

[Total No. of Questions - 9] [Total No. of Printed Pages - 2]

Dec.-23-0532

CS-605 (Data Mining & Data Warehousing)

B.Tech. 6th (CBCS)

Time : 3 Hours

Max. Marks : 60

The candidates shall limit their answers precisely within the answer-book (40 pages) issued to them and no supplementary/continuation sheet will be issued.

Note : Attempt five questions in all, selecting one question each from section A, B, C, and D. Section - E is compulsory.

SECTION - A

1. Answer the following:

- (a) Define Data-warehouse and explain four key features of data warehouse compared to other data repository systems. (5)
- (b) Explain 3-Tier architecture of data warehouse with neat diagram. (5)

2. Explain the following in OLAP:

- (a) Roll up operation
- (b) Drill Down operation
- (c) Slice operation
- (d) Dice operation
- (e) Pivot operation (5×2=10)

SECTION - B

3. What do you mean by Data Pre-processing? What are the operations in data pre-processing? Discuss each with example. (10)

2

CS-605

4. Explain star schema, snowflake schema and aggregate table with suitable examples. (10)

SECTION - C

5. What are the major clustering methods? Explain partitioning methods for clustering in detail. (10)
6. Answer the following:
- (a) What is Fuzzy Logic? Write a note on the role of Fuzzy logic in data mining with example. (5)
 - (b) What is Genetic Algorithm? Discuss its use in knowledge discovery with example. (5)

SECTION - D

7. What is web data mining? What are the aims of web data mining? Explain the challenges faced while mining complex data types. (10)
8. Write a short note on:
- (a) Spatial data mining.
 - (b) Multimedia databases. (2×5=10)

SECTION - E (Compulsory Questions)

9. Attempt all the following:
- (a) How is clustering different from classification?
 - (b) Define confidence of an association rule.
 - (c) Discuss K-mean with respect to clustering.
 - (d) Define Data cube. Give an example.
 - (e) What is Decision tree? List out the major steps used in decision tree classification. (5×4=20)

2. Explain the three-tier data warehouse architecture in detail. 10

Section B

3. Examine the steps involved for the design and construction of data warehouses. 10
4. List down the areas in which data warehouses are used in present and in future ? Justify your answer with suitable example. 10

Section C

5. Describe the steps involved in Knowledge Discovery in Databases (KDD). 10
6. Explain the algorithm for constructing a decision tree from training samples. 10

Section D

7. Discuss the method of mining spatial databases in detail. 10
8. Explain time series analysis with suitable example. 10

(Compulsory Question)

9. (a) What is meant by outlier ? How is these outliers detected using data mining ?
- (b) Explain various methods for data cleaning with suitable example.
- (c) Explain various strategies for data reduction with suitable example.
- (d) Define strong association rule with suitable example. 4×5=20

Feb-22

[Total No. of Questions - 9] [Total No. of Printed Pages - 2]

Dec.-22-0280

CS-605 (Data Mining and Data Warehousing)

B.Tech. 6th (CBCS)

Time : 3 Hours

Max. Marks : 60

The candidates shall limit their answers precisely within the answer-book (40 pages) issued to them and no supplementary/continuation sheet will be issued.

Note : Attempt five questions in all, selecting one question from each section A, B, C and D. Question no. 9 is compulsory.

SECTION - A

1. How is data warehouse different from a database? How are they similar? Describe the features of a data warehouse. What are ROLAP and MOLAP? Describe two approaches and list their advantages and disadvantages. (10)
2. Discuss the characteristics of OLAP. List the major differences between OLTP systems and OLAP systems. What are the benefits of implementing a data warehouse? What are the risks? (10)

SECTION - B

3. Describe how a data warehouse is modelled and implemented using star schema and the snowflake schema. Explain using an example. Are there any other modelling techniques for data warehouse? (10)
4. Explain various tools required to create and manage a data warehouse. What is the purpose of a data cube? Define the terminology used in it? Use an example to illustrate the use of a data cube. (10)

SECTION - C

5. Briefly explain data mining and define it. What is association rule mining? Explain why association rules cannot be used

2

CS-605

directly for prediction, without further analysis or domain knowledge. Explain the different types of cluster analysis methods and discuss their features. (10)

6. Explain the various steps followed in KDD (Knowledge Discovery in Data) Process in detail. Give real life applications explaining each step. Elaborate the various attributes of data and the proximity measures to find the similarity amongst data attributes. (10)

SECTION - D

7. Explain various Data Mining Techniques applicable for the following complex data types:
 - (i) Time Series Data
 - (ii) Sequence Data
 - (iii) Web Data
 - (iv) Textual Data(10)
8. What is web data mining? In what situations can web mining data techniques be useful? What are the aims of web data mining? Explain the challenges faced while mining complex datatypes. (10)

SECTION - E

9. (a) Define support and confidence for an association rule.
(b) What is the need of stemming in text mining?
(c) Differentiate Classification and Clustering.
(d) Briefly outline the major steps of decision tree classification.
(e) Explain the methodology for knowledge discovery through neural networks. (5×4=20)

Roll No.

