Mid-Semester Examination (March 2025)
MCAC 204: Artificial Intelligence and Machine Learning
Semester II
Year of admission: 2024

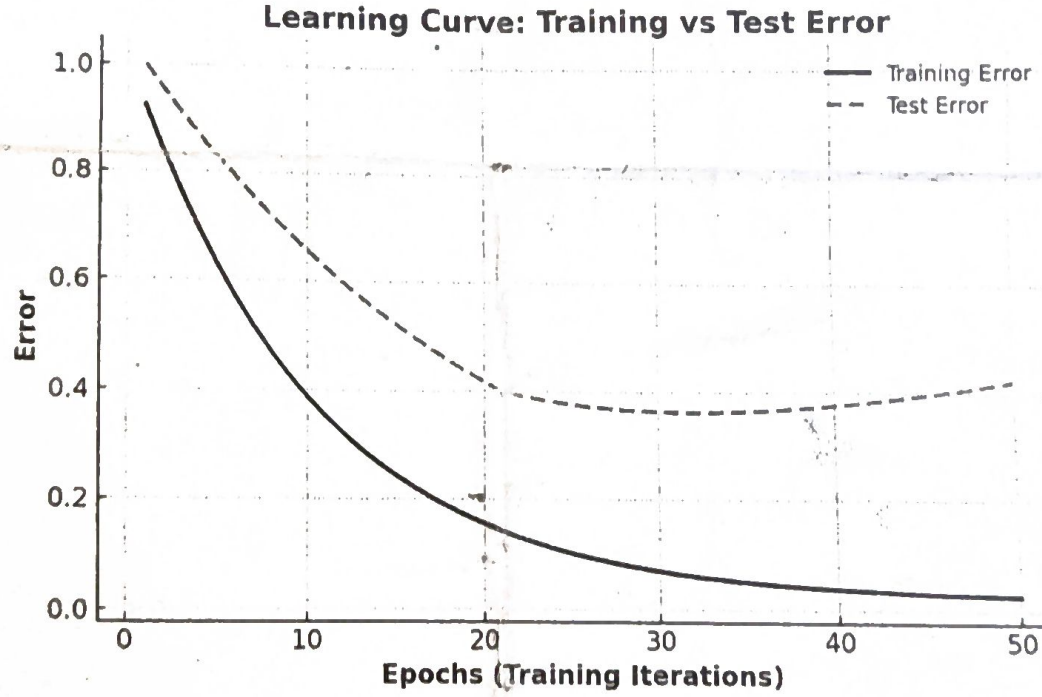
Time: One Hour **Maximum Marks: 25**

1. Consider the following dataset: (5)

Instance	Weather	Temperature	Humidity	Wind	PlayTennis
1	Sunny	Warm	Normal	Strong	Yes
2	Sunny	Warm	High	Strong	Yes
3	Rainy	Cold	High	Strong	No
4	Sunny	Warm	Normal	Weak	Yes

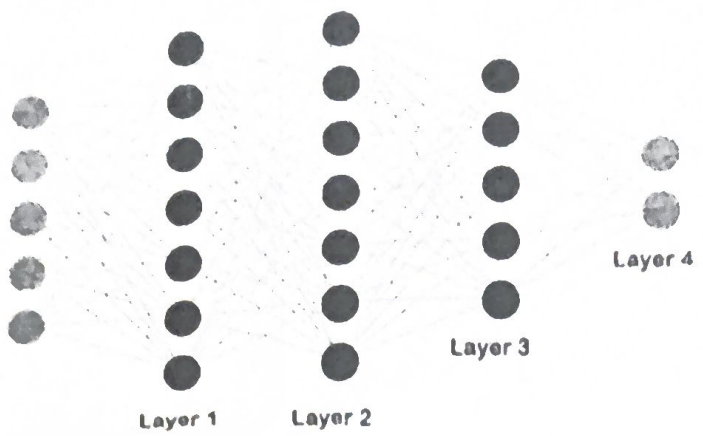
Apply the Candidate Elimination Algorithm to determine the version space for the above dataset.

2. Consider the learning curve of a linear regression model, which displays training error and test error over multiple epochs.



a) Do you observe any issue with the trend of training and test error as the number of epochs increases? If yes, clearly identify and explain the problem. (2)
b) Suggest an appropriate method to address this problem. Modify the cost function accordingly and explain how the modification helps. (3)

3. Consider the following neural network comprising four layers: layer 1, layer 2, layer 3, and layer 4. Layers 1, 2, and 3 are hidden layers, and layer 4 is the output layer.



Determine the following:

(i) dimensions of $W^{[1]}, b^{[1]}, W^{[2]}, b^{[2]}, W^{[3]}, b^{[3]}, W^{[4]}, b^{[4]}$ (3)

(ii) number of trainable parameters. (2)

4. A neural network is initialized with the following settings:
   - The input data is zero-centered.
   - Weights are initialized independently with mean *zero* and variance 0.001.
   - Biases are initialized to 0.
   - The learning rate is small and fixed (it cannot be tuned).

a) Derive the slope (first derivative) of the sigmoid and tanh activation functions. (3)

b) Which activation function is expected to result in a higher gradient during the first weight update? Justify your answer.. (2)

5. Consider a two-layer neural network with the following configuration:
   Input Layer: $X_1, X_2, \ldots, X_{10}$
   Hidden Layer:
   - *Five* activation units.
   - Activation function: ReLU
   - Parameters: $W^{[1]}$ *and* $b^{[1]}$

   Output Layer:
   - One activation unit
   - Activation function: Sigmoid
   - Parameters: $W^{[2]}$ *and* $b^{[2]}$

   The network uses the following loss function:

$$L(A^{[2]}, y) = -[y \log(A^{[2]}) + (1 - y) \log(1 - A^{[2]})]$$

   where y is the true label and $A^{[2]}$ is the predicted output.

a) Write the forward propagation rule for the above neural network. (2)

b) Determine the gradients computed during the backpropagation: $\frac{\partial L}{\partial W^{[2]}}$ and $\frac{\partial L}{\partial b^{[2]}}$. (3)

5x 10 10x1

5x1