MCS-221

MASTER OF COMPUTER APPLICATION (MCA-NEW)
Term-End Examination
June, 2024

MCS–221 : DATA WAREHOUSING AND DATA MINING

Time : 3 Hours   *Maximum Marks : 100*

Weightage : 70%

Note : *Question number-1 is compulsory. Attempt any three questions from the rest.*

1. (a) With the help of a diagram, explain the real-time data warehouse architecture. Also, discuss its trade-off.   10

    (b) Describe fact constallation scheme. How is it different from snow-flake scheme ? With the help of an example explain fact constallation schema. Mention its advantages and limitations.   10

(c) Write and explain K-NN classification algorithm in Data-Mining. Discuss its advantages and disadvantages.   10

(d) Discuss vector space modeling for representing text documents. With reference to this modeling, explain TF-IDF and Inverse Document Frequency (IDF). 10

2. (a) Discuss various factors to be considered to improve the performance of ETL.   7

(b) Define a Data Mart. How is it different from a centralized data warehouse ? Discuss the structure of Data Mart with the help of an example.   8

(c) Discuss the following OLAP data cube operations :   5

(i) Roll-up   (ii) Drill-down.

3. (a) Discuss the following Data Integration issues :   10

(i) Schema Integration and object modeling

(ii) Redundancy

      (iii) Detection and Resolution of Data Value Conflicts.

  (b) Write and explain Decision Tree Classifer algorithm with the help of an example.   10

4. (a) Discuss Linear Discriminant Analysis (LDA) and Principal Component Analysis (PDA) feature extraction techniques.   10

  (b) Explain Association Rule Mining (ARM). Also, explain the two phases involved in ARM process with the help of an example.

                                        10

5. Write short notes on the following : $4 \times 5 = 20$

  (a) ELT and its need

  (b) Metadata and Data Warehousing

  (c) Rule based classification

  (d) Naive-payes Classifier

MCS–221