GANDHI INSTITUTE OF ENGINEERING AND TECHNOLOGY UNIVERSITY, ODISHA, GUNUPUR
(GIET UNIVERSITY)

B. Tech (Fourth Semester - Regular) Examinations, April – 2025

23BCDPE24001 - Predictive Modeling and Analytics

(CSE-DS)

Time: 3 hrs Maximum: 60 Marks

Answer ALL questions

(The figures in the right hand margin indicate marks)

PART – A **(2 x 5 = 10 Marks)**

Q.1. Answer ***ALL*** questions	CO #	Blooms Level
a. Identify the potential issues in converting raw data into a matrix.	CO1	K3
b. Explain the role of Support Vector Machines in classification.	CO2	K2
c. Write a short note on Data Cleaning.	CO3	K1
d. Explain the term “hidden grouping” in visualization.	CO4	K2
e. Define regression. Implement a regression model using R.	CO5	K1

PART – B **(10 x 5 = 50 Marks)**

Answer ***ALL*** the questions	Marks	CO #	Blooms Level
2. a. Differentiate between structured and unstructured data types.	5	CO1	K1
b. Discuss in brief Item-Based Collaborative Filtering with a neat diagram.	5	CO1	K2
(OR)			
c. Explain briefly Targeting using predictive modelling.	5	CO1	K2
d. Write a short note on Content and Text Analytics with examples.	5	CO1	K1
3.a. Given the following training dataset:			

Point	X1	X2	Class
A	2	4	Yes
B	4	2	No
C	4	4	Yes
D	6	4	No
E	6	2	No

Using Euclidean distance and K = 3, classify the test point (5, 3). (i) Compute the distance from the test point to each training point. (ii) Identify the 3 nearest neighbors. (iii) Determine the majority class among the neighbors. (iv) Assign the class to the test point.	10	CO2	K4
(OR)			
b. Cluster the following eight points (with (x, y) representing locations) into three clusters: A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)	10	CO2	K5
4.a. Explain briefly the ETL Process.	5	CO3	K2
b. How do you convince the Management to adopt Predictive Analytics in their Organizations?	5	CO3	K2
(OR)			

c.	Briefly discuss the various steps used in building a Predictive Model.	10	CO3	K2
5.a.	Analyze visualization techniques for decision trees and outlier detection.	5	CO4	K4
b.	Explain the importance of data visualization in predictive analytics. Why does visualization matter when interpreting model results and analytical findings? Illustrate your answer with relevant examples	5	CO4	K2
	(OR)			
c.	Describe how data classification results can be visualized effectively. What visualization techniques are commonly used to represent classification performance, and how do they aid in model evaluation?	5	CO4	K2
d.	Write a short note on Data Visualization tools.	5	CO4	K1
6.a.	How do you handle Missing Values in a data set? Give the code in R Language.	5	CO5	K1
b.	What are the key steps in building a predictive model in R?	5	CO6	K1
	(OR)			
c.	What are some major trends in predictive analytics? Explain briefly.	5	CO6	K1
d.	What are the main components of the RStudio interface?	5	CO5	K1

-- End of Paper ---