MCS-221

MASTER OF COMPUTER APPLICATIONS (NEW) (MCA-NEW)

Term-End Examination

December, 2021

MCS-221 : DATA WAREHOUSING AND DATA MINING

Time : 3 hours   *Maximum Marks : 100*

(Weightage : 70%)

*Note : Question no. **1** is **compulsory**. Answer any **three** questions from the rest.*

1. (a) Define a Data Warehouse. List and explain the four characteristics of a Data Warehouse. *10*

   (b) Explain the following OLAP architectures and draw their architectural diagram : *10*

   (i) Multidimensional Online Analytical Processing (MOLAP)

   (ii) Hybrid Online Analytical Processing (HOLAP)

(c) Define "data cleaning" which is a data preprocessing technique. In this context, explain the concept of Noisy data cleaning along with some suitable examples. *10*

(d) What is a Decision Tree ? How is it useful in classification ? With the help of an example, explain the process of construction of a decision tree and its representation. *10*

2. (a) What is Text Mining ? Where is it used ? Explain any two text mining techniques with the help of a suitable example for each. *10*

(b) In the context of mining multimedia data on the web, explain the following terms : *10*

  (i) Page Rank

  (ii) Hits

  (iii) Page Layout Analysis

  (iv) Vision Page Segmentation

3. (a) Define a Data Lake. Explain the step-by-step process of creating a data lake. *10*

(b) With the help of an example, explain the star schema dimensional model. *10*

4. (a) What is Metadata ? What are its contents ? Justify how metadata can be an important component in Data Warehousing ? Also mention its types. *10*

(b) Explain the three components of ETL. Also mention how to improve the ETL performance. *10*

5. Write short notes on any *four* of the following : *4×5=20*

(a) Data Marts

(b) Association Rule Generation

(c) Apriori Algorithm

(d) Clustering and its Methods

(e) Bayes' Theorem

———