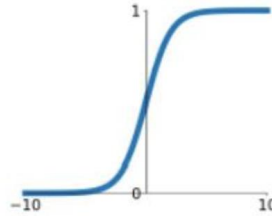


5. 简述神经网络需要激活函数的原因，并比较 sigmoid、tanh、ReLU 函数分别作为激活函数的效果有何区别。

- Sigmoid 函数：

**Sigmoid**

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



形状：

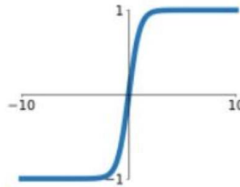
范围：(0, 1)。

特点：平滑且连续，但容易引起梯度消失问题，且输出不是以 0 为中心的。主要用于输出层，在二分类问题中尤为常见。

- Tanh 函数：

**tanh**

$$\tanh(x)$$



形状：

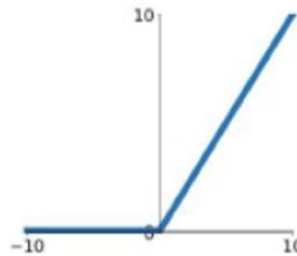
范围：(-1, 1)。

特点：输出以 0 为中心，可以缓解梯度消失的问题，但在极端值处仍然可能发生。隐藏层中比 sigmoid 更常见。

- ReLU 函数：

**ReLU**

$$\max(0, x)$$



形状：

范围：[0, ∞)。

特点：计算简单，加速神经网络的收敛，减轻了梯度消失问题。但有可能导致神经元“死亡”（在训练过程中某些神经元可能永远不被激活）。在现代神经网络架构中极为普遍，尤其是在隐藏层中。