Course 2 – Week 2 – Quiz

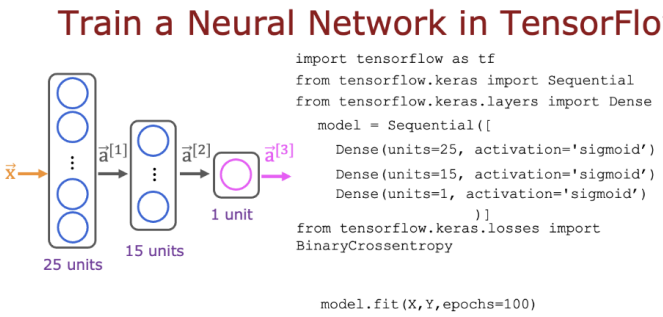
Quizz 1 – Neural Network Training

Congratulations! You passed!

Grade received 100% | Latest Submission Grade 100% | To pass 80% or higher

Go to next item

1. 1 / 1 point



Here is some code that you saw in the lecture:

```
model.compile(loss=BinaryCrossentropy())
```

For which type of task would you use the binary cross entropy loss function?

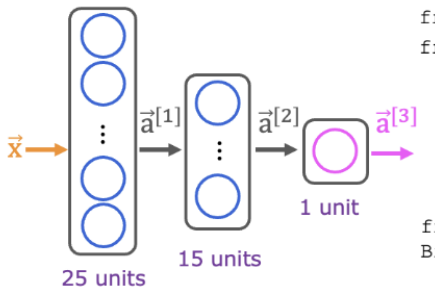
- (●) binary classification (classification with exactly 2 classes)
- ( ) regression tasks (tasks that predict a number)
- ( ) BinaryCrossentropy() should not be used for any task.
- ( ) A classification task that has 3 or more classes (categories)

Correct
Yes! Binary cross entropy, which we've also referred to as logistic loss, is used for classifying between two classes (two categories).

2.

1 / 1 pc

Train a Neural Network in TensorFlow




```
import tensorflow as tf
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense
   model = Sequential([
     Dense(units=25, activation='sigmoid')
     Dense(units=15, activation='sigmoid')
     Dense(units=1, activation='sigmoid')
                          )]
from tensorflow.keras.losses import
BinaryCrossentropy


   model.fit(X,Y,epochs=100)
```

Here is code that you saw in the lecture:

```
model = Sequential([
Dense(units=25, activation='sigmoid'),
Dense(units=15, activation='sigmoid'),
Dense(units=1, activation='sigmoid')
])

model.compile(loss=BinaryCrossentropy())
model.fit(X,y,epochs=100)
```

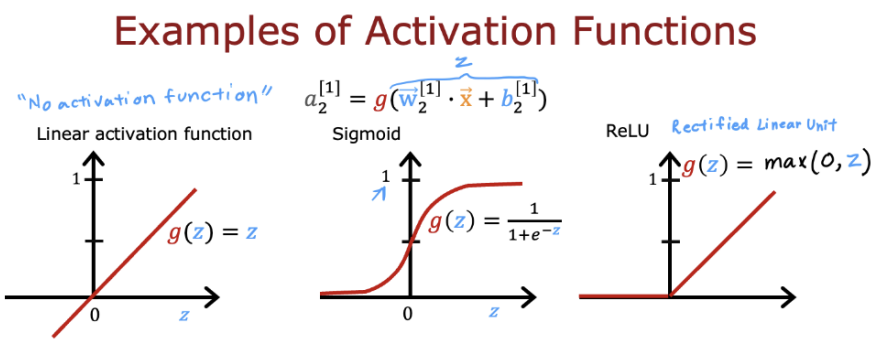
Which line of code updates the network parameters in order to reduce the cost?

- ○ model.compile(loss=BinaryCrossentropy())
- ○ model = Sequential([...])
- ◉ model.fit(X,y,epochs=100)
- ○ None of the above -- this code does not update the network parameters.

Correct
Yes! The third step of model training is to train the model on data in order to minimize the loss (and the cost)

