# 2001E1224A

Candidate's Seat No :_____

**MCA Sem.-2 (Rep) Examination**

**Machine Learning**

**Time : 3-00 Hours]** **January-2025** **[Max. Marks : 50**

**Instructions:**

- **Write both the Sections in the separate answer book.**
- **Both Sections having equal weightage.**
- **Draw Diagrams wherever necessary.**
- **Make Assumptions wherever necessary.**

## SECTION – I

Q.1 A scientist is investigating whether birds of prey exposed to pollutants lay eggs with thinner shells. A random sample of eggs from six nests is chosen and examined for polluted level p and thinning t of the shell. The results obtained are presented in the following table: [09]

| p | 3 | 8 | 30 | 25 | 15 | 12 |
|---|---|---|---|---|---|---|
| t | 1 | 3 | 9 | 10 | 5 | 6 |

(a) Draw a scatter diagram.

(b) Find the regression line of t on p of the form $t = a+b\,p$.

(c) Plot (average p, average t) and regression line on the scatter diagram.

(d) The scientist concludes from similar other researches that pollutant level above 18 is likely to result in the death of a chick soon after hatching. Estimate minimum thickness of its shell that is likely to result in the death of a chick.

OR

Q.1 **Attempt the following:** [09]

(a) Explain the concept of linear regression model. Why is it not used for classification problems?

(b) Explain Gradient Descent Method for optimization.

Q.2 Explain Random Forest algorithm in detail. Explain in brief Weighted Averaging technique used in ensemble learning. [08]

OR

Q.2 Explain in brief types of Decision Tree and give appropriate examples. Write algorithm steps for creating Decision Tree. [08]

(P.T.O.)

E122AA

Q.3 **Explain the following terms with appropriate example:** [08]

(a) Explain Non-Linear SVM and how one can solve it. What are the pros and cons of SVM.

(b) What is Ensemble Learning? Explain advance ensemble learning techniques.

## SECTION – II

Q.4 **Answer the following question:** [09]

(a) Explain the following (in one/two sentences)

i. TP

ii. TN

iii. FP

iv. FN

v. Confusion Matrix

(b) What is distance measure? Explain following distance measures with the help of examples (Any Two):

i. Manhattan distance

ii. Edit Distances

iii. Cosine Distance

Q.5 **Differentiate between the following (Any TWO) :** [08]

(a) Divisive Hierarchical Clustering & Agglomerative Hierarchical Clustering

(b) Over fitting & Under fitting

(c) Bagging & Boosting

Q.6 **Answer the following question (Any TWO):** [08]

(a) What is Logistic Regression? When do we use Logistic Regression?

(b) What is Kernel Trick? Explain different types of kernels with examples.

(c) Explain the techniques used to find the optimal number of clusters in Kmeans.