Crash Course GATE 2025
Machine Learning
MLP

Q1 Consider a single perceptron with sign activation function. The perceptron is represented by weight vector $[0.4 \; -0.3 \; 0.1]^t$ and a bias q = 0. If the input vector to the perceptron is X = [0.2 0.6 0.5] then the output of the perceptron is:

(A) 1 (B) 0

(C) –0.05 (D) – 1

Q2 Let W_{ij} represents weight between node i at layer k and node j at layer (k – 1) of a given multilayer perceptron.

The weight updation using gradient descent method is given by

Where a and E represents learning rate and Error in the output respectively.

(A) $W_{ij}(t+1) = W_{ij}(t) + \alpha\frac{\partial E}{\partial W_{ij}}, 0 \le \alpha \le 1$

(B) $W_{ij}(t+1) = W_{ij}(t) - \alpha\frac{\partial E}{\partial W_{ij}}, 0 \le \alpha \le 1$

(C) $W_{ij}(t+1) = \alpha\frac{\partial E}{\partial W_{ij}}, 0 \le \alpha \le 1$

(D) $W_{ij}(t+1) = -\alpha\frac{\partial E}{\partial W_{ij}}, 0 \le \alpha \le 1$

Q3 If you use 10 - fold cross - validation on a dataset of 1,000 samples, how many times will each sample be used for testing?

(A) 10 times (B) 1 time

(C) 5 times (D) 100 times

Q4 What is the function of the hidden layers in a multilayer perceptron (MLP)?

(A) Hidden layers perform feature extraction and transformation

(B) Hidden layers provide direct access to the input data

(C) Hidden layers perform the final classification or regression task

(D) Hidden layers are not necessary in an MLP

Q5 If a dataset is divided into 5 folds for cross-validation, how many data points are in each fold if the total number of data points is 500?

(A) 100 (B) 50

(C) 200 (D) 500

Q6 If a neural network with 4 hidden layers each having 50 neurons is trained with an input size of 20 and an output size of 10, how many parameters are there in the first hidden layer ?

(A) 1050 (B) 1050 + 50

(C) 1050 + 50 + 10 (D) 1050 + 50 + 50

Q7 What is the number of parameters in a neural network with 10 input neurons, 10 neurons in the hidden layer, and 2 output neurons, assuming the hidden layer uses a bias term?

(A) 120 (B) 132

(C) 140 (D) 150

Q8 For a neural network with 5 layers, if each layer has 100 neurons, and each neuron has connections to all neurons in the previous layer, how many parameters are there in the network ?

(A) 40400 (B) 50500

(C) 51000 (D) 50500 + 100

Q9 Consider a dataset with 1,000 samples and a k-fold cross-validation procedure with k=5. You are using a linear model with regularization (e.g., Lasso or Ridge regression) and performing cross-validation to tune the regularization parameter. Suppose the model's performance is evaluated using Mean Squared Error (MSE). What would be the total number of MSE values computed during this process ?

Q10 The Proposition (P Þ Q) Ù (Q Þ P). is a contingency.

(A) tautology (B) contradiction

(C) contingency (D) absurdity

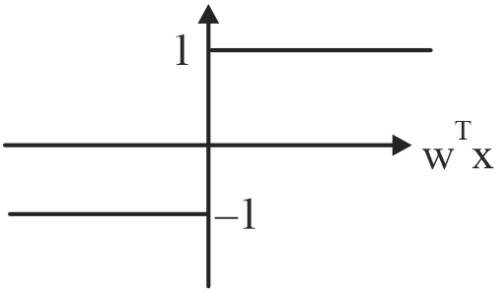
Answer Key

Q1	D
Q2	B
Q3	B
Q4	A
Q5	A
Q6	A
Q7	B
Q8	A
Q9	50
Q10	C

Hints & Solutions

Note: scan the QR code to watch video solution

Q1 Text Solution:



$w^T x \Rightarrow w_1x_1 + w_2x_2 + w_3x_3 \Rightarrow xw \Rightarrow$

$$(\cdot 2 \cdot 6 \cdot 5)\begin{bmatrix} \cdot 4 \\ -\cdot 3 \\ \cdot 1 \end{bmatrix}$$

$\Rightarrow -\cdot 05$

Q2 Text Solution:

In the context of weight update in a multilayer perceptron using gradient descent, the correct update rule is given by:

(B) $W_{ij}(t+1) = W_{ij}(t) - \alpha \frac{\partial E}{\partial W_{ij}}$Wij(t+1)=Wij(t)−α∂Wij∂E

Here's the breakdown:

- **Gradient Descent Method**: This is a method used to minimize the error function EE by iteratively adjusting the weights in the direction opposite to the gradient of the error function with respect to those weights.
- **Learning Rate (αα)**: This controls the size of the step we take in the direction of the negative gradient.
- **Weight Update**: The new weight $W_{ij}(t+1)$Wij(t+1) is computed by subtracting the product of the learning rate and the gradient of the error function with respect to the weight W_{ij} Wij from the current weight $W_{ij}(t)$Wij(t).

