

2. What is the purpose of the activation function in a neural network?

- A) To initialize weights
- B) To introduce non-linearity into the network
- C) To perform a summation of inputs
- D) To calculate the loss function

Answer: B

3. Which activation function is commonly used in hidden layers of neural networks?

- A) Sigmoid
- B) Tanh
- C) ReLU
- D) Softmax

Answer: C

4. What is backpropagation in the context of neural networks?

- A) A technique to prevent overfitting
- B) A method for initializing weights
- C) A process to calculate gradients for updating weights
- D) A type of neural network architecture

Answer: C

5. What is an epoch in training a neural network?

- A) A single pass through the entire training dataset
- B) The number of neurons in the hidden layer
- C) A technique to prevent overfitting
- D) A regularization method

Answer: A

7. In neural networks, what is a hyperparameter?

- A) A parameter that is learned during training
- B) A parameter that is set before training begins
- C) A parameter that controls the activation function
- D) A parameter that controls the learning rate

Answer: B

8. What is the role of the bias term in a neural network?

- A) To adjust the output of the activation function
- B) To prevent the network from overfitting
- C) To allow the model to fit the data better
- D) To normalize the input data

Answer: C

9. Which type of neural network is typically used for image classification tasks?

- A) Recurrent Neural Network (RNN)
- B) Convolutional Neural Network (CNN)
- C) Radial Basis Function Network (RBFN)
- D) Deep Belief Network (DBN)

Answer: B

10. What is the difference between a shallow and a deep neural network?

- A) The number of hidden layers
- B) The type of activation function used
- C) The size of the dataset

D) The type of data being processed

Answer: A

11. Which of the following statements about deep learning is true?

A) Deep learning models are shallow neural networks with few layers.

B) Deep learning requires large amounts of labeled data and computational power.

C) Deep learning models are always more accurate than traditional machine learning models.

D) Deep learning is only applicable to image data.

Answer: B

12. Suppose  $\hat{y}=0.8$  and  $y=1$ . What is the value of the “Logistic Loss”?

A) 0.198

B) 0.223

C) 0.999

D) 0.298

ANSWER: 0.223

14. Consider the numpy array x: `x=np.array([[[1],[3],[4],[5]],[[2],[3],[5],[8]])` What is the shape of x?

A) (4,5,6)

B) (4,2,1)

C) (1,2,4)

D) (2,4,1)

ANSWER: (2,4,1)

15. Consider the following random arrays a and b, and c:

`a=np.random.randn(2,3)`

`b=np.random.randn(3,5)`

`c = np.dot(a, b)`

What will be the shape of c?

A) 2\*5

B) 5\*2

C) 3\*4

D) 2\*3

ANSWER: 2\*5

16. Suppose our input batch consists of 10 RGB images, each of dimension 32x32. We reshape these images into feature column vectors  $y^{[i]}$ . Remember that  $Y=[y^{(1)}, y^{(2)}, \dots, y^{(10)}]$ . What is the dimension of Y?

