QP Code: SN22BTECH594 Reg. No | | | | | | | | | | | AR 21

Gandhi Institute of Engineering and Technology University, Odisha, Gunupur (GIET University)

B. Tech (Seventh Semester - Regular) Examinations, November – 2025

BCSPC6020– Data Analytics

(CSE)

Time: 3 hrs Maximum: 70 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.	What is bias-variance trade off?	CO1	K1
b.	Explain sigmoid activation function with its criteria.	CO2	K2
c.	How to evaluate a classification model? Explain with formulas.	CO2	K2
d.	Briefly describe decision boundary with its diagram.	CO3	K3
e.	Define support and confidence in association rule mining.	CO1	K1

PART – B **(15 x 4 = 60 Marks)**

Answer ***all*** the questions — Marks | CO # | Blooms Level

2. a. Construct a logistic regression model with predictors for the riding mower example with B_0 =-25.9382 and B_1=0.1109, B_2=0.9638, where B_1 and B_2 are for the variables X_1 and X_2 respectively. Use logistic regression model to classify the product will be purchased or not. — 8 | CO1 | K2

X_1	60	64	84	59	108	75
X_2	18.4	21.6	17.6	16.0	17.6	19.6
Purchased or not	YES	YES	NO	NO	YES	NO

b. Explain the stages of data analytics — 7 | CO3 | K1

(OR)

c. We want to fit a quadratic polynomial (degree 2) to the following data: — 10 | CO3 | K2

X: 1, 2, 3, 4, 5

Y: 2, 5, 6, 8, 10

The model will be: $\hat{y} = b_0 + b_1x + b_2x^2$

d. How can PCA be used for noise reduction in images or signals? — 5 | CO1 | K1

3.a. For the following data set apply KNN with k=3 and predict whether the person will buy the product or not. — 10 | CO3 | K2

Age:	25	63	33	42	50
Salary:	20	120	75	100	140
Purchased or not:	N	Y	N	Y	?

b. Evaluate whether Neural Networks or SVM would be better for image classification. — 5 | CO1 | K1

(OR)

c. Assume the data points are A(1, 1), B(1, 2), C(5, 5), D(6, 6), E(7, 7) , **K = 2**, Initial centroids: C1(2, 3), C2(9, 9). Implement k-means clustering technique to group the points. — 8 | CO3 | K3

d. Discuss the process of creating data for analytics through designed experiments. — 7 | CO2 | K1

4.a.	Given 2- class of points: Class 1: (1,2), (2,3), (3,3) Class 2: 6,8), (7,8), (8,9) Perform LDA to find the projection direction.	10	CO3	K3
b.	Explain the backpropagation algorithm in detail with mathematical derivations.	5	CO2	K1
	(OR)			
c.	To access the significant of possible variation in performance in a certain test between the convent schools of a city , a common test was given to a number of students taken at random from the 12th class of 3 schools. Concerned results are given below: A: 2,4,6 B: 3,5,7 C: 6,7,8 Make analysis of variance for the given data.	8	CO4	K2
d.	What is the role of activation functions in neural networks?	7	CO2	K1
5.a.	Consider the following data points given, find the optimal hyperplane for these data points. (1,0.5), (1,1),(1,-0.5), (-0.5,0.5), (0.5,0.5),(2,0),(4,0),(4.5,1),(4.5,5), (5,-1), (5.5,0)	8	CO4	K2
b.	Explain the challenges of Big-data.	7	CO2	K1
	(OR)			
c.	A car manufacturer claims their cars average 35 mpg. A sample of 16 cars gives: $\bar{x} = 33.2$ mpg, $s = 2.8$ mpg. Test at $\alpha = 0.05$ if mpg is less than claimed.	8	CO4	K3
d.	Create a complete classification pipeline using decision trees with cross-validation and performance metrics.	7	CO1	K1

--- End of Paper ---