

The background of the slide is a light blue color with a large, irregular, darker blue brushstroke shape in the center. The text is written in a white, bold, italicized font within this central shape.

NN RECAP

The output of a neural network is calculated during:

A) Forward propagation

B) Backpropagation

C) Weight initialization

D) Loss function evaluation

If weights in a neural network are initialized to zero, it:

A) Speeds up training

B) Leads to symmetry in updates and prevents learning

C) Reduces overfitting

D) Improves the effectiveness of activation functions

How Improves the effectiveness of activation functions

A) Measure the model's performance by comparing predictions to true values

B) Initialize the weights of the network

C) Update the input features for better learning

D) Determine the size of the neural network

What is the main reason behind using activation functions?

- A) To add dependencies**
- B) To break linearity**
- C) To add linearity**
- D) To break dependencies**

Which of the following represents a batch of size 32 of 16 pixels by 16 pixels grayscale images?

A) (32, 1, 16, 16)

B) (1, 32, 16, 16)

C) (1, 16, 32, 16)

D) (1, 16, 16, 32)

Suppose the value of z is a negative number. What will be the output of the relu activation function? i.e what will be the value of $relu(z)$?

A) The absolute value of z

B) The same value of z

C) Zero

D) Error as relu activation function does not take negative input

**The following are hyperparameters
EXCEPT:**

A) Epochs

B) Hidden Units

C) Learning Rate

D) Loss Function

A neural network has 2 input nodes, 3 hidden nodes, and 2 output nodes. How many learnable parameters are there?

A) 12

B) 10

C) 17

D) 16

Gradient descent optimizes the loss function by increasing its value :

True

False

A neural network with no activation functions behaves like a linear model:

True

False



***GOOD
LUCK***