

# Machine Learning

## Supervised Learning Quiz

### Set 1

#### Instructions:

- Answer all questions clearly and completely.
- Show your work for subjective questions.
- For multiple choice questions, **circle** the correct option.
- **Marks Distribution:** MCQ (9 marks) + Subjective (17 marks) = **26marks total**

### Multiple Choice Questions

**Q1****[2 marks]**

In a decision tree, which impurity measure is most commonly used for classification tasks?

- (A) R-squared
- (B) Mean Absolute Error (MAE)
- (C) Mean Squared Error (MSE)
- (D) Gini Impurity

**Q2****[3 marks]**

Given the confusion matrix below for a binary classification problem:

	Predicted 0	Predicted 1
Actual 0	85	15
Actual 1	10	90

What is the precision of the classifier?

- (A) 0.857
- (B) 0.875
- (C) 0.90
- (D) 0.85

Q3

[2 marks]

Which regularization technique adds a penalty term proportional to the sum of absolute values of parameters?

- (A) L1 Regularization (Lasso)
- (B) Elastic Net
- (C) L2 Regularization (Ridge)
- (D) Dropout

Q4

[2 marks]

Which of the following best describes the bias-variance tradeoff in machine learning?

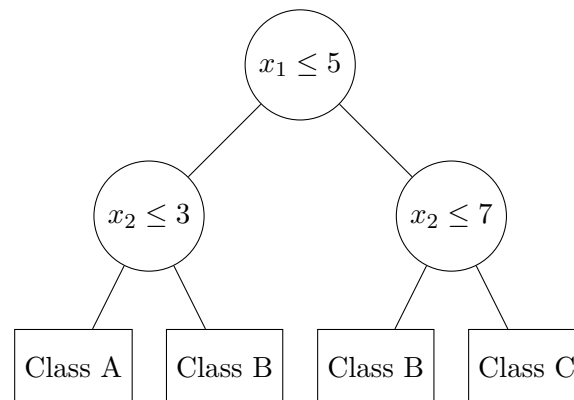
- (A) Bias and variance are independent and don't affect each other
- (B) Reducing bias typically increases variance, and vice versa
- (C) Variance only matters in unsupervised learning
- (D) High bias models always perform better than high variance models

## Subjective Questions

## Q5

[Total: 6 marks]

Analyze the decision tree structure below:



- What is the maximum depth of this tree? [1 mark]
- Calculate the Gini impurity for a node with class distribution: Class A: 40 samples, Class B: 30 samples, Class C: 10 samples. [3 marks]
- Explain why pruning might be beneficial for this tree. [2 marks]

## Q6

[Total: 5 marks]

Consider the following dataset for linear regression:

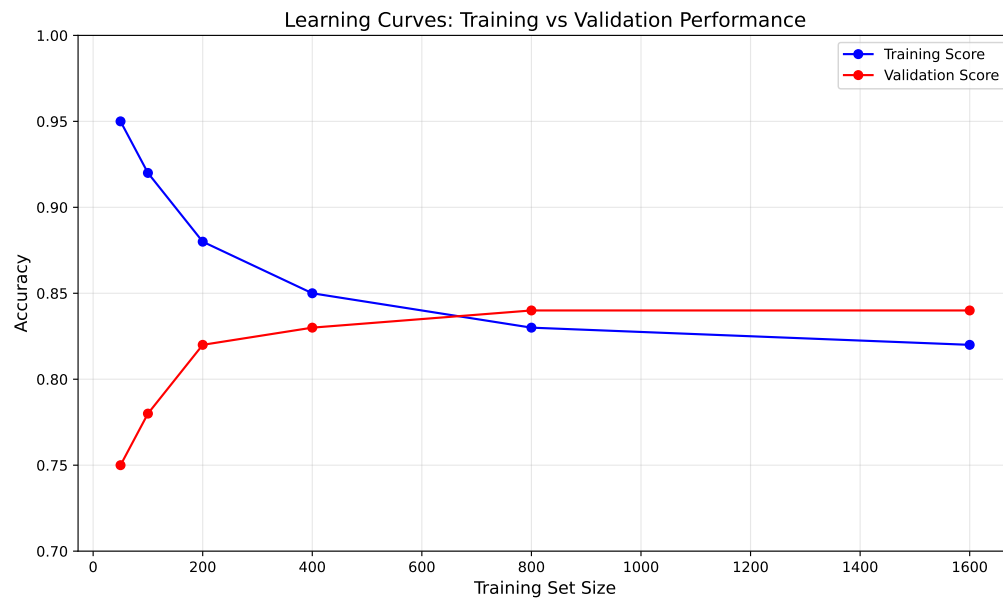
Sample	Feature 1	Feature 2	Target
1	1	2	6
2	3	1	7
3	2	3	9
4	4	2	10

- Calculate the mean squared error (MSE) if the model predicts  $\hat{y} = 5.9, 7.1, 8.8, 9.9$  respectively. [3 marks]
- If we use L2 regularization with  $\lambda = 0.05$ , write the complete loss function. [2 marks]

Q7

[Total: 6 marks]

Given the learning curves shown below:



- a) Identify whether the model is suffering from high bias or high variance. Justify your answer. [3 marks]
- b) Suggest two specific techniques to improve the model performance. [2 marks]
- c) If the training accuracy is 95% and validation accuracy is 78%, calculate the overfitting gap. [1 mark]