A) 2072\*10

B) 1072\*10

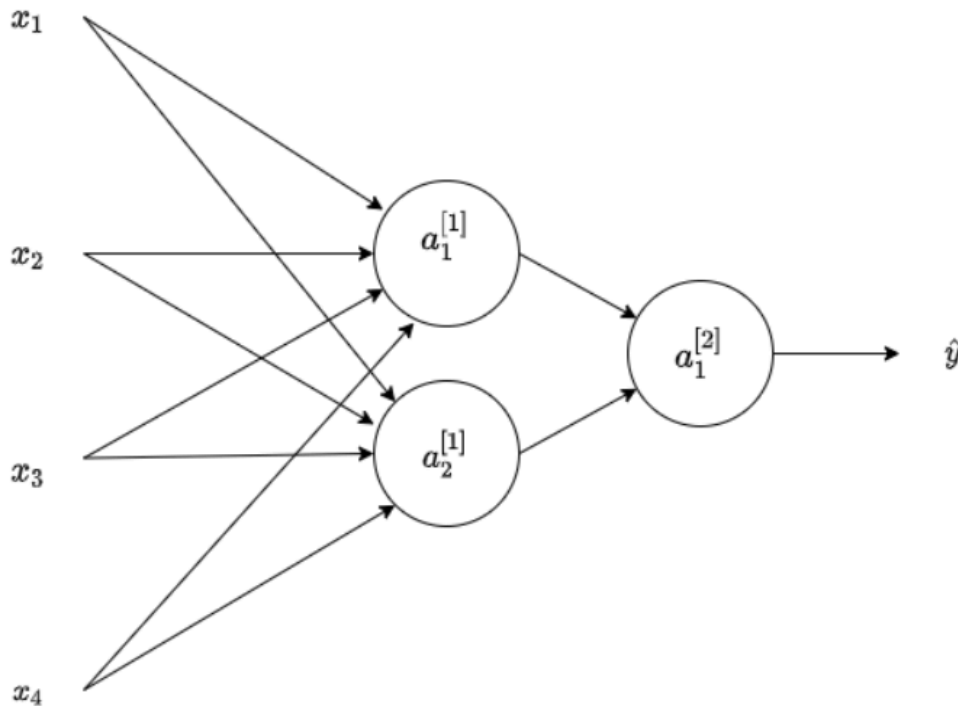
C) 3072\*10

D) 3072\*10

ANSWER :3072\*10

17.

If each input  $x_1, x_2, x_3, x_4$  is associated with a weight, how many weights are there between the input layer and the hidden layer?



- A) 12
- B) 4
- C) 8
- D) 10

ANSWER: 8

18. Consider the following code snippet:

```
x.shape = (5, 3)
```

```
y.shape = (3,)
```

```
for i in range(5):
```

```
    for j in range(3):
```

```
        z[i][j] = x[i][j] + y[j]
```

How do you vectorise this?

- A)  $z = x + y$
- B)  $z = x \times y$
- C)  $z = \text{np.dot}(x, y)$
- D) None

ANSWER:  $z = x + y$

20 You are developing a binary classifier to identify whether an image contains a cat ( $y=1$ ) or a dog ( $y=0$ ). Which activation function would you recommend using for the output layer of your neural network to achieve this binary classification?

- A) Relu
- B) Sigmoid
- C) Tanh
- D) None

ANSWER: Sigmoid function

21. Consider the following code:

```
A = np.random.randn(5,3)
B = np.sum(A, axis = 1, keepdims = True)
What will be B.shape?
```

- A) (1,5)
- B) (5,3)
- C) (5,2)
- D) (5,1)

ANSWER: (5,1)

22. For a given input matrix Z, if Z contains values  $\begin{bmatrix} -2 & 0 & 3 \\ 1 & -4 & 2 \end{bmatrix}$ , what will be the matrix after applying the ReLU activation function?

- A)  $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 2 \end{bmatrix}$
- B)  $\begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 2 \end{bmatrix}$
- C)  $\begin{bmatrix} -2 & 0 & 3 \\ 1 & -4 & 2 \end{bmatrix}$
- D)  $\begin{bmatrix} 2 & 0 & 3 \\ 1 & 4 & 2 \end{bmatrix}$

ANSWER:  $\begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 2 \end{bmatrix}$

23. In a neural network, which of the following is a primary advantage of using the ReLU activation function compared to the sigmoid activation function?

- a) ReLU activation function provides non-linearity to the model.
- b) ReLU activation function can introduce sparsity in the activation map.
- c) ReLU activation function ensures that outputs are bounded between 0 and 1.
- d) ReLU activation function avoids the vanishing gradient problem for positive input values.

