

(6 pages)

Reg. No. :

**Code No. : 30176 E Sub. Code : GMCS 63/
GMSE 63**

B.Sc. (CBCS) DEGREE EXAMINATION,
APRIL 2020.

Sixth Semester

Computer Science/Software Engineering — Main

DATA MINING

(For those who joined in July 2012 – 2015)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer :

1. Data mining is
 - (a) The actual discovery phase of a knowledge discovery process
 - (b) The state of selecting right data for a KDD process
 - (c) A subject oriented integrated time variant non volatile collection of data in support of management
 - (d) None of the above

2. Which stage of data mining involves preparation and collection of data?
 - (a) Validation
 - (b) Exploration
 - (c) Both (a) and (b)
 - (d) Collection

3. Which algorithm is used to find correlations among different attributes in a data set?
 - (a) Associative algorithm
 - (b) Association algorithm
 - (c) Time series algorithm
 - (d) Series algorithm

4. Which of the following is used for frequent item set mining and association rule learning over relational database?
 - (a) Fp-growth
 - (b) AIS
 - (c) Apriori
 - (d) SETM

5. Classification is
 - (a) A subdivision of a set of examples into no. of classes
 - (b) A measure of accuracy
 - (c) The task of assigning a classification to set of examples
 - (d) None of the above

6. A decision tree is a tree in which every node is either a _____ or a decision node.
- (a) Leaf node
 - (b) Root node
 - (c) Both (a) and (b)
 - (d) Sub node
7. Which method of analysis does not classify variables as dependant or independent?
- (a) Regression analysis
 - (b) Discriminate analysis
 - (c) Analysis of variance
 - (d) Cluster analysis
8. Cluster is
- (a) Group of similar objects that differ significantly from other objects
 - (b) Simply data in order prepare machine learning algorithm
 - (c) Symbolic representation of facts
 - (d) None of the above
9. A _____ technique is used to crawl through various web sources to collect required information, which enables a individual or company to promote business.
- (a) Text data mining
 - (b) Web data mining
 - (c) Spatial data mining
 - (d) Time series data mining

10. Web mining is the application of
- (a) Data mining
 - (b) Text mining
 - (c) Both (a) and (b)
 - (d) None of these

PART B — (5 × 5 = 25 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 250 words.

11. (a) What is data mining and Why data mining now?

Or

- (b) Give some applications of data mining.

12. (a) Write a note on FP- Tree Algorithm.

Or

- (b) Explain Apriori algorithm.

13. (a) What is classification and Define decision tree?

Or

- (b) What is Naive Bayes algorithm?

14. (a) What is cluster analysis and what are the types of cluster analysis methods?

Or

- (b) Elucidate the idea behind logistic regression.

15. (a) Define web mining and list out the types of web mining.

Or

- (b) Explain web content mining.

PART C — (5 × 8 = 40 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 600 words.

16. (a) Guidelines for successful data mining.

Or

- (b) List and explain Data mining Applications.

17. (a) Explain Association rule.

Or

- (b) Discuss and detail about OLAP.

18. (a) A novel association rule mining approach using TID.

Or

(b) Decision tree and Naivebayes classifier.

19. (a) Divide and conquer approach and parallel clustering approach to deal large database.

Or

(b) Explain the K-means method.

20. (a) Explain Web Usage Mining.

Or

(b) Explain Web Technology and its Characteristics in Web Mining.
