

# Introduction to Neural Networks

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Agenda

- Neural Networks Quiz
- Deep Learning vs Machine Learning
- Neural Network Architecture
- Activation Functions
- Loss Functions
- Gradient Descent and Backpropagation

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Let's begin the discussion by answering a few questions on neural networks

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following statements are true about Deep Learning in general?

A

In Deep Learning, feature extraction is performed manually, and in Machine Learning it is performed automatically

B

In Deep Learning, feature extraction is performed automatically, and in Machine Learning it is performed manually

C

Deep Learning models are typically less interpretable when compared to Machine Learning models

D

Deep Learning requires comparatively less computational power than Machine Learning

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following statements are true about Deep Learning in general?

A

In Deep Learning, feature extraction is performed manually, and in Machine Learning it is performed automatically

B

In Deep Learning, feature extraction is performed automatically, and in Machine Learning it is performed manually

C

Deep Learning models are typically less interpretable when compared to Machine Learning models

D

Deep Learning requires comparatively less computational power than Machine Learning

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deep Learning vs Machine Learning

## Machine Learning

Requires manual feature extraction and selection

Less effective when working with unstructured data

Requires comparatively less computational power and memory

Are generally more interpretable

## Deep Learning

Automatically does feature extraction

More effective when working with unstructured data

Requires comparatively more computational power and memory\*

Are generally less interpretable

\* Depends on the neural network architecture

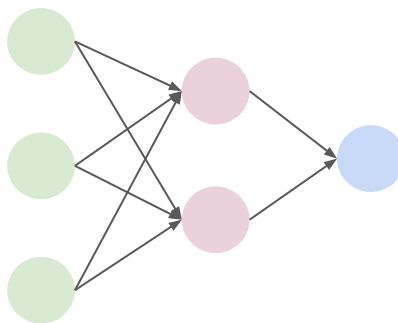
This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

How many parameters (including bias) will be trained in a fully connected neural network architecture shown below?



A

8

B

11

C

6

D

9

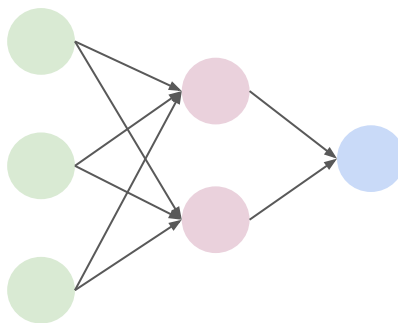
This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

How many parameters (including bias) will be trained in a fully connected neural network architecture shown below?



A

8

B

11

C

6

D

9

This file is meant for personal use by lhughey@gmail.com only.

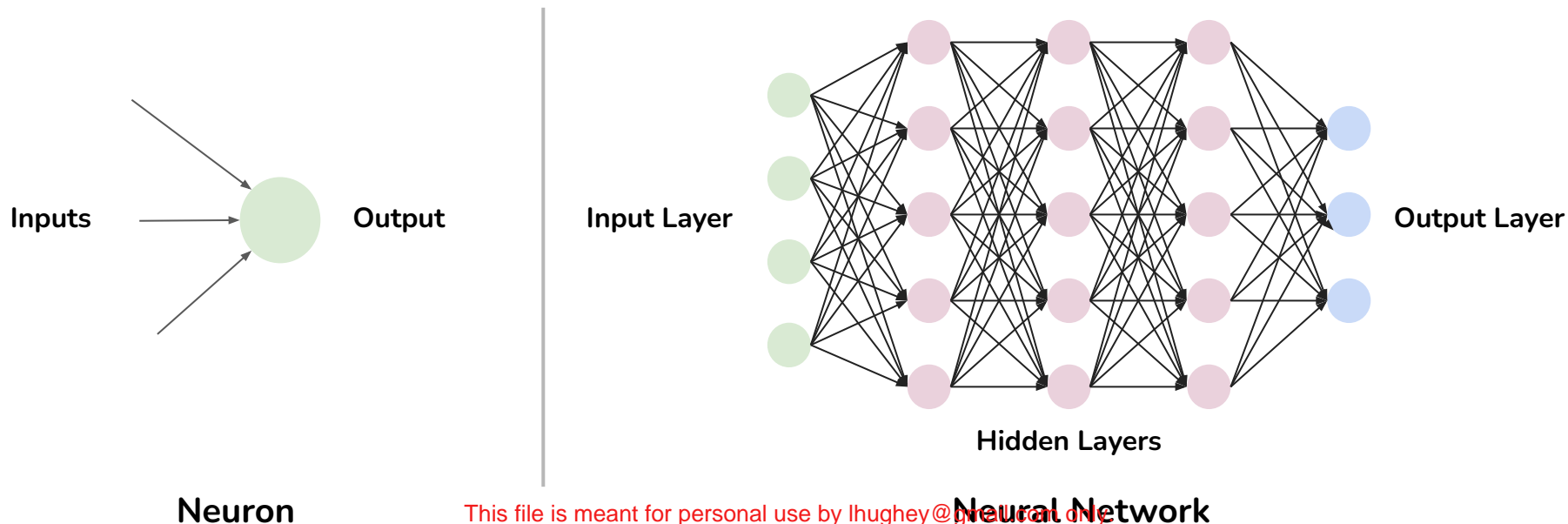
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

