

Roll No. :

MID SEMESTER EXAMINATION MARCH 2025

Name of the Program: **B.Tech**

Semester: VI

Name of the Course: Advanced Machine Learning

Course Code: TCS 682

Time : 90 Minutes

Maximum Marks : 50

Note:

- i. Answer all the questions by choosing any one of the sub questions
- ii. Each questions carries 10 marks

Q1	(10 Marks)	CO1									
(a)	What are the properties of a probability distribution? Discuss with examples of uniform and normal distributions.										
	OR										
(b)	Write a note on Decision Theory. Explain what the misclassification rate (error rate) is in the context of classification problems. How is it calculated, and what are its implications on model performance?	CO2									
Q2	(10 Marks)	CO2									
(a)	Compare the advantages and disadvantages of parametric and non-parametric methods in terms of model flexibility, interpretability, and computation.										
	OR										
(b)	Consider a dataset with 1000 samples and 50 features. If you increase the dimensionality to 1000 features without increasing the number of samples, explain the impact on the performance of a clustering algorithm like K-means.	CO3									
Q3	(10 Marks)	CO2									
(a)	Explain the difference between the Maximum Likelihood Estimation (MLE) approach and the Bayesian approach to parameter estimation.										
	OR										
(b)	For a classification problem with a confusion matrix: <table><tr><td></td><td>Predicted Positive</td><td>Predicted Negative</td></tr><tr><td>Actual Positive</td><td>40</td><td>10</td></tr><tr><td>Actual Negative</td><td>20</td><td>30</td></tr></table> Calculate Precision, Recall, and Accuracy.		Predicted Positive	Predicted Negative	Actual Positive	40	10	Actual Negative	20	30	CO3
	Predicted Positive	Predicted Negative									
Actual Positive	40	10									
Actual Negative	20	30									
Q4	(10 Marks)	CO1									
(a)	You have trained a machine learning model on a dataset and computed the bias and variance for the model's predictions. The bias is high, and the variance is low. How would you modify the model to improve its performance?										
	OR										
(b)	What is Lasso regression, and how does it help in feature selection?	CO1									
Q5	(10 Marks)	CO2									
(a)	Given a set of data points $X=\{3,4,5,6,7\}$, apply Kernel Density Estimation (KDE) with a Gaussian kernel and bandwidth of 1.0. Calculate the estimated density at $x=5$.										
	OR										

(b)	What is entropy in information theory? How does information gain relate to entropy?	CO2 ,3
-----	---	-----------