Neural Network Review Session

First, you work on these questions by yourself.
Second, you discuss these questions with your peers in a group (2-3).
Third, reflect and revise your answers.

1. What is the best conceptual way to understand a neural network according to the lecture?
    A. As a biological device
    B. As an electronic device
    C. As a mathematical function
    D. As a network diagram
2. Which of the following domains commonly uses neural networks?
    A. Gaming
    B. Speech recognition
    C. Computer vision
    D. All of the above
3. What is the main advantage of neural networks once they are trained?
    A. They require less memory
    B. They predict very fast
    C. They need more labeled data
    D. They avoid overfitting automatically
4. What limitation of perceptrons motivated the development of feedforward neural networks?
    A. Unable to classify nonlinear functions
    B. Too slow to train
    C. High computational cost
    D. Lack of interpretability
5. In gradient descent, what does the term η represent?
    A. Momentum
    B. Step size or learning rate
    C. Gradient direction
    D. Error magnitude

6. For logistic regression, what is the activation function used in the prediction of y?
   A. tanh
   B. ReLU
   C. Sigmoid
   D. Maxout

7. What is the derivative of the sigmoid function $\sigma(z)$?
   A. $\sigma'(z) = 1 - \sigma(z)$
   B. $\sigma'(z) = z(1 - z)$
   C. $\sigma'(z) = \sigma(z)(1 - \sigma(z))$
   D. $\sigma'(z) = e^{-z}$

8. In a one-hidden-layer network, which of the following represents the forward computation correctly?
   A. $h = Wx + b$, $\hat{y} = g(h)$
   B. $z^{[1]} = W^{[1]}x + b^{[1]}$, $h^{[1]} = g(z^{[1]})$, $\hat{y} = g(W^{[2]}h^{[1]} + b^{[2]})$
   C. $\hat{y} = W^{[2]}h^{[1]}$ only
   D. $h^{[1]} = x + b$, $\hat{y} = g(W^{[2]})$

9. What role do hidden layers play in neural networks?
   A. They reduce overfitting
   B. They provide non-linear feature transformations
   C. They act as memory buffers
   D. They are only used for dimensionality reduction

10. List three common activation functions discussed in the slides and describe one property of each.