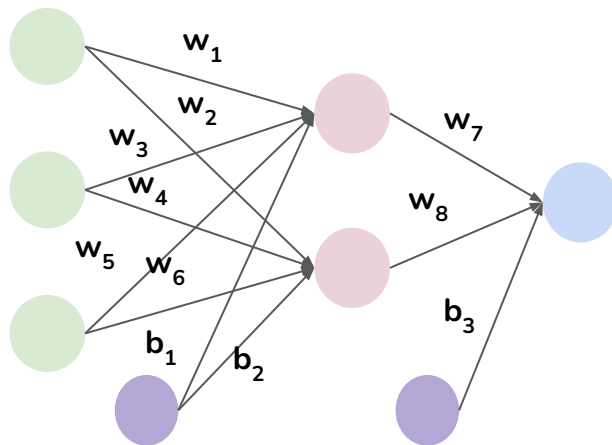


# Neural Network

A neural network architecture comprising interconnected input, hidden, and output layers, facilitating the learning of complex relationships between input and output data



# Neural Network Architecture



**Input Layer x Hidden Layer:** 3 neurons of input layer connected to 2 neurons in the hidden layer |  $3 \times 2 = 6$  connections, i.e., 6 weight parameters | 1 bias parameter for each neuron in the hidden layer |  $6 + 2 = 8$  parameters in total between input and hidden layers

**Hidden Layer x Output Layer:** 2 neurons of hidden layer connected to 1 neuron in the output layer |  $2 \times 1 = 2$  connections, i.e., 2 weight parameters | 1 bias parameter for the neuron in the output layer |  $2 + 1 = 3$  parameters in total between hidden and output layers

So, we have a total of 11 parameters

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

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

A

To compute the gradient of the loss function during backpropagation.

B

To regulate the magnitude of weights and biases in the network.

C

To introduce non-linearity into the network, enabling it to learn complex patterns.

D

To normalize the input data before feeding it into the neural network.

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

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

A

To compute the gradient of the loss function during backpropagation.

B

To regulate the magnitude of weights and biases in the network.

C

To introduce non-linearity into the network, enabling it to learn complex patterns.

D

To normalize the input data before feeding it into the neural network.

This file is meant for personal use by lhughey@gmail.com only.

