

Reg. No.																				
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code	12852
---------------------	-------

B.E. / B.Tech. - DEGREE EXAMINATIONS, APRIL / MAY 2024

Third Semester

Artificial Intelligence and Data Science

20AIPC302 - FUNDAMENTALS OF MACHINE LEARNING TECHNIQUES

Regulations - 2020

Duration: 3 Hours

Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

	Marks	K- Level	CO
1. What are the types of machine learning?	2	K1	CO1
2. List out any 5 real-time applications of Machine Learning.	2	K1	CO1
3. How to handle Missing or Corrupted Data in a Dataset?	2	K1	CO2
4. How classification varies from regression?	2	K1	CO2
5. State TP, FP, FN, TN for disease prediction into Benign & Malignant tumor assuming 'Benign' as a class of win.	2	K1	CO3
6. What is Data Pre-processing?	2	K1	CO3
7. How can the accuracy of Simple Linear regression be improved?	2	K2	CO4
8. List some of the common regression Algorithms.	2	K1	CO4
9. What are Clustering and its types?	2	K1	CO5
10. State the main difference in the approach of k-means and k-medoids algorithms.	2	K2	CO5

PART - B (5 × 13 = 65 Marks)

Answer ALL Questions

11. a) Explain in detail about the types of Machine learning with necessary diagrams. 13 K2 CO1

OR

b) How Machine Learning algorithms help in detecting fraudulent activities in Banking? Also state some of the real-time software used. 13 K2 CO1

12. a) i) How the Performance of the Regression is calculated? 7 K2 CO2

ii) Distinguish between Lazy Learner and Eager Learner. 6 K2 CO2

OR

b) How Numerical Data is Explored? Explain in detail. 13 K2 CO2

13. a) Consider the training dataset given in the following table. Use Weighted k-NN and determine the class. Test instance (7.6, 60, 8) and $K=3$. 13 K3 CO3

S.No.	CGPA	Assessment	Project Submitted	Result
1	9.2	85	8	Pass
2	8	80	7	Pass
3	8.5	81	8	Pass
4	6	45	5	Fail
5	6.5	50	4	Fail
6	8.2	72	7	Pass
7	5.8	38	5	Fail
8	8.9	91	9	Pass

OR

- b) Explain in detail about Support Vector Machines with algorithms and state the role of hyper planes. 13 K2 CO3
14. a) Define simple linear regression using a graph explaining slope and intercept also Explain rise, run, and slope in a graph. 13 K2 CO4
- OR**
- b) Elaborate the purpose of ridge and Lasso regression, its advantages and disadvantages. 13 K2 CO4
15. a) List out the broad three categories of clustering techniques? Explain the characteristics of each briefly. 13 K2 CO5
- OR**
- b) Explain about Partitioning Clustering Algorithm and its various types of partitioning algorithms. 13 K2 CO5

PART - C (1 × 15 = 15 Marks)

16. a) Apply Apriori Principle for Association rule learning with any Supermarket Dataset. Formulate the strong rules for the considered dataset. 15 K3 CO6
- OR**
- b) You are given a set of one-dimensional data points: {5, 10, 15, 20, 25, 30, 35}. Assume that $k = 2$ and first set of random centroid is selected as {15, 32} and then it is refined with {12, 30}. 15 K3 CO6
- (i) Create two clusters with each set of centroid mentioned above following the k-means approach.
- (ii) Calculate the SSE for each set of centroid.