

C

Reg. No. :

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

**Question Paper Code: 99209**

B.E./B.Tech. DEGREE EXAMINATION, MAY 2022

Elective

Computer Science and Engineering

19UCS909- Data Mining

(Régulations 2019)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (5 x 1 = 5 Marks)

- Strategic value of data mining is \_\_\_\_\_. CO1- U  
(a) Cost-sensitive (b) Work-sensitive  
(c) Time- sensitive (d) Technical- sensitive
- If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam, 10000 transaction contain both bread and jam. Then the support of bread and jam is \_\_\_\_\_. CO2- App  
(a) 2%. (b) 20% (c) 3%. (d) 30%.
- Which of the following criteria is not used to decide which attribute to split next in a decision tree: CO1- U  
(a) Gini index (b) Information gain (c) Entropy (d) Scatter
- Which is needed by K-means clustering? CO1- U  
(a) defined distance metric (b) number of clusters  
(c) initial guess as to cluster centroids (d) all of the above
- Data mining can be used to improve \_\_\_\_\_. CO1- U  
(a) Efficiency (b) Quality of data (c) Marketing (d) All of the above

PART – B (5 x 3= 15 Marks)

- State the various issues in data mining? CO1- U
- What is meant by constraint based mining? CO1- U
- What is rule based classification? How the rule is assessed? CO1- U
- State the various requirements of clustering CO1- U

10. What is meant by text mining? CO1- U
- PART – C (5 x 16= 80 Marks)
11. (a) Draw the architecture and describe the steps involved in data mining when viewed as a process of knowledge discovery. CO1- U (16)  
 Or  
 (b) Explain the process of data cleaning. CO2- U (16)
12. (a) Explain various kinds of Association Rules Mining CO1- U (16)  
 Or  
 (b) Describe efficient and scalable frequent item set using candidate generation Mining Methods. CO1- U (16)
13. (a) Explain with an example how classification is done using decision tree induction method. CO1- U (16)  
 Or  
 (b) Illustrate with an example how classification is done using Bayesian classification method. CO1- U (16)
14. (a) Explain different types of data used in cluster analysis & Explain the K-Means partitioning method CO1- U (16)  
 Or  
 (b) Explain the various methods for detecting outliers CO1- U (16)
15. (a) Discuss about Text Mining CO1- U (16)  
 Or  
 (b) Explain how data mining is used in health care analysis CO1- U (16)