1.

Question 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.

Question 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.

Question 3

Implement the ReLU activation function using numpy by replacing **None** in the code bellow.

2 / 2 points

```
1
2
3
4
5
6
7
```

```
import numpy as np

def ReLU(x):

  y = x*(x>=0)

  return y
```

Run

Reset

Correct
Good job!
4.

Question 4

The main building blocks of a machine learning system are: (Check all that apply.)
0 / 1 point

Hidden layers

A Model

An Optimization Procedure

Correct
Correct!

Output Layers

A loss function

Correct
Correct!
You didn't select all the correct answers
5.

Question 5

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

Softmax output units with Cross-Entropy Loss

Linear output units with Mean Squared Error Loss

Sigmoid output units with Mean Squared Error Loss

Linear output units with Cross-Entropy Loss

Correct
Correct!
6.

Question 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
Correct!

7.

Question 7

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

1 / 1 point

The magnitude of the change in parameter values

Correct
Correct!

The number of iterations or epochs

Correct
Correct!

The magnitude of change in loss function value

Correct
Correct!

The value of the training loss

8.

Question 8

How are neural network **bias** parameters usually initialized at the beginning of training?

1 / 1 point

Initialized to samples from a standard normal distribution.

Initialized to 0.

Initialized to -1.

Initialized to samples from a standard uniform distribution.

Correct

9.

Question 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.

Question 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.

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

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

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

Correct
Correct!
11.
Question 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:

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.)
2 / 2 points

Collect more training data.

Add more layers to your neural network.

Correct
Correct!

Train your neural network longer.

Correct
Correct!

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

True

False

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

Training the neural network longer

Early Stopping

Correct
Correct!

Dropout

Correct
Correct!

Increasing the number of parameters in the neural network architecture

Norm Penalties

Correct
Correct!
14.
Question 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.
Question 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.
Question 16
The input to a pooling layer has a **width, height and depth** of 224x224x3 respectively. The pooling layer has the following properties:

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

What is the width of the output of this pooling layer?
2 / 2 points

112

Correct
Correct!
17.
Question 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!