

Reg. No. :

(7 pages)

Code No. : 6939

Sub. Code : HCAM 52

M.C.A. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2014.

Fifth Semester

Computer Application

DATA MINING AND DATA WAREHOUSING

(For those who joined in July 2012 onwards)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer :

1. A _____ is a reporting database that contains relatively recent as well as historical data and may also contains a aggregate data
- (a) Relational Database
 - (b) Data Warehouse
 - (c) Operational Data Stores
 - (d) OLTP.

2. _____ is a collection of techniques for efficient automated discovery of previously unknown, valid, novel and useful patterns in large databases

- (a) Data Warehousing
- (b) Data Mart
- (c) Data Mining
- (d) Data Preprocessing.

3. _____ combines the measures for the interestingness of the rules

- (a) Antecedent and Consequent
- (b) Support and Confidence
- (c) Market — basket analysis
- (d) Frequent and Infrequent.

4. _____ is the separation of ordering of objects into classes

- (a) Clustering
- (b) Classification
- (c) Supervised learning
- (d) Decision Tress.



5. _____ is the accuracy method that does not rely on a single test set

- (a) Holdout Method
- (b) K-fold Cross Validation
- (c) Random sub sampling Method
- (d) Bootstrap Method.

6. Weight, marks, height, price, salary these indicates the type of _____ data

- (a) Numerical
- (b) Binary
- (c) Qualitative
- (d) Ordinal

7. _____ is the application of data mining techniques to find interesting and potentially useful knowledge from web data

- (a) Web Mining
- (b) Web Content Mining
- (c) Web Usage Mining
- (d) Web Structure Mining.

Page 3

Code No. : 6939

8. _____ is an identifier for an abstract or physical resource

- (a) Browser
- (b) Uniform Resource Locator
- (c) Client
- (d) Proxy.

9. _____ is the branch of data mining that deals with location or geo- referenced data

- (a) Text Mining
- (b) Temporal Mining
- (c) Spatial Mining
- (d) Web Mining.

10. _____ is a description of the spatial and non-spatial properties, which are typical for the target objects

- (a) Spatial characterization
- (b) Spatial Trends
- (c) Spatial Clustering
- (d) Spatial Trend Detection.

Page 4

Code No. : 6939

[P.T.O.]



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 Metadata in the context of Data warehouse?

Or

- (b) List out the steps in Data Mining Process.

12. (a) Explain Direct Hashing and Pruning algorithm briefly.

Or

- (b) What are the steps included in building a decision tree with tree induction algorithm?

13. (a) How evaluation criteria are classified for classification methods?

Or

- (b) What are the methods used to compute distance in cluster analysis?

14. (a) Write brief notes on web mining and its categories.

Or

- (b) What is search engine? Explain its functionalities.

Page 5 Code No. : 6939

15. (a) How spatial clustering is performed? Illustrate.

Or

- (b) What are the tasks performed in spatial mining? Briefly explain.

PART C — (5 × 8 = 40 marks)

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

Each answer should not exceed 600 words.

16. (a) With a neat sketch explain Data warehouses architecture.

Or

- (b) What are the various Data Mining Techniques? Explain any one.

17. (a) How Apriori Algorithm is used to find Association rules? Explain.

Or

- (b) Provide a detailed study on Split Algorithm based on Information Theory.

18. (a) Explain Naïve Bayes classification method with example.

Or

- (b) Explain K-Means clustering method in detail.

Page 6 Code No. : 6939



19. (a) Discuss the characteristics of web mining process.

Or

- (b) Write short notes on Web Document Clustering.

20. (a) What are all the Spatial Data Mining Primitives? Explain.

Or

- (b) Explain Spatial Clustering Algorithm in detail.
-

