Your grade: 90%

Your latest: **90%** • Your highest: **90%** • To pass you need at least 80%. We keep your highest score.

Next item

coach

Almost there! Let's review your answers and then try again.

Your work showed a strong understanding of neural networks and their components, but there is room for improvement in understanding the specifics of convolutional neural networks. Before retrying this assignment, we recommend focusing on this key area:

- **Convolutional Neural Networks**: In Question 16, your calculation of the output width of a pooling layer was incorrect. Try reviewing Lesson 6: Convolutional Neural Networks to refresh your understanding of how pooling layers work, including how to calculate the dimensions of their output.

Good job on the rest of the assignment!

Retry assignment

1. A feedforward neural network has an input layer, 5 hidden layers and an output layer. What is the **depth** of this neural network? 1 / 1 point

6

Correct

2. During training, the training data specifies the exact form of the hidden layers of a neural network. 1 / 1 point

- ○ True
- ◉ False

Correct

3. Implement the ReLU activation function using numpy by replacing **None** in the code bellow. 2 / 2 points

```
1  import numpy as np
2
3  def ReLU(x):
4
5      y = np.maximum(0, x)
6
7      return y
```

Run
Reset

Correct
Good job!

4. The main building blocks of a machine learning system are: (Check all that apply.) 1 / 1 point

- ☑ A loss function

Correct

- ☐ Output Layers
- ☑ A Model

Correct

- ☐ Hidden layers
- ☑ An Optimization Procedure

Correct

5. Which output unit/loss function pair is usually used for regression tasks that use neural networks? 1 / 1 point

- ◉ Linear output units with Mean Squared Error Loss
- ○ Softmax output units with Cross-Entropy Loss
- ○ Sigmoid output units with Mean Squared Error Loss
- ○ Linear output units with Cross-Entropy Loss

Correct

6. The softmax output layer with cross-entropy loss is used to model the mean of a Gaussian distribution. 1 / 1 point

- ○ True
- ◉ False

Correct

7. Which of the following might be used as a stopping condition for gradient descent. (Check all that apply.) 1 / 1 point

- [ ] The value of the training loss
- [x] The number of iterations

Correct

- [x] The number of epochs

Correct

- [x] The magnitude of the change in parameter values

Correct

- [x] The magnitude of change in loss function value

Correct

8. How are neural network **bias** parameters usually initialized at the beginning of training? 1 / 1 point

- (•) Initialized to 0.
- ( ) Initialized to -1.
- ( ) Initialized to samples from a standard normal distribution.
- ( ) Initialized to samples from a standard uniform distribution.

Correct

9. Using all samples to estimate the gradient of the loss function with respect to the parameter results in less than linear return in accuracy of this estimate. 1 / 1 point

- (•) True
- ( ) False

Correct

10. You are working on a self-driving car project and want to train a neural network to perform traffic sign classification. You collect images with corresponding traffic sign labels, and want to determine the number of frames you will use for training. Given that you have around **one million** images with labels, what training/validation/testing data split would you use? 1 / 1 point

- ( ) 60% training, 20% validation, 20% testing.
- ( ) 20% training, 40% validation, 40% testing.
- (•) 96% training, 2% validation, 2% testing.
- ( ) 100% training, 0% validation, 0% testing.

Correct

11. You finish training your traffic sign classifier, and want to evaluate its performance. You compute the classification accuracy on the training, validation, and testing data splits and get the following results: 2 / 2 points

Data Split	Training	Validation	Testing
Accuracy	90%	88%	87%

You know that a human has an accuracy of around 98% on the traffic sign classification task. What are things you might try to achieve human level performance? (Check all that apply.)

- [x] Add more layers to your neural network.

Correct

- [ ] Add regularization to your neural network.
- [ ] Collect more training data.
- [x] Train your neural network longer.

Correct

12. When a neural network overfits the training data, the generalization gap is usually very small. 1 / 1 point

- ( ) True
- (•) False

Correct

13. Which of the following strategies are used for regularization in neural networks? (Check all that apply.) 1 / 1 point

- [ ] Increasing the number of parameters in the neural network architecture
- [ ] Training the neural network longer
- [x] Norm Penalties

Correct

- [x] Early Stopping

Correct

- [x] Dropout

Correct

14. Dropout significantly limit the type of neural network models that can be used, and hence is usually used for specific architectures. 1 / 1 point

- ( ) True
- (•) False

Correct

15. The name convolutional neural networks comes from the fact that these neural networks use a **convolution operation** instead of general matrix multiplication. 1 / 1 point

- ( ) True
- (•) False

Correct

16. The input to a pooling layer has a **width, height and depth** of 224x224x3 respectively. The pooling layer has the following properties: 0 / 2 points

- **Kernel shape:** 4x4
- **Stride**: 2

What is the width of the output of this pooling layer?

110

Incorrect
This answer is incorrect. Please refer back to Lesson 6 in this module to review this material.

17. Using convolutions might reduce overfitting, as the number of parameters in convolutional layers is **less** than the number of parameters in fully connected layers. 1 / 1 point

- (•) True
- ( ) False

Correct