1. Answer: **B**

2. Answer: **B**

3. Answer: **A, B**

4. Answer: **A**

5.

(1) Use mean squared error as loss function:

$$MSE = \frac{1}{P}\sum_{i=1}^{P}(\tilde{Y}_i - Y_i)^2$$

(2) Use **Early Stopping** to prevent converging to a trivial and useless solution.

6.

(1)

(2) For this loss function, **w1 = w2 = 0** is a saddle point.

7. **Yes.** It is possible.

First, training another neural network on the same task.

Then, construct an adversarial example to this network.

Thus, This adversarial example is also an adversarial example of the encrypted network.

8. **No,** this approach won’t help.

Because the model is underfitting (high bias). Adding more data can reduce variance but not bias.

To solve this problem, we can try to add more layers and parameters.

9. If the training error is low but dev error is high, then it is overfitting.

10. Benefit of convolutional layers:

1. The amount of parameters is reduced.
2. CNN is invariant to translation (Translation invariance)