Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE
100%

Feed-Forward Neural Networks

LATEST SUBMISSION GRADE

100%

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 = x*(x>0)
    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 Model

> ✓ **Correct**
> Correct!

☑ A loss function

> ✓ **Correct**
> Correct!

☑ An Optimization Procedure

> ✓ **Correct**
> Correct!

☐ Output Layers

☐ Hidden layers

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

○ Sigmoid output units with Mean Squared Error Loss

○ Softmax output units with Cross-Entropy Loss

◉ Linear output units with Mean Squared Error Loss

○ Linear output units with Cross-Entropy Loss

> ✓ **Correct**
> Correct!

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

    ⚪ True

    ⦿ False

    ✓ **Correct**
      Correct!

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

    ☑ The magnitude of the change in parameter values

    ✓ **Correct**
      Correct!

    ☑ The number of iterations or epochs

    ✓ **Correct**
      Correct!

    ☐ The value of the training loss

    ☑ The magnitude of change in loss function value

    ✓ **Correct**
      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**

    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**

  ○ 100% training, 0% validation, 0% testing.

  ○ 20% training, 40% validation, 40% testing.

  ◉ 96% training, 2% validation, 2% testing.

  ○ 60% training, 20% validation, 20% testing.

    ✓ **Correct**

    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	70%	68%	67%

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.)

  ☐ Collect more training data.

☑ Train your neural network longer.

✓ **Correct**
Correct!

☐ Add regularization to your neural network.

☑ Add more layers to your neural network.

✓ **Correct**
Correct!

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

○ True

◉ False

✓ **Correct**
Correct!

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

☑ Early Stopping

✓ **Correct**
Correct!

☐ Increasing the number of parameters in the neural network architecture

☑ Dropout

✓ **Correct**
Correct!

☑ Norm Penalties

✓ **Correct**
Correct!

☐ Training the neural network longer

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**
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**
Correct!

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

2 / 2 points

- **Kernel shape:** 2x2
- **Stride**: 2

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

112

✓ **Correct**
Correct!

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**
Correct!