

15. (a) Write about DTS TASKS.

Or

(b) Discuss about data mining applications using the schema rowsets.

PART C — (5 × 8 = 40 marks)

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

Each answer should not exceed 600 words.

16. (a) Briefly describe the basic data mining tasks in detail with examples.

Or

(b) Write about Neural Networks in data mining.

17. (a) Explain decision tree based algorithm.

Or

(b) Explain rule-based algorithms.

18. (a) Describe the parallel and distributed algorithms.

Or

(b) Describe any two advanced association rule techniques.

19. (a) Discuss the web content mining.

Or

(b) Explain the spatial data overview.

20. (a) Explain about DTS to create a data mining model with diagram.

Or

(b) How to create DTS packages workflow. Give an example.

Reg. No. :

Code No. : 7514

Sub. Code : PCAE 41

M.C.A. (CBCS) DEGREE EXAMINATION,
APRIL 2019.

Fourth Semester

Computer Application

Elective – DATA MINING.

(For those who joined in July 2017 onwards)

Time : Three hours Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer :

1. The techniques for adjusting weight is called _____ techniques.

- (a) Adjustment (b) Learning
(c) Propagation (d) Activation

2. A _____ identifies patterns or relationships in data.

- (a) Descriptive model
(b) Predictive model
(c) Classification model
(d) Regression model

3. _____ curve shows the relationship between false positives or true positives.

- (a) OC
(b) Stahshcal based curve
(c) Division
(d) Predichon



4. Distance-based algorithms use _____ measures to perform classification.
 (a) Similarity (b) Distance
 (c) Only (b) (d) Both (a) and (b)
5. Clustering algorithm alternatively assumes that the cluster is represented by one centrally located object in cluster called a _____.
 (a) Centric (b) Mediod
 (c) Outlier (d) All the above
6. A self-organizing feature map in a neural network approach that uses competitive _____ learning.
 (a) Unsupervised (b) Supervised
 (c) Selection (d) Evolutionary
7. In association rules, the _____ of an item is the percentage of transaction in which the item occurs.
 (a) Item set (b) Confidence
 (c) Relationship (d) Support
8. When computing distance in cluster analysis, distance is always _____.
 (a) Positive (b) Negative
 (c) Zero (d) positive or negative
9. _____ retrieve relevant document documents using a keyword based technique.
 (a) Search engines (b) Query
 (c) Data set (d) Cursor

10. _____ is based on the use of caching indexing and crawling.
 (a) Crawlers
 (b) Harvest system
 (c) Virtual web view
 (d) Personalization

PART B — (5 × 5 = 25 marks)

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

Each answer should not exceed 250 words.

11. (a) Discuss data mining versus knowledge discovery in databases.
 Or
 (b) Explain the Bayes theorem.
12. (a) Explain regression in statistical based algorithm.
 Or
 (b) What are the combining techniques in Classification?
13. (a) Explain the apriori algorithm with example.
 Or
 (b) What are used to measure the quality of rules?
14. (a) What is web structure mining?
 Or
 (b) Write about spatial rules and spatial association rules.

