1.

Examples of Activation Functions

"No activation function"

$$a_2^{[1]} = g(\overbrace{\vec{w}_2^{[1]} \cdot \vec{x} + b_2^{[1]}}^{z})$$

Linear activation function

Sigmoid

ReLU *Rectified Linear Unit*

$g(z) = z$

$g(z) = \frac{1}{1+e^{-z}}$

$g(z) = max(0, z)$
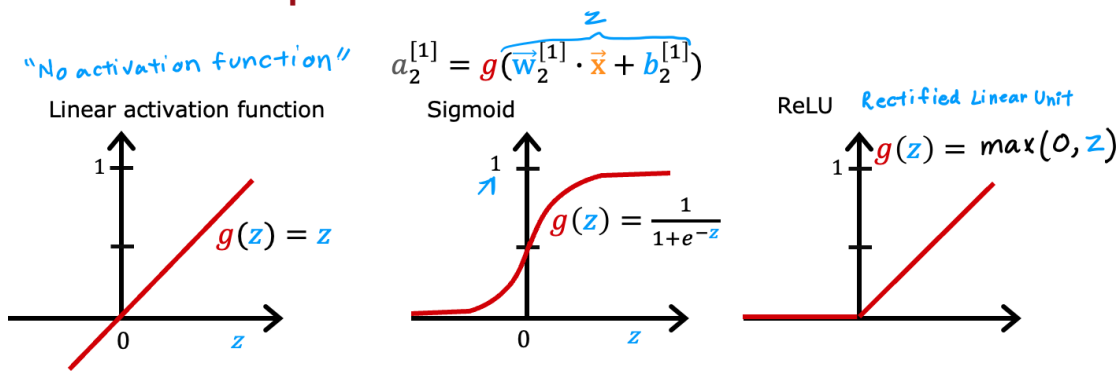
Which of the following activation functions is the most common choice for the hidden layers of a neural network?

- ○ Sigmoid
- ● ReLU (rectified linear unit)
- ○ Linear
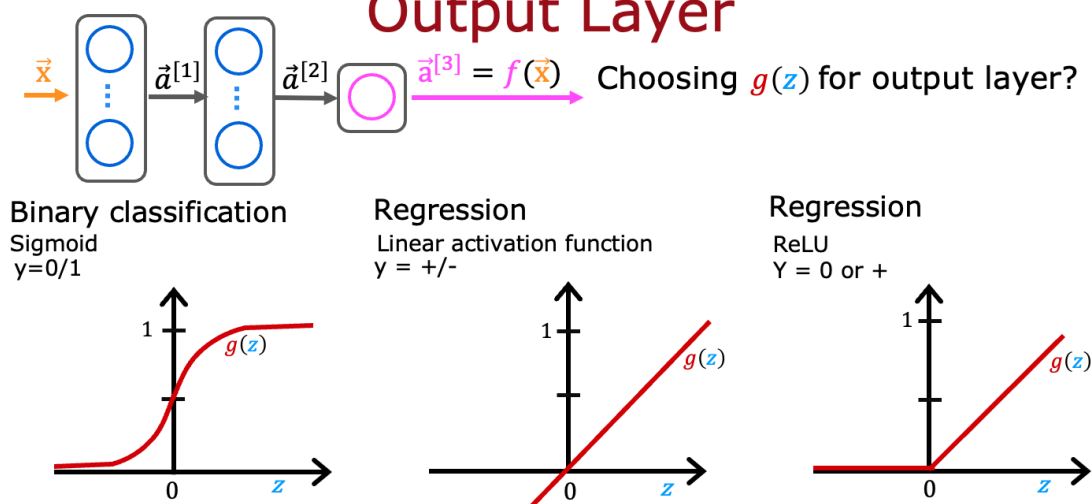- ○ Most hidden layers do not use any activation function

✓ **Correcto**

Yes! A ReLU is most often used because it is faster to train compared to the sigmoid. This is because the ReLU is only flat on one side (the left side) whereas the sigmoid goes flat (horizontal, slope approaching zero) on both sides of the curve.

2.

Output Layer

$\vec{x}$ → $\vec{a}^{[1]}$ → $\vec{a}^{[2]}$ → $\vec{a}^{[3]} = f(\vec{x})$ Choosing $g(z)$ for output layer?

Binary classification
Sigmoid
y=0/1

$g(z)$

Regression
Linear activation function
y = +/-

$g(z)$

Regression
ReLU
Y = 0 or +

$g(z)$

For the task of predicting housing prices, which activation functions could you choose for the output layer? Choose the 2 options that apply.

☑ ReLU

⊘ **Correcto**
Yes! ReLU outputs values 0 or greater, and housing prices are positive values.

☑ linear

⊘ **Correcto**
Yes! A linear activation function can be used for a regression task where the output can be both negative and positive, but it's also possible to use it for a task where the output is 0 or greater (like with house prices).

☐ Sigmoid

3. True/False? A neural network with many layers but no activation function (in the hidden layers) is not effective; that's why we should instead use the linear activation function in every hidden layer.

1 / 1 punto

🔘 False

○ True

⊘ **Correcto**
Yes! A neural network with many layers but no activation function is not effective. A linear activation is the same as "no activation function".