Quiz 2 – Activation Functions

1. 1 / 1 point

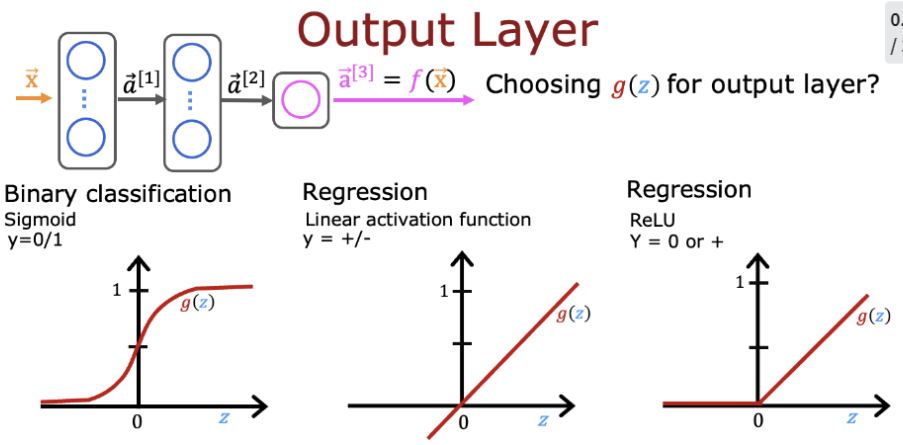


Which of the following activation functions is the most common choice for the hidden layers of a neural network?

- ○ Sigmoid
- ◉ ReLU (rectified linear unit)
- ○ Most hidden layers do not use any activation function
- ○ Linear

✓ **Correct**
Yes! A ReLU is most often used because it is faster to train compared to the sigmoid. This is because the ReLU is only flat on one side (the left side) whereas the sigmoid goes flat (horizontal, slope approaching zero) on both sides of the curve.

2. 0.6666666666666666 / 1 point



For the task of predicting housing prices, which activation functions could you choose for the output layer? Choose the 2 options that apply.

- [ ] Sigmoid
- [x] ReLU

Correct
Yes! ReLU outputs values 0 or greater, and housing prices are positive values.

- [ ] linear

You didn't select all the correct answers

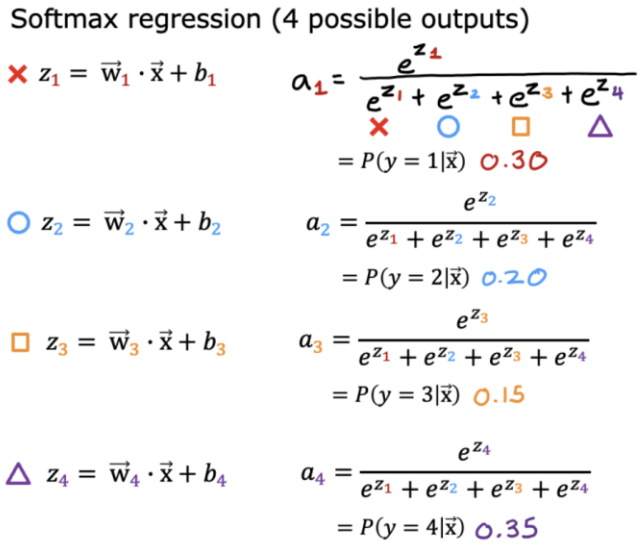
3. True/False? A neural network with many layers but no activation function (in the hidden layers) is not effective; that's why we should instead use the linear activation function in every hidden layer. 1 / 1 point

- ( ) True
- (•) False

Correct
Yes! A neural network with many layers but no activation function is not effective. A linear activation is the same as "no activation function".

Quiz 3 – Multiclass Classification

1. 1 / 1 point



For a multiclass classification task that has 4 possible outputs, the sum of all the activations adds up to 1. For a multiclass classification task that has 3 possible outputs, the sum of all the activations should add up to ....

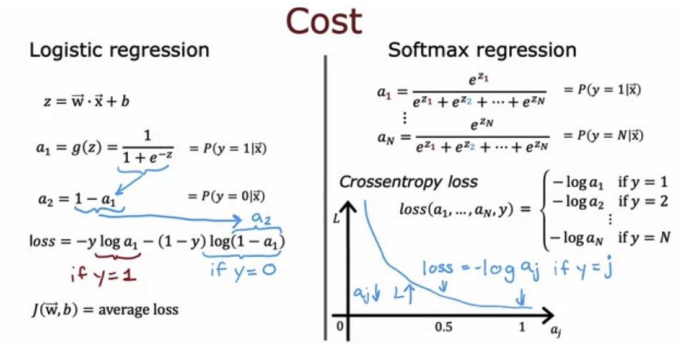
- ○ More than 1
- ◉ 1
- ○ It will vary, depending on the input x.
- ○ Less than 1

✓ Correct

Yes! The sum of all the softmax activations should add up to 1. One way to see this is that if $e^{z_1} = 10, e^{z_2} = 20, e^{z_3} = 30$, then the sum of $a_1 + a_2 + a_3$ is equal to $\frac{e^{z_1}+e^{z_2}+e^{z_3}}{e^{z_1}+e^{z_2}+e^{z_3}}$ which is 1.

2.