Total Pages : 03

July-22-00359

B. Tech. EXAMINATION, 2022

Semester VI (CBCS)

DATA MINING AND DATA WAREHOUSING

CS-605

Time : 3 Hours

Maximum Marks : 60

The candidates shall limit their answers precisely within the answer-book (40 pages) issued to them and no supplementary/continuation sheet will be issued.

Note : Attempt *Five* questions in all, selecting *one* question from each Section A, B, C and D. Q. No. 9 is compulsory.

Section A

1. Compare OLTP and OLAP Systems on the bases of advantages and disadvantages. 10

P.T.O.

- (b) Discuss about SDD with suitable example and differentiate between S-attribute and L-attribute. (5)

SECTION - D

7. (a) Construct the DAG for the following basic block

$$d = b * c$$

$$e = a + b$$

$$b = b * c$$

$$a = e - d$$

and simplify the above three address code of assuming:

(i) Only **a** is live on the exit from the block.

(ii) **a**, **b**, and **c** are live on exit from the block. (5)

- (b) Discuss the partitioning of three-address instructions into basic blocks and representation by flow graph using suitable example. (5)

8. (a) Translate the following three-address statements into machine-code instructions and show the register and address descriptors before and after the translation of each three-address instruction.

$$t = a - b$$

$$u = a - c$$

$$v = t + u$$

$$a = d$$

$$d = u + v \quad (5)$$

- (b) Construct the DAG following expressions.

$$((x + y) - ((x + y) * (x - y))) + ((x + y) * (x - y)) \quad (5)$$

SECTION - E

9. (a) Discuss in brief about any compiler writing tool.
- (b) Compare DFA and NDFA with suitable example.
- (c) What do you mean by bootstrapping? Discuss with example.
- (d) Explain backtracking with suitable example.
- (e) Describe the context free grammar and parse tree.
- (f) Construct the syntax tree for the expression:
 $a + a * (b - c)$.
- (g) Compare the quadruples and triples representation of three-address code.
- (h) Translate the arithmetic expression $a = b * -c + b * -c$ into three-address code where $-c$ is a unary operator.
- (i) What are the primary tasks of code generator in compiler design?
- (j) What is the Peephole Optimization? (10×2=20)

[Total No. of Questions - 9] [Total No. of Printed Pages - 2]

Dec-24-0467 (CBCS)

CS-605 (Data Mining & Data Warehousing)

B.Tech-6th (CBCS)

Time : 3 Hours

Max. Marks : 60

The candidates shall limit their answers precisely within the answer-book (40 pages) issued to them and no supplementary/continuation sheet will be issued.

Note: Attempt five questions in all, select one question from each section A, B, C, D. Section E (Question-9) is compulsory.

SECTION-A

1. Define Data Warehouse. What are the basic analytical operations of OLAP? Apply these OLAP operations for Trend Analysis. (10)
2. Draw a 3-tier data warehouse architecture and briefly define the working of each tier. (10)

SECTION-B

3. A) Explain the different features of data in data warehousing. (5)
B) What do you mean by distributed and virtual data warehouse? (5)
4. A) Explain the process of data cleaning and transportation. (5)
B) What is dimensional modelling? Write design methodology for the data warehouse. (5)

SECTION- C

5. A) What is frequent pattern mining? Also gives its applications. (5)

2

CS-605

- B) Draw the labelled architecture of data mining and explain it. (5)
6. A) What is the process of integrating data mining and data warehousing? Explain. (5)
B) Give DBMS vs Data Mining. (5)

SECTION-D

7. Write a note on i) Text mining ii) Web mining. (10)
8. Explain time series analysis and spatial data mining. (10)

Section-E (Compulsory)

9. i What do you mean by data granularity?
ii Define data cleaning.
iii What is the need to handle missing values in data mining?
iv What is sequential rule mining?
v Explain the need for creating a virtual data warehouse?
vi Define ETL. What is the significance of ETL tool in data warehouse?
vii Define frequent sets, confidence, and support.
viii Give the importance of Association Rule Mining.
ix "Classification is supervised learning". Justify.
x Explain challenges in data mining. (10×2=20)