

Reg. No. :

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

Question Paper Code: 59204

B.E./B.Tech. DEGREE EXAMINATION, SEP 2020

Elective

Computer Science and Engineering

15UCS904- DATA MINING

(Regulation 2015)

Duration: One hour

Maximum: 30 Marks

PART A - (6 x 1 = 6 Marks)

(Answer any Six of the following Questions)

1. Which of the following is not a data mining metric CO1- U
(a) Space complexity (b) Time complexity (c) ROI (d) All of the above
2. Strategic value of data mining is _____. CO1- U
(a) Cost-sensitive (b) Work-sensitive
(c) Time- sensitive (d) Technical- sensitive
3. Principal component analysis is a technique for performing? CO2- U
(a) Sampling (b) Discretization
(c) Dimensionality reduction (d) Aggregation
4. Online transaction processing is used because _____. CO2- U
(a) It is efficient (b) Disk is used for storing files
(c) It can handle random queries (d) Transactions occur in batches
5. If a store has N items, the number of possible item sets is CO3- R
(a) $2^N - 1$ (b) $N - 1$ (c) $N/2$ (d) None of the above
6. The absolute number of transactions supporting X in T is called _____. CO3- R
(a) Confidence (b) Support (c) Support count (d) None of the above

7. 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 _____.
 (a) 2% (b) 20% (c) 3% (d) 30% CO3- R
8. For support vectors X_j in a hard margin SVM, we have
 (a) $|W^T X_i + b| = 0$ (b) $|W^T X_i + b| < 1$ (c) $|W^T X_i + b| = 1$ (d) None of the above CO4- R
9. If the hyper plane $W^T X + b = 0$ separates all the training points (X_i, Y_i) , where $y_i = \{+1, -1\}$, then:
 (a) $\|W\| = 0$ (b) $b = 0$ (c) $W^T X_i + b \geq 0 \forall i$ (d) $y_i(W^T X_i + b) \geq 0 \forall i$ CO4- R
10. In a multiclass classification problem, Bayes classifier assigns an instance to the class corresponding to
 (a) Highest aposteriori probability (b) Highest apriori probability
 (c) Lowest aposteriori probability (d) Lowest apriori probability CO4- R

PART – B (3 x 8 = 24 Marks)

(Answer any Three of the following Questions)

11. Explain data mining task primitives and functionalities. CO1- U (8)
12. Explain the various systems on which data mining can be done. CO1- U (8)
13. Explain the process of data cleaning. CO2- U (8)
14. Explain various kinds of Association Rules Mining. CO3- U (8)
15. Briefly outline the major steps of decision tree classification CO4- U (8)