1 / 1 point



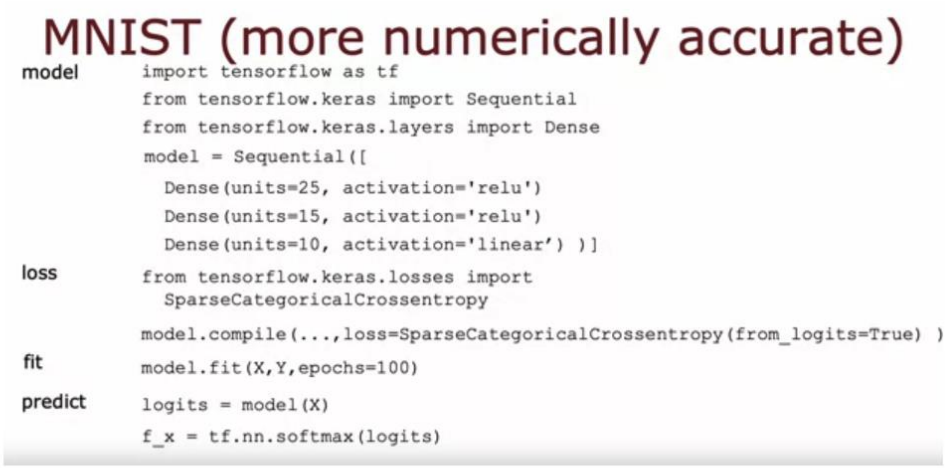
For multiclass classification, the cross entropy loss is used for training the model. If there are 4 possible classes for the output, and for a particular training example, the true class of the example is class 3 (y=3), then what does the cross entropy loss simplify to? [Hint: This loss should get smaller when a_3 gets larger.]

- ○ z_3/(z_1+z_2+z_3+z_4)
- ○ z_3
- ○ $\frac{-log(a_1)+-log(a_2)+-log(a_3)+-log(a_4)}{4}$
- ◉ $-log(a_3)$

✓ **Correct**
Correct. When the true label is 3, then the cross entropy loss for that training example is just the negative of the log of the activation for the third neuron of the softmax. All other terms of the cross entropy loss equation ($-log(a_1), -log(a_2), and -log(a_4)$) are ignored

3.

1 / 1 point



For multiclass classification, the recommended way to implement softmax regression is to set from_logits=True in the loss function, and also to define the model's output layer with…

- ○ a 'softmax' activation
- ◉ a 'linear' activation

✓ Correct
Yes! Set the output as linear, because the loss function handles the calculation of the softmax with a more numerically stable method.

Quiz 4 – Additional Neural Network Concepts

1.

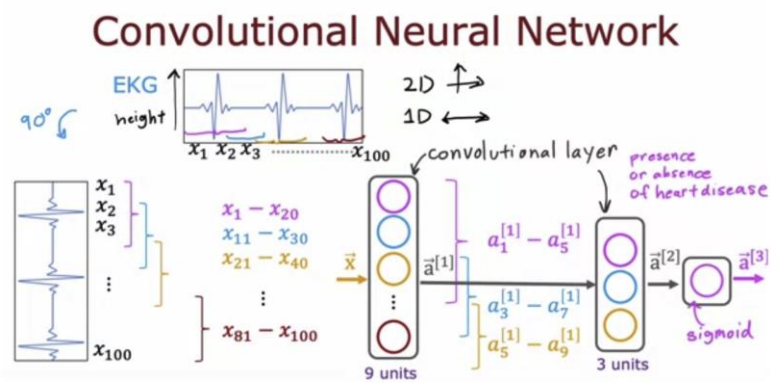
1 / 1 point

The Adam optimizer is the recommended optimizer for finding the optimal parameters of the model. How do you use the Adam optimizer in TensorFlow?

- (●) When calling model.compile, set optimizer=tf.keras.optimizers.Adam(learning_rate=1e-3).
- ( ) The call to model.compile() uses the Adam optimizer by default
- ( ) The Adam optimizer works only with Softmax outputs. So if a neural network has a Softmax output layer, TensorFlow will automatically pick the Adam optimizer.
- ( ) The call to model.compile() will automatically pick the best optimizer, whether it is gradient descent, Adam or something else. So there's no need to pick an optimizer manually.

✓ Correct
Correct. Set the optimizer to Adam.

2.



1 / 1 point

The lecture covered a different layer type where each single neuron of the layer does not look at all the values of the input vector that is fed into that layer. What is this name of the layer type discussed in lecture?

- ( ) 1D layer or 2D layer (depending on the input dimension)
- (●) convolutional layer
- ( ) Image layer
- ( ) A fully connected layer

✓ Correct
Correct. For a convolutional layer, each neuron takes as input a subset of the vector that is fed into that layer.