

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



B. Tech (Sixth Semester - Regular) Examinations, April 2026
23BCSPC36002 – Introduction to Machine Learning
(CSE)

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

- a. Write the cost function of lasso and ridge regression
- b. Given cost function: $J(\theta) = \theta^2 + 4\theta + 4$
Find gradient and update θ using learning rate 0.2, initial $\theta = 2$.
- c. Explain training, validation, and test sets with diagram.
- d. Define overfitting and underfitting.
- e. Differentiate the labelled data and unlabelled data with suitable examples

CO #	Blooms Level
CO1	K2
CO1	K3
CO1	K2
CO3	K1
CO1	K1

PART – B**(10 x 5 = 50 Marks)**Answer **ALL** the questions

2. a. Consider cost function:
 $J(\theta) = (\theta - 6)^2 + 4$. Initial $\theta = 2$, learning rate $\alpha = 0.2$. Perform the following tasks :
 - a. Compute gradient $\partial J / \partial \theta$
 - b. Perform two iterations of Gradient Descent
 - c. Identify whether the function is **convex or non-convex**
 - d. State whether the algorithm converges to local or global minimum
- b. Define convex and non-convex functions and differentiate between Local minima and Global minima

Marks	CO #	Blooms Level
5	CO1	K3

(OR)

- c. A model makes predictions on a dataset and perform the following task:
 1. Compute Mean Squared Error (MSE)
 2. Compute Mean Absolute Error (MAE)
 3. Identify which loss penalizes large errors more and justify
 4. Explain how this relates to Empirical Risk Minimization (ERM)
- d. Explain Bias-variance tradeoff with a suitable example
- 3.a. A dataset is given as the following and apply Statistical Feature Selection Techniques as Chi-Square test of goodness fit where the Critical Value At: $\alpha = 0.05$ is 5.99. Conclude whether the feature is relevant or independent

Actual (y)	Predicted ($\hat{y}$)
3	2.5
-1	0
2	2
7	8

5 CO1 K3

Feature (Color)	Class = Yes	Class = No
Red	20	10
Blue	15	15
Green	5	15

5 CO1 K2

5 CO2 K3

- b. A feature is divided into 3 groups based on class: Apply one-way ANOVA and Compute F-statistic to decide if the feature is significant. F-critical (from F-table at $\alpha = 0.05$ is ≈ 5.14)

Class A	Class B	Class C
5	7	6
6	8	5
7	9	4

5 CO2 K3

(OR)

- c.

X	10	20	30	40	50
---	----	----	----	----	----

5 CO2 K3

Let us consider the below dataset having the

Y	100	80	60	40	20
---	-----	----	----	----	----

one feature and a target . Perform Spearman rank correlation (ρ) and Decide usefulness of feature

d. Given the following joint probability distribution and compute the followings : 5 CO2 K3

1. Compute marginal probabilities $P(X)$ and $P(Y)$
2. Calculate Mutual Information $MI(X, Y)$
3. State whether X and Y are independent or dependent
4. Comment on whether X is a useful feature for predicting Y

X	Y	P(X,Y)
0	0	0.2
0	1	0.3
1	0	0.1
1	1	0.4

4.a. The dataset consists of one feature(X) and a target (Y) . Apply the gradient descent technique to compute the cost function(MSE) .Initial parameters: $\theta_0 = 0$, $\theta_1 = 1$,Learning rate $\alpha = 0.1$ 5 CO3 K3

X	Y
1	2
2	4

b. Explain overfitting and underfitting with suitable diagrams. 5 CO3 K2

(OR)

c. Given dataset where $X(1,2,3)$ and $Y(1,4,9)$.Perform the Quadratic Polynomial Regression to Predict value of Y when $X=4$ 5 CO3 K3

d. Describe the following regularization techniques: Ridge Regression (L_2) , Lasso Regression (L_1) , Elastic Net 5 CO3 K2

5.a. Explain briefly Logistic Regression and explain how it differs from Linear Regression. Derive the logistic (sigmoid) function and explain its properties. 5 Co4 K2

b. Consider the following training dataset: 5 CO4 K3

1. This phone is excellent
2. I like this phone
3. This phone is good
4. This phone is bad
5. I dislike this phone

Apply the NB classifier to classify the new instance : ***“This product is good”***

(OR)

c. From the above dataset ,apply the K-Mean clustering to Classify the point $P(3, 3)$ ($k = 3$) 5 CO5 K3

Point	X	Y	Class
A	1	2	Red
B	2	3	Red
C	6	7	Blue

d. Write a short notes on the followings : 5 CO4 K2

- a. Hard margin vs soft margin, Kernel
- b. functions(Linear, Polynomial, Radial Basis Function)

6.a. Consider the following 2D dataset and apply the PCA of the following:- 5 CO5 K3

X	2	0	3	1
Y	0	2	1	3

- a. Calculate the covariance matrix.
- b. Find the eigenvalues and eigenvectors.
- c. Reduce the data to 1 principal component (1D).

b. Write a short notes on the following: Bagging, Boosting, 5 CO5 K2

(OR)

c. Given the following data points: $A(1,1)$, $B(2,2)$, $C(5,5)$, $D(6,6)$ Use Agglomerative Hierarchical Clustering with Euclidean distance and Single Linkage. Demonstrate the dendrogram 5 CO5 K3

d. Explain briefly L_1 and L_2 regularization with a suitable example 5 CO3 K2

--- End of Paper ---