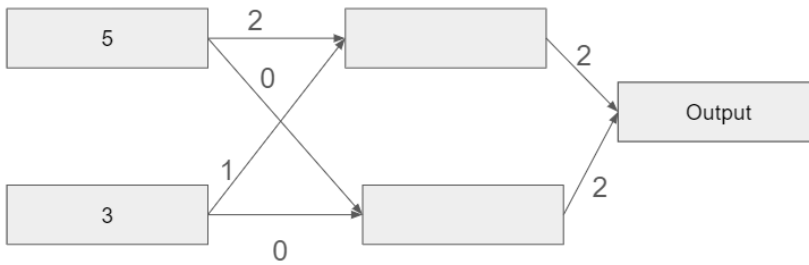
ANSWER:d

24. Which of the following is true for tanh function:

- a. Squashes the value to a range between 0 and 1
- b. Squashes the value to a range between -1 and 1
- c. Outputs zero for negative inputs and input itself for positive input
- d. None

ANSWER: B

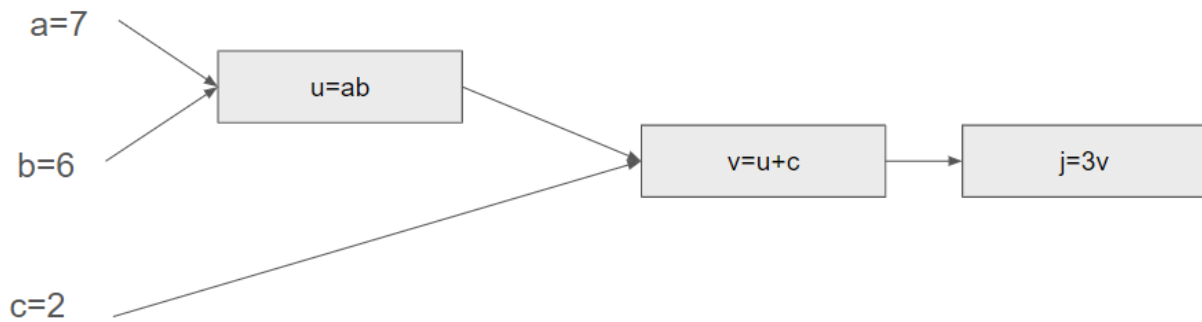
25. What is the error (predicted - actual) for the following network using the ReLU activation function when the input data is  $[5, 3]$  and the actual value of the target (what you are trying to predict) is 5?



- A) 20
- B) 21
- C) 23
- D) 24

ANSWER:21

26. Calculate derivative of  $J$  with respect to  $a$ , when  $J(a,b,c)=3(ab+c)$



- A) 20
- B) 31
- C) 18
- D) 16

ANSWER:18

28. What is the derivative of  $\tanh$  function?

- A)  $1-\tanh(x)$
- B) 1

- C) 0
- D)  $1 - \tanh^2(x)$

ANSWER=  $1 - \tanh^2(x)$

30. In a shallow neural network with one hidden layer, if the input layer has 3 neurons and the hidden layer has 4 neurons, how many weights are there between the input layer and the hidden layer?

- A) 11
- B) 14
- C) 12
- D) 10

ANSWER:12

31 .Who proposed a highly simplified computational model of the brain in 1943?

- A) Alan Turing
- B) John von Neumann
- C) Warren McCulloch and Walter Pitts
- D) Claude Shannon

Answer: C

32. Choose from below all options that are "hyperparameters" of a neural network. Note that you will get marks only if you choose all of the correct answers.

