

1. Which of the following is the most commonly used activation function in deep neural networks?

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

2. What is the main advantage of the ReLU activation function?

- A) It prevents the vanishing gradient problem
- B) It is computationally expensive
- C) It outputs values between 0 and 1
- D) It works only for ~~binary~~ classification

$0 \rightarrow +\infty$

3. Which of the following activation functions is commonly used for binary classification?

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



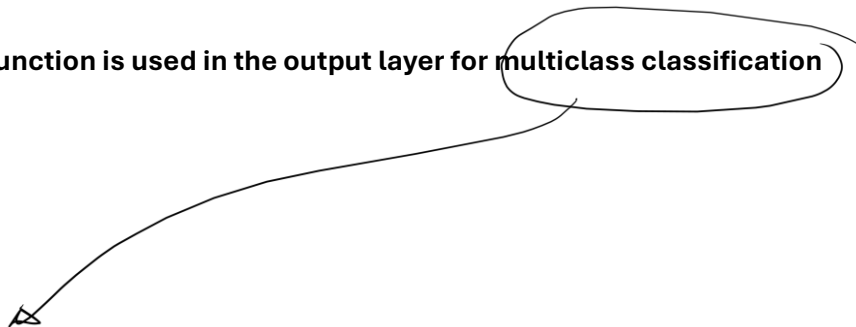
$0 \rightarrow 1$
 $> 0.5 \rightarrow 1$
 $< 0.5 \rightarrow 0$

4. The tanh activation function outputs values in the range of:

- A) $[0, 1]$
- B) $[-1, 1]$
- C) $[0, \infty)$
- D) $(-\infty, \infty)$

5. Which activation function is used in the output layer for multiclass classification problems?

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



6. The output of a ReLU activation function is:

- (A) Zero for negative inputs, and the same value for positive inputs
- B) A constant for all inputs
- C) Between 0 and 1 for all inputs
- D) The square of the input

7. Which activation function is most likely to suffer from the vanishing gradient problem?

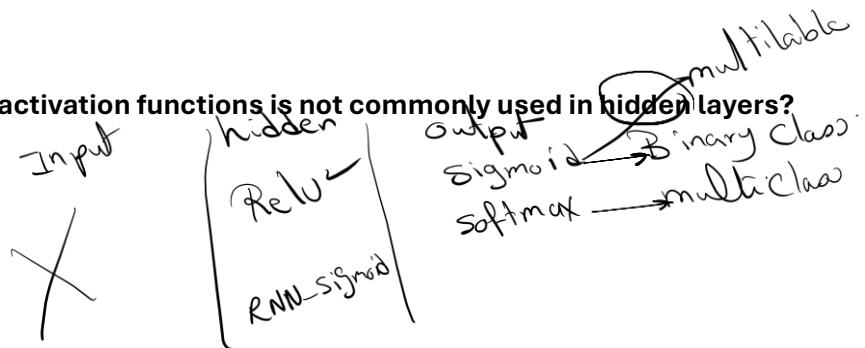
- A) ReLU
- B) Sigmoid
- C) Softmax
- D) Leaky ReLU

8. Which activation function is considered a variant of ReLU and allows small negative values?

- A) ELU
- B) Sigmoid
- (C) Leaky ReLU
- D) Softmax

9. Which of the following activation functions is not commonly used in hidden layers?

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



10. Which activation function is often used in generative models like GANs?

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

11. What is the range of the output for the Sigmoid activation function?

- A) $[0, 1]$ ✓
- B) $[-1, 1]$
- C) $(-\infty, \infty)$
- D) $[0, \infty)$

Sigmoid — $[0, 1]$
Tanh — $[-1, 1]$
ReLU — $[0, \infty)$
ReLU6 — $[0, 6]$

12. Which activation function is best suited for the output layer in a binary classification task?

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

13. Which activation function is defined as the maximum of 0 and the input?

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

$\max(0, z)$

14. Which of the following is a drawback of the sigmoid activation function?

- A) It is computationally expensive
- B) It suffers from the vanishing gradient problem
- C) It produces outputs that are not differentiable
- D) It is unsuitable for multi-class classification

15. The output of the Softmax activation function:

- A) Outputs values between 0 and 1, but the sum of all outputs is not constrained
- B) Outputs values between 0 and 1, and the sum of all outputs is 1
- C) Can be used in hidden layers
- D) Is not differentiable