Therefore, the update formula subtracts the gradient (scaled by the learning rate) from the current weight to decrease the error, making option (B) the correct choice.

Q3 Text Solution:

In 10-fold cross-validation, each sample in the dataset is used exactly once as a part of the test set and k-1 times as part of the training set. Therefore, each of the 1,000 samples will be used for testing only once.

Q4 Text Solution:

The hidden layers in a multilayer perceptron (MLP) perform feature extraction and transformation by applying nonlinear transformations to the input data. Each hidden layer learns increasingly abstract and complex representations of the input data, allowing the network to capture intricate patterns and relationships in the data.

Q5 Text Solution:

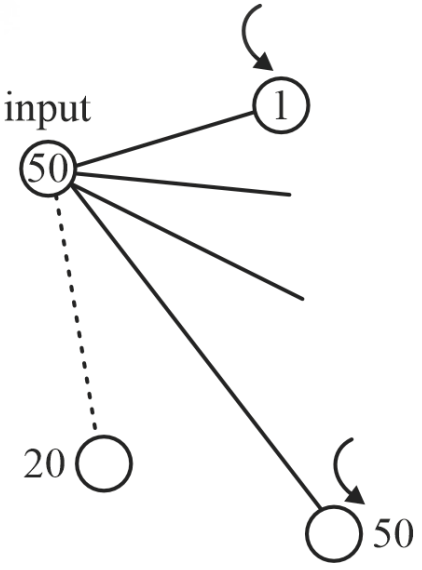
Each fold will contain $\frac{500}{5} = 100$ data points.

Q6 Text Solution:

50 × 20

1000 + 50

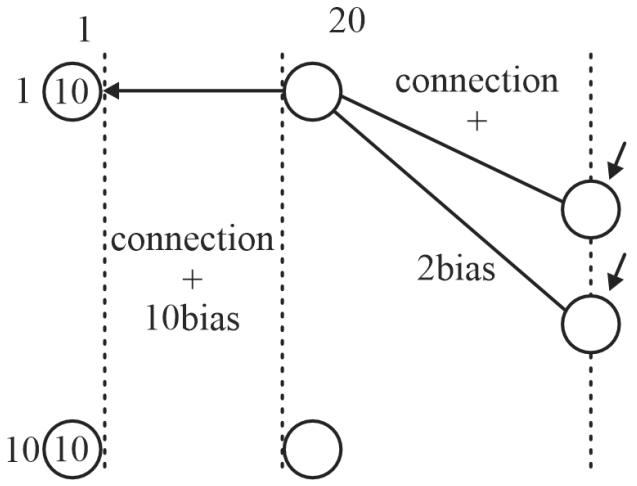
= 1050



Q7 Text Solution:

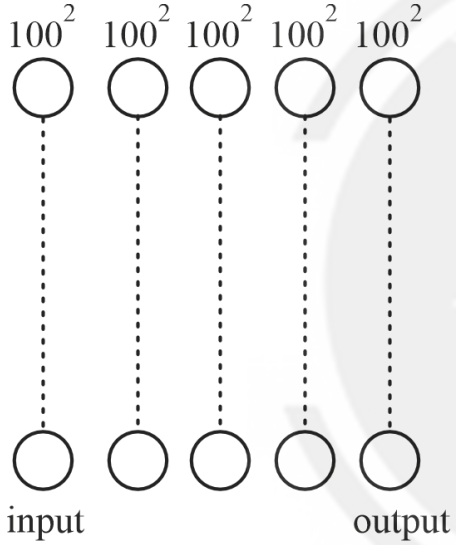
100 + 10 + 20 + 2

$\Rightarrow$ 132



Q8 Text Solution:

(Assume 1st layer as input layer and 5th as output layer)



$4 \times 100^2 + 4 \times 100$

$\Rightarrow$ 40000 + 400

$\Rightarrow$ 40400.

Q9 Text Solution:

5 fold Gross validation

For each l we do 5 time testing 5 MSE

for 10 different l 50 MSE.

Q10 Text Solution:

P	Q	P Q	Q P	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
0	0	1	1	1
0	1	1	0	0
1	0	0	1	0
1	1	1	1	1