

2017

(5th Semester)

BACHELOR OF COMPUTER APPLICATION

Paper No : BCA-5E3

(Data Mining and Warehousing)

Full Marks : 75

Time : 3 hours

(PART : B—DESCRIPTIVE)

(Marks : 50)

*The figures in the margin indicate full marks
for the questions*

1. (a) Explain the steps of knowledge discovery in databases. 6
- (b) Discuss about the different issues of data mining. 4
- Or
- (c) List and discuss the classification of data mining systems. 10

2. (a) Explain web content mining techniques. 10

Or

- (b) What are the characteristics of web data? 4
- (c) Discuss the steps that involved in web log data preprocessing. 6

3. (a) Illustrate the structure and functions of data warehouse. 10

Or

- (b) Explain the host-based data warehouse. 10

4. (a) What is a data warehouse? Discuss the data warehouse architecture with diagram. 2+8=10

Or

- (b) What is metadata? Describe the three categories of metadata. 2+8=10

5. (a) What is STAR schema? What are component tables? 2+8=10

Or

- (b) Describe the composition of the primary keys for the dimension and fact tables. 10

★★★

Subject Code : V/BCA/5E3

Booklet No. A

Date Stamp

.....

To be filled in by the Candidate

DEGREE 5th Semester
(Arts / Science / Commerce /
.....) Exam., **2017**

Subject

Paper

INSTRUCTIONS TO CANDIDATES

- 1. The Booklet No. of this script should be quoted in the answer script meant for descriptive type questions and vice versa.**
- 2. This paper should be ANSWERED FIRST and submitted within 1 (one) Hour of the commencement of the Examination.**
- 3. While answering the questions of this booklet, any cutting, erasing, over-writing or furnishing more than one answer is prohibited. Any rough work, if required, should be done only on the main Answer Book. Instructions given in each question should be followed for answering that question only.**

To be filled in by the Candidate

DEGREE 5th Semester
(Arts / Science / Commerce /
.....) Exam., **2017**

Roll No.

Regn. No.

Subject

Paper

Descriptive Type

Booklet No. B

*Signature of
Scrutiniser(s)*

*Signature of
Examiner(s)*

*Signature of
Invigilator(s)*

/278

V/BCA/5E3

2 0 1 7

(5th Semester)

BACHELOR OF COMPUTER APPLICATION

Paper No : BCA-5E3

(Data Mining and Warehousing)

(PART : A—OBJECTIVE)

(Marks : 25)

The figures in the margin indicate full marks for the questions

SECTION—I

(Marks : 15)

I. Put a Tick (✓) mark against the correct answer in the brackets provided : 1×10=10

1. A goal of data mining includes which of the following?

- (a) To explain some observed events or conditions ()
- (b) To confirm that data exists ()
- (c) To analyze data for expected relationships ()
- (d) To create a new data warehouse ()

(2)

2. ____ is described as extracting the information from large amount of data in a database.

(a) Data modeling ()

(b) Data warehousing ()

(c) Data mining ()

(d) None of the above ()

3. Which metadata links business and warehouse terminology and describes the data's technical structure?

(a) Usage metadata ()

(b) Control metadata ()

(c) Operational metadata ()

(d) Build time metadata ()

4. Which of the following is used to examine data collected by search engines and web spiders?

(a) Web structure mining ()

(b) Web content mining ()

(c) Web usage mining ()

(d) None of the above ()

V/BCA/5E3/278

(3)

5. Which of the following is the process of detecting and correcting the wrong data?
- (a) Data selection ()
 - (b) Data cleaning ()
 - (c) Data integration ()
 - (d) None of the above ()
6. The information gathered through web mining is evaluated by using
- (a) clustering ()
 - (b) classification ()
 - (c) association ()
 - (d) All of the above ()
7. Which is an application that modifies the data whenever it retrieved and has large number of simultaneous users?
- (a) OLAP ()
 - (b) OLTP ()
 - (c) ODS ()
 - (d) Fact table ()

V/BCA/5E3/278

(4)

8. High speed, single pass, parallelizable, multitable join is
- (a) snowflake ()
 - (b) star index ()
 - (c) star join ()
 - (d) star schema ()
9. Which data design includes determination of the various data elements that are needed and combination of the data elements into structures of data?
- (a) Logical ()
 - (b) Physical ()
 - (c) Conceptual ()
 - (d) Operational ()
10. ____ is the process of using graph theory to analyze the node and connection structure of a web site.
- (a) Web structure mining ()
 - (b) Web content mining ()
 - (c) Web usage mining ()
 - (d) None of the above ()

V/BCA/5E3/278

(5)

II. State whether the following statements are *True (T)* or *False (F)* by putting a Tick (✓) mark in the brackets provided : 1×5=5

(a) Foreign key is nothing but a substitute for the natural primary key.

(T / F)

(b) Data dictionary contains the information about the project information, graph and server information.

(T / F)

(c) PageRank or PR(A) can be calculated using a simple iterative algorithm and corresponds to the principal eigenvector of the normalized link matrix of the web.

(T / F)

(d) Association is the process of making a group of abstract objects into classes of similar objects.

(T / F)

(e) Data warehouse is the actual discovery phase of a knowledge discovery process.

(T / F)

V/BCA/5E3/278

(6)

SECTION—II

(Marks : 10)

III. Answer the following questions :

2×5=10

1. Define data mining.

V/BCA/5E3/**278**

(7)

2. What is rule-based filtering?

V/BCA/5E3/**278**

(8)

3. What are the steps to build the data warehouse?

4. What are the types of data modeling?

V/BCA/5E3/**278**

(9)

5. List out the functionalities of metadata.

★ ★ ★

8G—70/**278**

V/BCA/5E3