Reg. No. | | | | | | | | | | | |

| Question Paper Code | 12345 |

**B.E. / B.Tech - DEGREE EXAMINATIONS, NOV / DEC 2023**
Third Semester
**Artificial Intelligence and Data Science**
(Common to Computer Science and Engineering (AIML))
**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 is Machine learning? What is the need of it? — *2,K1,CO1*
2. Define the steps involved in the basic machine learning process. — *2,K1,CO1*
3. What is a Confusion Matrix? — *2,K1,CO2*
4. Differentiate Model underfitting and Model overfitting. — *2,K2,CO2*
5. What are the strengths and weaknesses of KNN Algorithms? — *2,K1,CO3*
6. Give short notes on Entropy & Information Gain. — *2,K1,CO3*
7. Write a formula for Lasso regression. What is Variance/Bias TradeOFF? — *2,K1,CO4*
8. Define Polynomial Regression. — *2,K1,CO4*
9. What is K-Means Clustering? — *2,K1,CO5*
10. List some of the applications of Unsupervised Learning. — *2,K1,CO5*

**PART - B (5 × 13 = 65 Marks)**
Answer ALL Questions

11. a) Illustrate the types of machine learning with neat diagrams. — *13,K2,CO1*

**OR**

b) Explain in detail about the applications of machine learning in the Healthcare domain. — *13,K2,CO1*

12. a) Explain in detail about the basic types of data in machine learning. — *13,K2,CO2*

**OR**

b) Let's assume the confusion matrix of the win/loss prediction of cricket match problem to be as below: — *13,K3,CO2*

|  | Actual Win | Actual Loss |
|---|---|---|
| Predicted Win | 85 | 4 |
| Predictive Loss | 2 | 9 |

Show the parameters involved in calculating the performance of the model.

*K1 – Remember; K2 – Understand; K3 – Apply; K4 – Analyze; K5 – Evaluate; K6 – Create*                    **12345**

13. a) Discuss the random forest model in detail. What are the strengths and weaknesses of it?

*13,K2,CO3*

**OR**

b) Describe in detail about the support vector machines with algorithms.

*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) Discuss in detail about logistic regression and draw different scenarios for slopes.

*13,K2,CO4*

15. a) List out the broad three categories of clustering techniques? Explain with the characteristics of each in briefly.

*13,K2,CO5*

**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}.
  (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.

*13,K3,CO5*

## PART - C (1 × 15 = 15 Marks)

16. a) Illustrate how the bagging, gradient boosting works with ensemble learning.

*15,K3,CO6*

**OR**

b) Demonstrate in detail about the instance based Learning (Memory-based learning).

*15,K3,CO6*