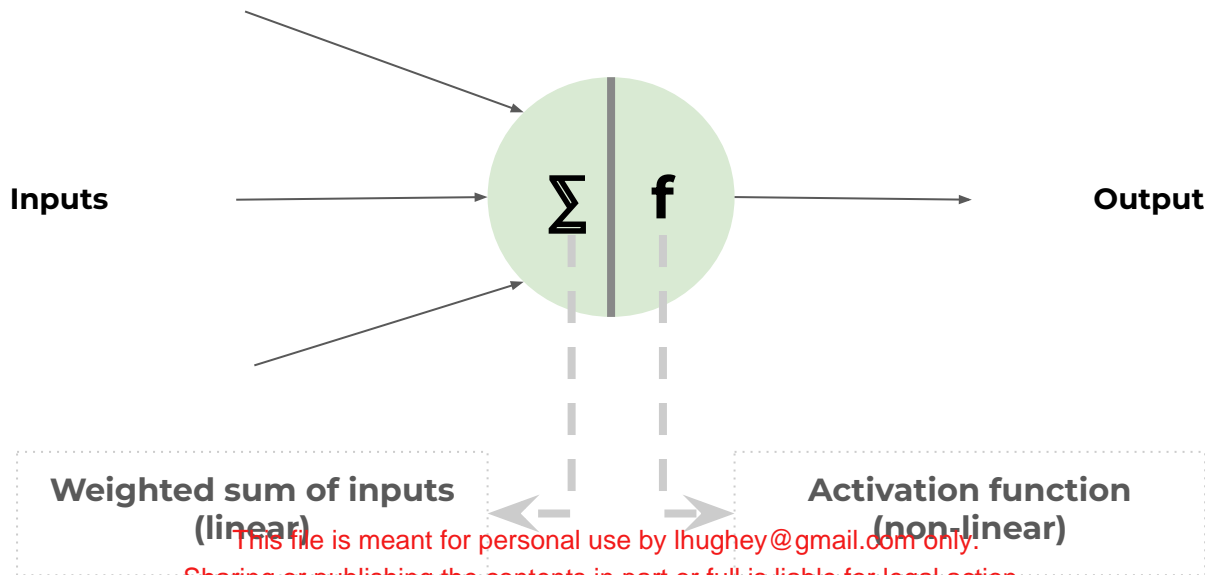
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Activation Function

Determines the output of a neuron based on the input

**Introduces non-linearity** and enables the neural network to learn complex patterns in the data



# Neural Networks Quiz

Which of the following activation functions can be used in a hidden layer?

A

Sigmoid

B

TanH

C

ReLU

D

Leaky ReLU

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following activation functions can be used in a hidden layer?

A

Sigmoid

B

TanH

C

ReLU

D

Leaky ReLU

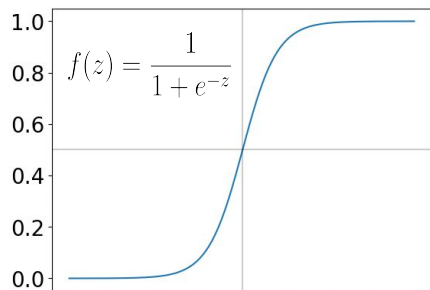
This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

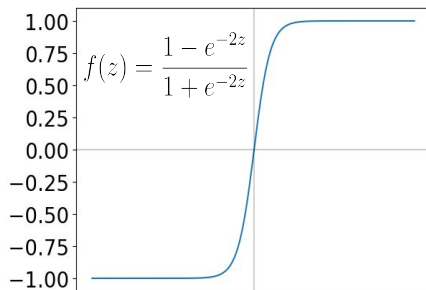
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Activation Functions for Hidden Layers

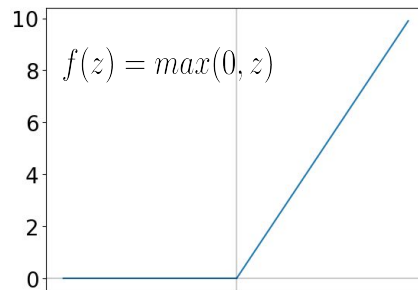
Sigmoid



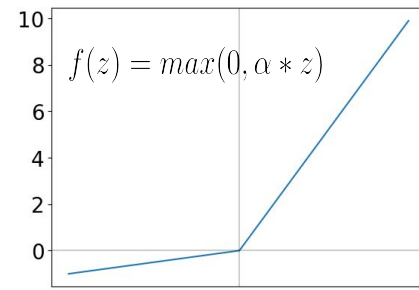
Tanh



ReLU



Leaky ReLU



This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.



