CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
IV–B.TECH–I–Semester End Examinations (Regular) - December- 2025
PREDICTIVE ANALYTICS
(CSD)

[Time: 3 Hours]   **[Max. Marks: 60]**

Note: This question paper contains two parts A and B.
   Part A is compulsory which carries 10 marks. Answer all questions in Part A.
   Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART-A   (10 Marks)

1. a) Define ridge regression and mention when it is used.   [1M]
   b) Define logistic regression.   [1M]
   c) What is the bias-variance trade-off?   [1M]
   d) What is the purpose of bootstrapping?   [1M]
   e) What is gradient boosting?   [1M]
   f) Define AdaBoost.   [1M]
   g) Define the concept of backpropagation.   [1M]
   h) Define is K-nearest neighbor classifier.   [1M]
   i) Define unsupervised learning.   [1M]
   j) List the applications of association rule mining.   [1M]

PART-B   (50 Marks)

2. a) How does LDA differ from logistic regression for classification?   [5M]
   b) Analyze the effect of regularization parameters in Lasso and Ridge regression. [5M]

OR

3. a) What is Lasso regression and Ridge regression? Explain.   [5M]
   b) Derive the cost function of logistic regression.   [5M]

4.   Demonstrate K-fold cross-validation algorithm with an example.   [10M]

OR

5. a) Derive the relationship between bias, variance, and mean squared error.   [5M]
   b) How can effective number of parameters be used in model selection?   [5M]

6. a) Discuss bagging and boosting techniques.   [5M]
   b) Illustrate the use of additive models for nonlinear relationships.   [5M]

OR

7.   Compare additive models, regression trees, and boosting in predicting newzealand fish size data. [10M]

8.   Illustrate the working of KNN with a numerical example.   [10M]

OR

9. a) Compare SVM and neural networks for classification tasks.   [5M]
   b) Write the issues related to overfitting in neural networks.   [5M]

10. Discuss how random forests handle missing data and feature correlation with suitable examples. [10M]

OR

11. a) Write the hierarchical clustering and its linkage methods.   [5M]
    b) Compare random forests with decision trees.   [5M]

****