

**A**

Reg. No. :

--	--	--	--	--	--	--	--	--	--

**Question Paper Code: 55802**

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2018

Fifth Semester

Information Technology

15UIT502 - DATA WAREHOUSING AND DATA MINING

(Regulation 2015)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

1. A data warehouse is which of the following? CO1- R
  - (a) Can be updated by end users.
  - (b) Contains numerous naming conventions and formats
  - (c) Organized around important subject areas.
  - (d) Contains only current data
  
2. The data is stored, retrieved & updated in \_\_\_\_\_ CO1- R
  - (a) OLAP
  - (b) OLTP
  - (c) FTP
  - (d) SMTP
  
3. A goal of data mining includes which of the following? CO2- R
  - (a) To explain some observed event or condition
  - (b) To confirm that data exists
  - (c) To analyze data for expected relationships
  - (d) To create a new data warehouse
  
4. The various aspects of data mining methodologies is/are \_\_\_\_\_ CO2- R
  - (i) Mining various and new kinds of knowledge
  - (ii) Mining knowledge in multidimensional space
  - (iii) Pattern evaluation and pattern or constraint-guided mining.
  - (iv) Handling uncertainty, noise, or incompleteness of data
  - (a) i, ii and iv only
  - (b) ii, iii and iv only
  - (c) i, ii and iii only
  - (d) All i, ii, iii and iv

5. Support Vector Machines (SVMs, also support vector networks) are CO3- R  
 (a) Supervised learning (b) Unsupervised learning  
 (c) Both a and b (d) None of the above
6. A decision tree is a decision support tool that uses CO3- R  
 (a) Tree like graph (b) Linked representation  
 (c) Binary tree (d) None of the above
7. An Outlier is a CO4- R  
 (a) Rare chance of occurrence within a given data set  
 (b) Observation point that is distant from other observations.  
 (c) Both a and b  
 (d) None of the above
8. \_\_\_\_\_ is a method of incremental conceptual clustering. CO4- R  
 (a) COBWEB (b) OLAP (c) CORBA (d) STING
9. Temporal data mining refers to extract \_\_\_\_\_ information CO5- R  
 from large collections of temporal data  
 (a) Useful abstract (b) Non-trivial (c) Implicit (d) All the above
10. What are the different types of web mining? CO5- R  
 (a) Web Content Mining (b) Web Usage Mining  
 (c) Web Structure Mining (d) All the above

PART – B (5 x 2= 10Marks)

11. What is data warehouse metadata? CO1- R
12. What is Data preprocessing? CO2- R
13. What is the objective function of the K-means algorithm? CO3- R
14. List the requirements of clustering in data mining. CO4- R
15. What is meant by Spatial mining? CO5- R

PART – C (5 x 16= 80 Marks)

16. (a) What is the significance of OLAP in data warehouse? Describe OLAP operations for multidimensional data with necessary diagram/example. CO1- U (16)
- Or
- (b) With a neat sketch, Explain three tier architecture and different models of data ware housing. CO1- U (16)
17. (a) Elucidate various Classification of Data Mining Systems and its functionality in detail. CO2 -U (16)
- Or
- (b) Analyze the various tasks and issues involved to Integrate a data Mining system with a Data Warehouse. CO2- Ana (16)
18. (a) Explain decision tree induction algorithm for classifying data tuples and apply it on a suitable example. CO3- App (16)
- Or
- (b) Elaborate in detail about to mine closed frequent data item sets and Constraint Based Association Mining. CO3- Ana (16)
19. (a) Explain K-means clustering algorithm and Write the key issue in hierarchical clustering algorithm. CO4- U (16)
- Or
- (b) What is outlier detection? Explain distance based outlier detection in detail. Apply it on a suitable example. CO4- App (16)
20. (a) Explain the need and applications of Web mining and Web content mining in detail. CO5- U (16)
- Or
- (b) Write short notes on the following with its relevant applications CO5- U (16)
- (i) Spatial clustering algorithm
  - (ii) Temporal Mining