# Neural Networks Quiz

Consider a neural network built to predict whether a student will pass or fail in an exam. Which of the following activation functions would be most appropriate to use in the output layer?

A

Sigmoid

B

Softmax

C

Linear

D

Leaky ReLU

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Consider a neural network built to predict whether a student will pass or fail in an exam. Which of the following activation functions would be most appropriate to use in the output layer?

A

Sigmoid

B

Softmax

C

Linear

D

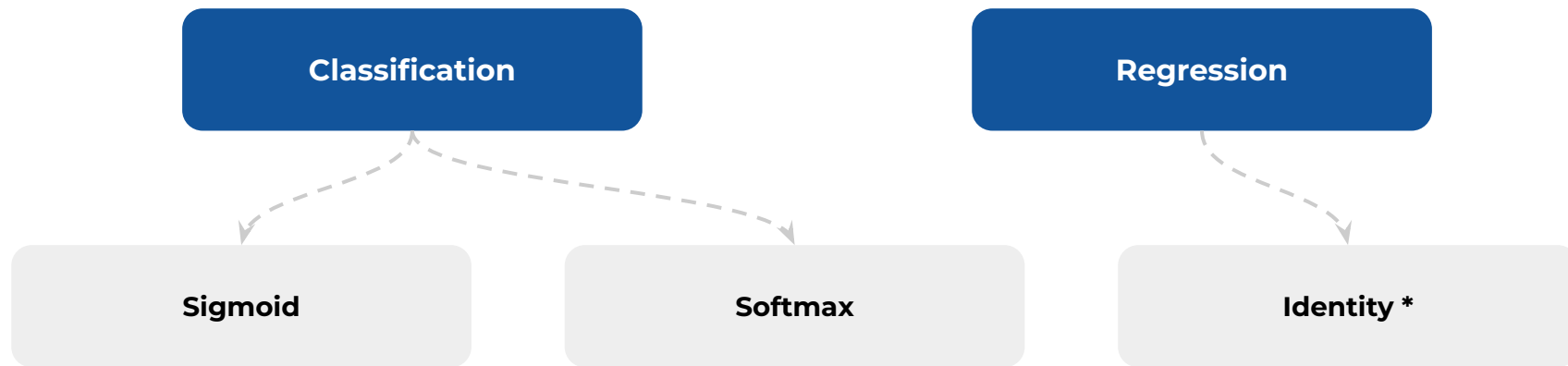
Leaky ReLU

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Activation Functions for Output Layer



$$f(z) = \frac{1}{1 + e^{-z}} \quad f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad \text{for } i = 1, 2, \dots, K \quad f(z) = z$$

\* This means no activation is applied in the output layer

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following statements best describes the role of a loss function in training a neural network?

A

The loss function adjusts the learning rate during training

B

The loss function measures the accuracy of the model's predictions compared to the true labels

C

The loss function initializes the weights and biases of the neural network

D

The loss function measures the accuracy of the model

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following statements best describes the role of a loss function in training a neural network?

A

The loss function adjusts the learning rate during training

B

The loss function measures the accuracy of the model's predictions compared to the true labels

C

The loss function initializes the weights and biases of the neural network

D

The loss function measures the accuracy of the model

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Loss Functions

In training a neural network, the loss function plays a critical role in evaluating how well the model is performing by measuring the difference between the predicted output of the model and the actual target output (true labels).

## Regression

### Mean Squared Error

$$\sum_{i=1}^D (x_i - y_i)^2$$

## Classification

### Cross-Entropy Loss

$$-\sum_{c=1}^M y_{o,c} \log(p_{o,c})$$

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following describes the learning rate in gradient descent?

A

The speed at which the model learns

B

The magnitude of the gradient

C

The step size for each iteration

D

The number of iterations required for convergence

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

Which of the following describes the learning rate in gradient descent?

