

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY-GURUJADA VIZINAGARAM
III B. Tech II Semester Regular/Supplementary Examinations, November -2025
MACHINE LEARNING
(CSE)

Time: 3 hours
70

Max. Marks:

Answer any **FIVE** Questions **ONE** Question from **Each unit**
All Questions Carry Equal Marks

		UNIT-I																					
1.	a)	Compare and contrast the problem-solving approaches of machine learning and traditional programming, delineating the steps integral to designing a learning system. Explain with an example that show how training experience affects the learning outcomes in terms of feedback provided by training experience, interaction with learning system and hypothesis space considered during the system design.	[7M]																				
	b)	Analyze the impact of high-dimensional data on machine learning models. What challenges arise, and what techniques can be used to mitigate these issues?	[7M]																				
		(OR)																					
2.	a)	Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For the healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the precision and recall for the data.	[7M]																				
	b)	What is the difference between unsupervised and supervised learning? How is reinforcement learning different from these two methods? Explain.	[7M]																				
		UNIT-II																					
3.	a)	(i) How can we help ensure that our supervised learning models are not overfit? (ii) In linear regression, what does the error term, ϵ , capture? What about R^2 ?	[4+4 M]																				
	b)	Discuss the applications, strength and weakness of KNN, Logistic regression and Naïve Bayes.	[6M]																				
		(OR)																					
4.	a)	Using the dataset besides, we want to build a decision tree which classifies Y as T/F given the binary variables A, B, C. Draw the tree that would be learned by the greedy algorithm with zero training error. You do not need to show any computation.	[7M]																				
		<table><tr><td>A</td><td>B</td><td>C</td><td>Y</td></tr><tr><td>F</td><td>F</td><td>F</td><td>F</td></tr><tr><td>T</td><td>F</td><td>T</td><td>T</td></tr><tr><td>T</td><td>T</td><td>F</td><td>T</td></tr><tr><td>T</td><td>T</td><td>T</td><td>F</td></tr></table>	A	B	C	Y	F	F	F	F	T	F	T	T	T	T	F	T	T	T	T	F	
A	B	C	Y																				
F	F	F	F																				
T	F	T	T																				
T	T	F	T																				
T	T	T	F																				
	b)	Is this tree (build in above part) optimal (i.e. does it get zero training error with minimal depth)? Explain in less than two sentences. If it is not optimal, draw the optimal tree as well.	[7M]																				
		UNIT-III																					
5.	a)	Describe the main differences between Linear SVM Classification and Nonlinear SVM Classification. How does the kernel trick enable nonlinear	[7M]																				

		decision boundaries?	
	b)	Evaluate the limitations of linear regression for classification tasks. Additionally, demonstrate how logistic regression improves classification performance compared to linear regression. Provide the cost function used in logistic regression.	[7M]
		(OR)	
6.	a)	Differentiate between Bagging and Pasting in ensemble learning. Under what circumstances would one be preferred over the other?	[7M]
	b)	Consider a binary classification problem with two classes: Class 0 and Class 1. You are given a dataset with the following points: Class 0: (1, 2), (2, 3), (3, 3) Class 1: (2, 1), (3, 2), (4, 1) Using a linear SVM classifier, find the equation of the decision boundary (hyperplane) separating the two classes. Assume $C = 1$ for this calculation. Also discuss the soft margin and kernel trick in details.	[7M]
		UNIT-IV	
7.	a)	Consider the following design matrix, representing four sample points $X_i \in \mathbb{R}^2$. We want to represent the data in only one dimension, so we turn to principal components analysis (PCA). Compute the unit-length principal component directions of X , and state which one the PCA algorithm would choose if you request just one principal component. Please provide an exact answer, without approximation.	[7M]
		$X = \begin{bmatrix} 4 & 1 \\ 2 & 3 \\ 5 & 4 \\ 1 & 0 \end{bmatrix}$	
	b)	Apply K(=2) means algorithm over the data (185, 72), (170, 56), (168, 60), (179, 68), (182, 72), (188, 77) up to two iterations and show the clusters. Initially chose first two object as initial centroid.	[7M]
		(OR)	
8.	a)	Consider a dataset with the following points: Points: (2, 2), (2, 3), (3, 2), (8, 7), (9, 7), (8, 8), (20, 20), (21, 21), (22, 22). Apply the DBSCAN algorithm to this dataset with the parameters epsilon (ϵ) = 2 and minimum points (MinPts) = 3. Identify the clusters formed by the algorithm and classify any points that are labeled as noise.	[7M]
	b)	Discuss the application of unsupervised learning techniques in data preprocessing, focusing on how tools Transform facilitate scalable feature extraction and provide examples of real-world applications where these techniques have been instrumental in improving data quality and model performance.	[7M]
		UNIT-V	
9.	a)	Explain the role of optimization in training the neural network. Why it is important for improving model performance?	[7M]
	b)	Illustrate the significance of backpropagation in training neural networks. How does it adjust weights?	[7M]
		(OR)	
10.	a)	Differentiate between the following optimization algorithm and describe the suitable use case for each: (i) SGD (ii) RMSprop (iii) Adagrad (iv) Adam	[7M]
	b)	How does TensorFlow 2 handle data loading and preprocessing? Explain with an example.	[7M]
