Professional Course Examination, November 2018

(5th Semester)

BACHELOR OF COMPUTER APPLICATIONS

Course : BCA 5E3

(Data Mining and Warehousing)

(Revised)

Full Marks : 75

Time : 3 hours

(PART : A—OBJECTIVE)

(*Marks* : 25)

The figures in the margin indicate full marks for the questions

SECTION—A

(*Marks* : 15)

Tick (✓) the correct answer in the brackets provided : 1×10=10

1. The synonym for data mining is

(a) data warehouse ( )

(b) knowledge discovery in database ( )

(c) extract, transform and load (ETL) ( )

(d) business intelligence ( )

2. Which of the following is ***not*** a function of data warehouses?

(a) Characterization and discrimination ( )

(b) Classification and regression ( )

(c) Selection and interpretation ( )

(d) Clustering and analysis ( )

3. An operational system is which of the following?

(a) The business in real time and is based on historical data ( )

(b) The business in real time and is based on current data ( )

(c) Support decision-making and is based on current data ( )

(d) Support decision-making and is based on historical data ( )

4. _____ is a good alternative to the star schema.

(a) Fact constellation ( )

(b) Star schema ( )

(c) Snowflake schema ( )

(d) Star-snowflake schema ( )

5. A data warehouse is which of the following?

(a) It can be updated by end users ( )

(b) It contains numerous naming conventions and formats ( )

(c) It can be organized around important subject areas ( )

(d) It contains only current data ( )

6. Which of the following technologies is not well-suited for data mining?

(a) Expert system technology ( )

(b) Data visualization ( )

(c) Parallel architecture ( )

(d) Technology limited to specific data types such as numeric data types ( )

7. The most common kind of queries in a data warehouse are

(a) inside-out queries ( )

(b) outside-in queries ( )

(c) browse queries ( )

(d) range queries ( )

8. Which of the following features usually applies to data in a data warehouse?

(a) Data are often deleted ( )

(b) Data are rarely deleted ( )

(c) Most applications consist of transactions ( )

(d) Relatively few records are processed by applications ( )

9. Concept description is the basic form of the

(a) predictive data mining ( )

(b) descriptive data mining ( )

(c) data warehouse ( )

(d) relational database ( )

10. Web mining techniques do not include

(a) cookies ( )

(b) clustering ( )

(c) user identification ( )

(d) Web content personalization ( )

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

1. Data warehouse architecture is based on RDBMS. *(T / F)*

2. The type of relationship in star schema is many to many. *(T / F)*

3. A goal of data mining is to explain some observed event or condition. *(T / F)*

4. Transaction processing is not a kind of data warehouse application. *(T / F)*

5. Stationary database is not a type of data warehouses. *(T / F)*

SECTION—B

(*Marks* : 10)

Answer the following questions : 2×5=10

1. What is knowledge discovery in database?

2. Why is it important for decision support?

3. What are the two major issues in data mining?

4. What are two applications of data mining?

5. Define single-stage data warehousing.

(PART : B—DESCRIPTIVE)

(*Marks* : 50)

The figures in the margin indicate full marks for the questions

Answer **all** questions

1. *(a)* What is data warehouse? Explain the design and construction of data warehouse in detail. 2+8=10

OR

(b) What is data mining? Explain the architecture of data mining system with suitable diagram. 2+8=10

2. *(a)* Explain the following data mining methodologies with suitable example : 5+5=10

(i) Decision trees

(ii) Classification rules

OR

(b) Differentiate between the following terms : 5+5=10

(i) Operational system and Information system

(ii) Statistical models and Linear models

3. *(a)* What is dimensional modelling? Explain dimension tables and aggregate fact tables with suitable demonstration. 2+8=10

OR

(b) What is schema in the context of dimensional modelling? Explain the concepts of STAR schema with suitable demonstration. 2+8=10

4. *(a)* Explain Web content mining techniques. 10

OR

(b) Mention the characteristics of Web data. 10

5. Mention the following types of data warehouses : 5+5=10

(a) Host-based data warehouses

(b) LAN-based data warehouses

OR

(c) Virtual data warehouses

(d) Distributed warehouses

★ ★ ★