A

The speed at which the model learns

B

The magnitude of the gradient

C

The step size for each iteration

D

The number of iterations required for convergence

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

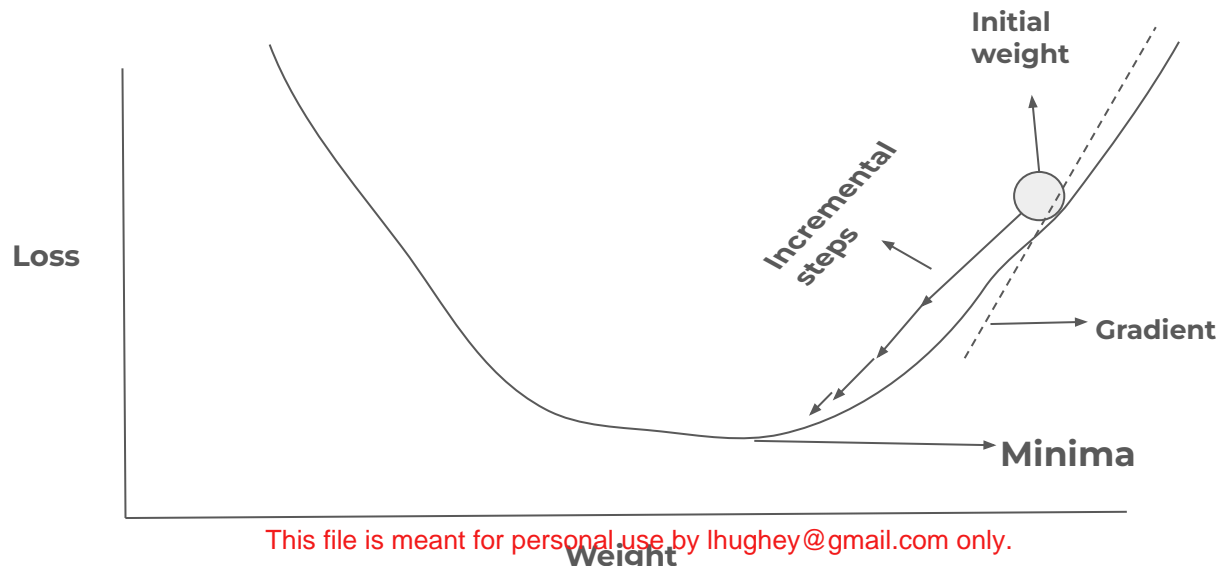
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.



# Gradient Descent

An **optimization algorithm** used to minimize the loss function

**Adjusts model parameters iteratively** in the direction of steepest descent of the gradient



This file is meant for personal use by lhughey@gmail.com only.

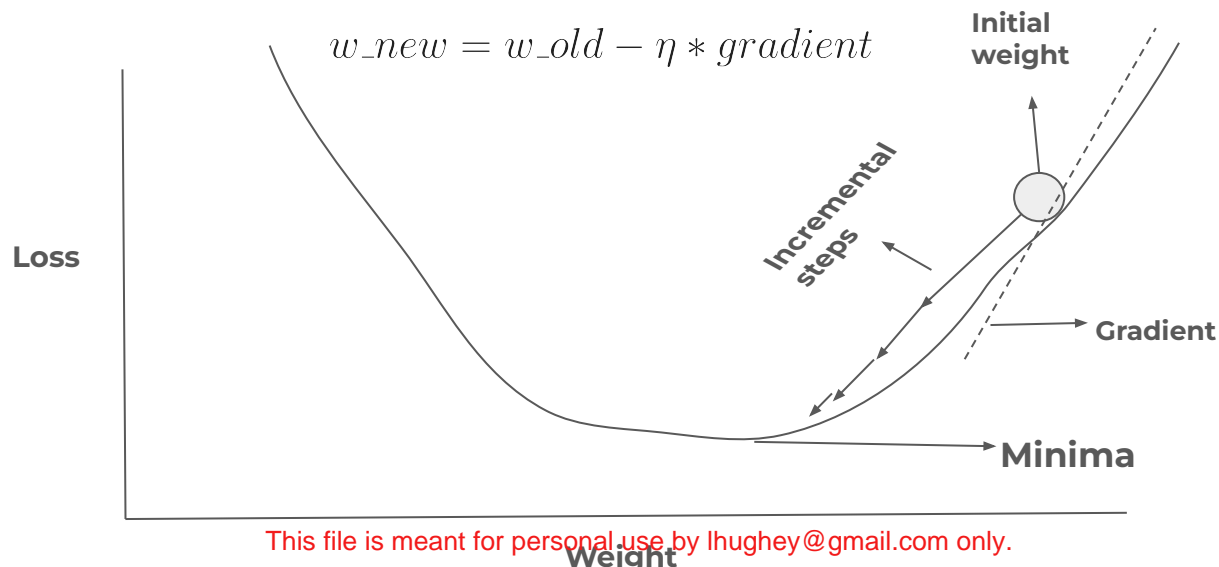
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Learning Rate

