

(6 pages)

Reg. No. :

**Code No. : 30386 E Sub. Code : JMCS 64/
JMSE 64**

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

Sixth Semester

Computer Science/Software Engineering – Main

DATA MINING

(For those who joined in July 2016 only)

Time : Three hours Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL the questions.

Choose the correct answer :

1. The enterprise decision makers need to formulate goals that the data mining process is expected to achieve is _____.
 - (a) Requirement analysis
 - (b) Data selection and collection
 - (c) Data validation
 - (d) Data mining exploration

2. Decision tree is _____ technique.
- (a) Association rules mining
 - (b) Cluster analysis
 - (c) Supervised classification
 - (d) Search engine
3. Support (X) = ?
- (a) Number of times X appears
 - (b) Total number of transactions (N)
 - (c) Number of times X appears/N
 - (d) N+X
4. Given the association rules $X \rightarrow YZ$ and $AB \rightarrow C$, which one of following is not true
- (a) $X \rightarrow Y$ (b) $A \rightarrow C$
 - (c) $B \rightarrow AC$ (d) $X \rightarrow Z$
5. If the classes are created without looking at the data, the classification is called _____.
- (a) objects
 - (b) apriori classification
 - (c) posteriori classification
 - (d) none

6. Decision tree is a _____ model.
- (a) predictive (b) descriptive
(c) learning (d) both (a) and (b)
7. An ideal cluster analysis method should have _____.
- (a) scalability
(b) minimal input parameters
(c) ability to stop and resume
(d) all the above
8. In cluster analysis, Distance is always _____.
- (a) zero (b) positive
(c) negative (d) none
9. Which algorithms is used for web content mining?
- (a) DIPRE (b) Path traversal
(c) HITS (d) None
10. _____ deals with understanding user behaviour in interacting with the web or with a web site.
- (a) Web Content Mining
(b) Web Structure Mining
(c) Web Usage Mining
(d) Hyperlink

PART B — (5 × 5 = 25 marks)

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

Each answer should not exceed 250 words.

11. (a) Describe the steps in typical data mining process.

Or

- (b) Describe the guidelines for successful data mining.

12. (a) What is association rules mining? Explain the basics and give examples.

Or

- (b) Explain Naive algorithm with an example.

13. (a) What do you understand about Decision Tree? Explain Tree Induction Algorithm.

Or

- (b) Describe the evaluation criteria for classification methods.

14. (a) Describe the features of cluster analysis.

Or

- (b) Explain the types of cluster analysis methods.

15. (a) What is Web Mining? Describe the categories.

Or

(b) Explain about Web Document Clustering.

PART C — (5 × 8 = 40 marks)

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

Each answer should not exceed 600 words.

16. (a) What kind of tasks in data mining suitable for? Discuss.

Or

(b) What are the aims of Data Mining Techniques? Explain.

17. (a) Explain Apriori Algorithm with an example.

Or

(b) Explain about FP-Growth.

18. (a) Explain partitioned methods in cluster analysis.

Or

(b) Briefly explain Hierarchical methods.

19. (a) Explain Split Algorithm based on information theory.

Or

(b) Explain Naive Bayes method for classification.

20. (a) Describe the importance of web usage mining.

Or

(b) Write about web structure mining.