A)The weight matrices  $W^{[l]}$

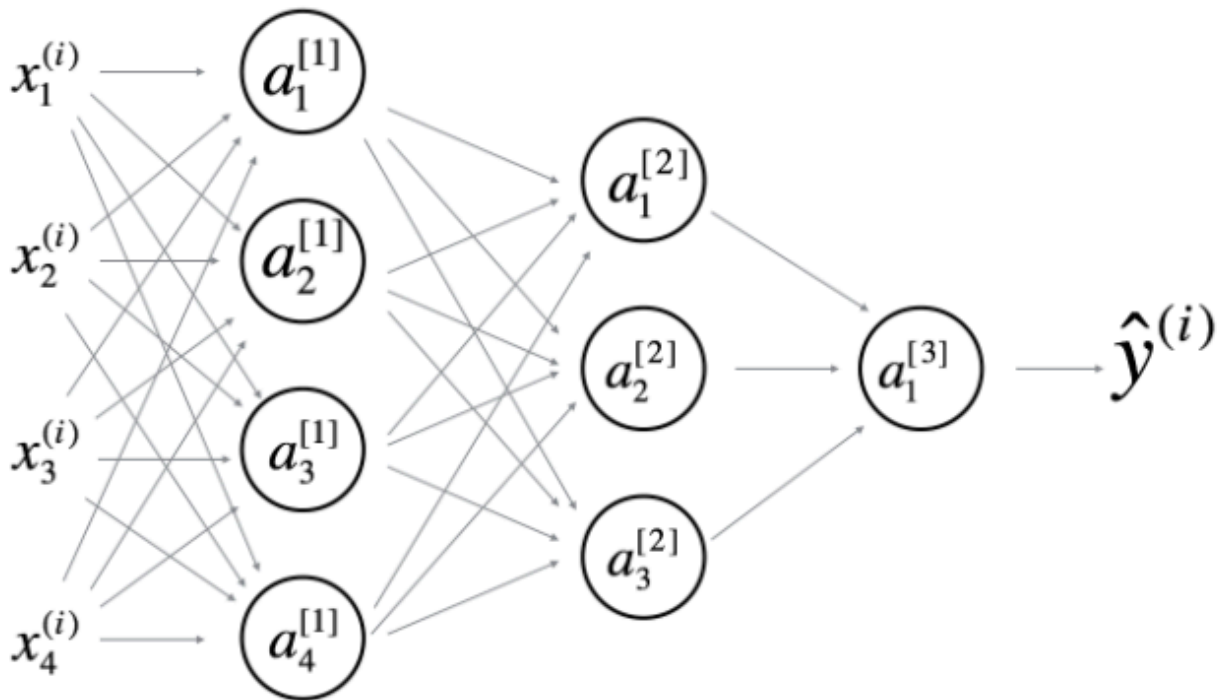
B)The bias vector  $b^l$ .

C)The learning rate  $\alpha$

D)The activation functions  $g^{[l]}$

Answer: C, D

34. Consider the following neural network with 2 hidden layers:



Which of the following statement is true?

- A)  $W^{[1]}$  will have shape (4,4)
- B)  $W^{[1]}$  will have shape (3,4)
- C)  $W^{[2]}$  will have shape (2,3)
- D) None of the options

Answer: A

During the forward propagation in a neural network, why are certain intermediate values, known as the "cache," stored for later use during backpropagation?

- A) To reduce the memory usage of the neural network.
- B) To speed up the forward propagation process.
- C) To ensure that the activation function is applied correctly.
- D) To efficiently compute gradients needed for updating weights and biases.

Answer: D

Which of the following is a key difference between traditional machine learning and deep learning in terms of scalability and hardware requirements?

- A. Traditional machine learning models often require specialized hardware and infrastructure to scale effectively.
- B. Deep learning models can easily scale with less hardware requirements.
- C. Traditional machine learning models are easier to scale and require less specialized hardware.
- D. Deep learning models struggle with scaling due to the need for manual feature selection and tuning.

Answer: C

37. In a deep neural network with 3 layers, where each layer contains 5 neurons, and each neuron is fully connected to the neurons in the previous layer, what is the total number of weights (connections) in the network?

- a) 50
- b) 75
- c) 100
- d) 125

ANSWER: 50

38. In the context of forward propagation in neural network, what does the equation  $Z^{[1]} = W^{[1]}X + b^{[1]}$  represent?

- A) The activation of the first layer
- B) The linear transformation for the first layer
- C) The cost function
- D) The final output of the network

Answer: B

39. In the third layer of a neural network, what does the equation  $A^{[3]} = g^{[3]}(Z^{[3]})$  represent?

- A) The output of the network
- B) The activation of the second layer
- C) The linear transformation of the third layer
- D) The activation of the third layer

Answer: D



40. In a neural network, what is the purpose of the activation function  $\sigma(z_i^{[1]})$  applied to the output of each node in the hidden layer?

- A) To add noise to the input data.
- B) To introduce non-linearity into the model.
- C) To reduce the dimensionality of the input.
- D) To remove irrelevant features from the input data.

Answer: B