How big a step to take?  $\Rightarrow$  Decided by the **learning rate**

Smaller the learning rate, smaller the step



This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

During backpropagation, which of the following parameters of the model gets updated?

A

Input layer parameters

B

Output layer parameters

C

Hidden layer parameters

D

All the model parameters

This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Neural Networks Quiz

During backpropagation, which of the following parameters of the model gets updated?

A

Input layer parameters

B

Output layer parameters

C

Hidden layer parameters

D

All the model parameters

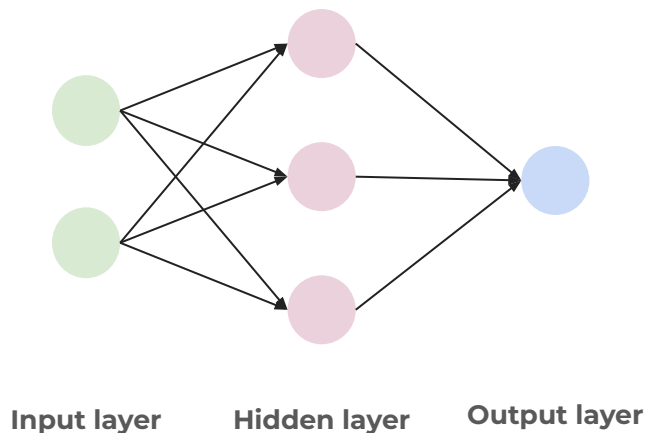
This file is meant for personal use by lhughey@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

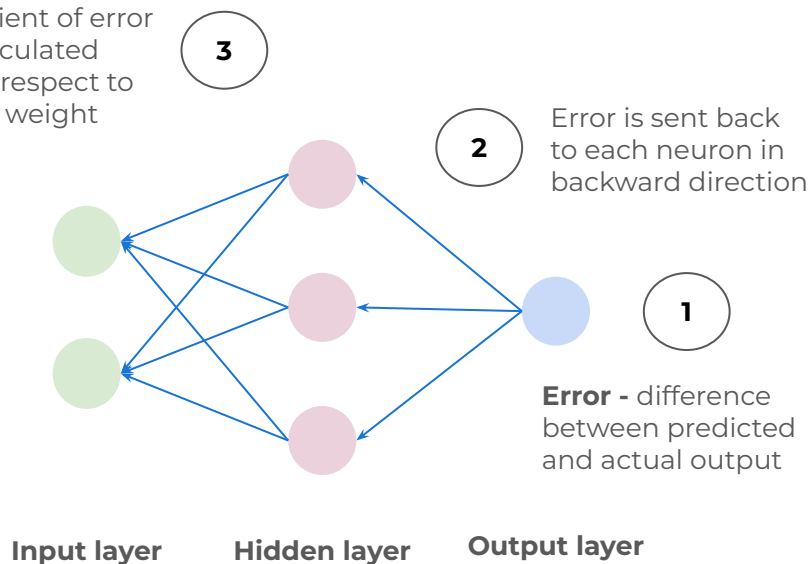
# Backpropagation

An **algorithm** used to efficiently compute gradients of the loss function with respect to model parameters



**Forward Propagation**

Gradient of error  
is calculated  
with respect to  
each weight



**Backpropagation**



# Happy Learning !

