CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
III–B.TECH–I–Semester End Examinations (Supply)–June - 2025
INTELLIGENT DATABASE SYSTEMS
(CSD)

[Time: 3 Hours]   [Max. Marks: 70]

Note: This question paper contains two parts A and B.

    Part A is compulsory which carries 20 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.

PART-A   (20 Marks)

1. a)   What is Data Science?   [2M]
   b)   List the different fields of data science.   [2M]
   c)   Define Data Cleaning.   [2M]
   d)   What is Data Reduction?   [2M]
   e)   Define Pivot table.   [2M]
   f)   What is the use of Heat Map?   [2M]
   g)   What is the use of Pipeline?   [2M]
   h)   What is Visualization?   [2M]
   i)   What is the purpose of Grid Search?   [2M]
   j)   Define Ridge Regression.   [2M]

PART-B   (50 Marks)

2.   Discuss the measures required to ensure data security in a data science project.   [10M]

OR

3.   Describe the importance of each stage in the data science project lifecycle with an Example.   [10M]

4.   Evaluate different methods of data integration and their effectiveness in creating a unified dataset.   [10M]

OR

5.   Analyze the effects of discretizing continuous data on the performance of machine learning models.   [10M]

6. a)   What is the standard deviation measure in a data set?   [5M]
   b)   Compute the standard deviation for the data set [4, 8, 6, 5, 9].   [5M]

OR

7. a)   How does ANOVA help in comparing multiple groups?   [5M]
   b)   Perform a one-way ANOVA on the data sets Group A: [5, 7, 8], Group B: [6, 9, 11], Group C: [4, 8, 10].   [5M]

8. a)   Describe how a distribution plot helps in model evaluation.   [5M]
   b)   Explain how in-sample evaluation differs from out-of-sample evaluation.   [5M]

OR

9. a)   Analyze the strengths and weaknesses of using visualizations for model evaluation.   [5M]
   b)   Describe the role of regression models in decision making.   [5M]

10.   Discuss the advantages and disadvantages of using cross-validation compared to a simple train-test split.   [10M]

OR

11.   Compare the causes and effects of under fitting and over fitting.   [10M]

****