aggregation
  - (B) Segmentation
  - (C) Classification
  - (D) Regression
74. What is the purpose of data cleaning in data warehousing ?
- (A) Adding noise to the dataset
  - (B) Increasing the size of the dataset
  - (C) Enhancing data security
  - (D) Removing irrelevant data
75. What does OLAP allow users to do ?
- (A) Conduct online transactions
  - (B) Analyze multidimensional data interactively
  - (C) Encrypt data for security purposes
  - (D) Create random datasets
76. Which of the following is not a component of data warehousing ?
- (A) OLAP server
  - (B) Database management system
  - (C) Web browser
  - (D) Data warehouse
77. Data warehouse contains ..... data that is never found in the operational environment.
- (A) normalized
  - (B) informational
  - (C) summary
  - (D) denormalized
78. Which of the following is an essential process in which the intelligent methods are applied to extract data patterns ?
- (A) Warehousing
  - (B) Text Mining
  - (C) Data Selection
  - (D) Data Mining

79. What is KDD in data mining ?  
 (A) Knowledge Discovery Data  
 (B) Knowledge Data Definition  
 (C) Knowledge Discovery Database  
 (D) Knowledge Data House
80. For what purpose, the analysis tools pre-compute the summaries of the huge amount of data ?  
 (A) To obtain the queries response  
 (B) In order to maintain consistency  
 (C) For authentication  
 (D) For data access
81. The Data Warehouse is .....  
 (A) write only  
 (B) read write only  
 (C) read only  
 (D) None of the above
82. Expansion for DSS in Data Warehouse is .....  
 (A) Decision Single System  
 (B) Data Storable System  
 (C) Data Support System  
 (D) Decision Support System
83. .... describes the data contained in the data warehouse.  
 (A) Relational data  
 (B) Metadata  
 (C) Operational data  
 (D) Informational data
84. .... predicts future trends and behaviours, allowing business managers to make proactive, knowledge-driven decisions.  
 (A) Data warehouse  
 (B) Datamarts  
 (C) Data mining  
 (D) Metadata
85. .... is the heart of the warehouse.  
 (A) Data warehouse database servers  
 (B) Datamining database servers  
 (C) Datamart database servers  
 (D) Relational database servers
86. .... is the specialized data warehouse database.  
 (A) Oracle  
 (B) Redbrick  
 (C) DBZ  
 (D) Informix
87. .... defines the structure of the data held in operational databases and used by operational applications.  
 (A) Operational metadata  
 (B) User-level metadata  
 (C) Data warehouse metadata  
 (D) Data mining metadata

88. .... is held in the catalog of the warehouse database system.
- (A) Application level metadata
  - (B) Departmental level metadata
  - (C) Core warehouse metadata
  - (D) Algorithmic level metadata
89. .... maps the core warehouse metadata to business concepts, familiar and useful to endusers.
- (A) User level metadata
  - (B) End user level metadata
  - (C) Application level metadata
  - (D) Core level metadata
90. What is the difference between OLTP and OLAP ?
- (A) OLTP is optimized for analytical processing, while OLAP is optimized for transaction processing.
  - (B) OLTP is optimized for transaction processing, while OLAP is optimized for analytical processing.
  - (C) OLTP and OLAP are the same thing.
  - (D) OLTP and OLAP are both optimized for transaction processing.
91. What is a dimension table ?
- (A) A table that stores transactional data.
  - (B) A table that stores details about products.
  - (C) A table that stores metadata.
  - (D) A table that stores details about customers.
92. Which of the following is not a basic data mining task ?
- (A) Prediction
  - (B) Classification
  - (C) Spooling
  - (D) Clustering
93. Which of the following is not an issue in data mining ?
- (A) Overfitting
  - (B) Outliers
  - (C) High dimensionality
  - (D) Shortage of data



94. Which of the following is a subset of data warehouse focused on a specific functional area ?
- (A) Association rules
  - (B) Datamart
  - (C) Flat files
  - (D) Database
95. What do data warehouses support ?
- (A) OLTP
  - (B) OLAP
  - (C) OLAP and OLTP
  - (D) Operational databases
96. Which one of the following statements is not correct about the data cleaning ?
- (A) It refers to the process of data cleaning.
  - (B) It refers to the transformation of wrong data into correct data.
  - (C) It refers to correcting inconsistent data.
  - (D) All of the above
97. The issues like efficiency, scalability of data mining algorithms comes under :
- (A) Diverse data type issues
  - (B) Performance issues
  - (C) Mining methodology and user interaction
  - (D) All of the above
98. Which one of the following can be considered as the correct application of the data mining ?
- (A) Fraud detection
  - (B) Corporate analysis and Risk management
  - (C) Management and Market analysis
  - (D) All of the above
99. .... is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management decisions.
- (A) Data warehousing
  - (B) Data mining
  - (C) Web mining
  - (D) Text mining
100. Record cannot be updated in .....
- (A) files
  - (B) data warehouse
  - (C) rdbms
  - (D) None of the above