

(6 pages)

Reg. No. :

Code No. : 41304 E Sub. Code : JMCA 5 C

B.C.A. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2018.

Fifth Semester

Computer Application – Main

Elective II — DATA MINING

(For those who joined in July 2016 onwards)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer.

1. What is the principle of sorting through large amounts of data and picking out relevant Information?
 - (a) databases
 - (b) knowledge discovery
 - (c) data mining
 - (d) extraction

2. Which type of tasks characterize the general properties of the database?
 - (a) descriptive mining
 - (b) predictive mining
 - (c) data mining
 - (d) knowledge mining
3. Which one appear frequently in a dataset?
 - (a) itemset
 - (b) correlations
 - (c) frequent patterns
 - (d) data
4. Which algorithm is a seminal algorithm for mining frequent itemsets for Boolean Association rules?
 - (a) Correlation
 - (b) Apriori
 - (c) K-means
 - (d) Association rule
5. What is a form of data analysis that extracts models describing important databases?
 - (a) classification
 - (b) correlation
 - (c) association
 - (d) pattern

Page 2 Code No. : 41304 E



6. Which classifier uses a set of IF-THEN rules in classification?
 - (a) Decision tree
 - (b) Bayes method
 - (c) Naive method
 - (d) Rule-based
7. What is the process of partitioning a set of data objects into subsets?
 - (a) cluster analysis
 - (b) classification
 - (c) association
 - (d) rule-mining
8. Which method creates a hierarchical decomposition of the given set of data objects?
 - (a) Association
 - (b) Clustering
 - (c) Hierarchical method
 - (d) Pattern
9. Which mining contains a rich and dynamic collection of Information about web page contents with hypertext structures and multimedia?
 - (a) Text Data
 - (b) Web Data
 - (c) Structure Data
 - (d) Global Data
10. Which mining is the process of extracting useful Information from server logs?
 - (a) web data
 - (b) text data
 - (c) content
 - (d) web usage

Page 3 Code No. : 41304 E

PART B — (5 × 5 = 25 marks)

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

Each answer should not exceed 250 words.

11. (a) What are all the applications of Data mining? Explain it.
 Or
 (b) Write short note on Data Mining software.
12. (a) Write short note on the basics of Association Rule Mining.
 Or
 (b) Explain briefly on the performance evaluation of Algorithms.
13. (a) Write short note on Decision Tree rules.
 Or
 (b) Write short note on classification methods.
14. (a) Explain the types of data.
 Or
 (b) Write short note on partitional methods.

Page 4 Code No. : 41304 E

[P.T.O.]



15. (a) Write short note on Web Technology and its characteristics.

Or

- (b) Write short note on Web usage mining.

PART C — ($5 \times 8 = 40$ marks)

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

Each answer should not exceed 600 words.

16. (a) Describe the process of Data Mining.

Or

- (b) Explain in detail about Data Mining techniques.

17. (a) Discuss elaborately about the Apriori Algorithm.

Or

- (b) Explain Mining frequent pattern without candidate generation.

18. (a) Explain split Algorithm based on the Information theory.

Or

- (b) Explain Evaluation criteria for classification methods.

Page 5 Code No. : 41304 E

19. (a) Describe desired features of cluster analysis.

Or

- (b) How to deal with large databases? Explain the process.

20. (a) Explain Web content mining.

Or

- (b) Explain Web structure mining.

Page 6 Code No. : 41304 E

