Q4

Due 12 Feb at 5:30 **Points** 180 **Questions** 12 **Available** until 12 Feb at 5:30 **Time limit** 45 Minutes

Submission details:

Time:	14 minutes
Current score:	141 out of 180
Kept score:	141 out of 180

Instructions

1. You are 45 minutes to take the quiz
2. Keep a calculator handy
3. Make sure you have read and understood the things in "Italics" in the notes shared.
4. Once you start the quiz, you cannot go back and re-attempt it
5. You will not find answers online, so please make sure you are ready for the quiz
6. For Multiple Answer Questions, ALL the answers must be correct to score any point

This quiz was locked 12 Feb at 5:30.

Attempt history

	Attempt	Time	Score
LATEST	Attempt 1	14 minutes	141 out of 180

Score for this quiz: **141** out of 180
Submitted 12 Feb at 4:07
This attempt took 14 minutes.

Question 1

10 / 10 pts

When you read "Those circles are "temporary" values that will be stored. Once you train the model, *lines are what all matters!!* in the notes, what is the meaning of temporary?

- They are temporary because we can use sources as well to represent the weights
- Circles represent weights and since they are changing, circles represent temporary values

Correct! ✓ Circles represent the calculated neuron value, or the channel's pixel value. These values are temporary as they will change with every image and are dumped out of memory after every inference.

Correct! ✓ Circles represents the values calculated after multiplying the input with the weights (represented by the lines). Since Inputs will change, multiplying the Inputs with weights will also change. Hence they are temporary

Question 2

5 / 10 pts

When you read "Those circles are "temporary" values that will be stored. Once you train the model, *lines are what all matters!!* in the notes, what is the meaning of "lines are what all matter"?

- Lines are all matter because they are the routes through which the input values are transferred to the next layer as it is.

Correct! ✓ Lines are what matter, as they not only represent the weights which we want to train, they also represent how "dense" our connections are. More the lines, denser the network. "Denseness" has direct implication on the model type.

Correct! ✓ Lines represent the weights, and it for achieving correct weights we are training the model. Hence finally it is those lines which matter.

You Answered ✓ Lines are what matter because without those lines circles will fall.

Question 3

10 / 10 pts

When you read "Exactly, that's the point. " what was meant by it?

- That 1D pattern created by converting a 2D pattern has retained it's spatial information

Correct! ✓ That a 1D pattern created by converting 2D pattern has lost its spatial meaning.

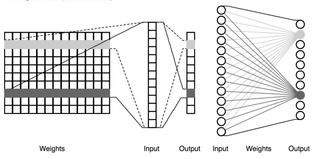
Converting 2D patterns into 1D patterns allows the network to keep spatial pattern, and that is why we need to convert 2D patterns into 1D patterns, especially when we are working on "vision" dnn.

Correct! ✓ Converting 2D pattern into a 1D pattern throws away the "spatial information". And without spatial information it wouldn't be ideal to train a "vision" dnn.

Question 4

25 / 25 pts

In the image shown below (don't consider biases):



Correct! ✓ Total weights used are 130

Correct! ✓ The output size is 10d

- The weight matrix is 10x13

Correct! ✓ If we connect all the input circles to the output circles (right part of the image), we will end up drawing 130 lines.

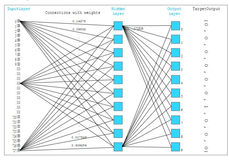
Correct! ✓ The input size is 13d

Correct! ✓ The weight matrix is 13x10

Question 5

15 / 15 pts

In the image below (don't consider biases):



- total 7370 weights are used

Correct! ✓ Target Output is shown as a One Hot Vector

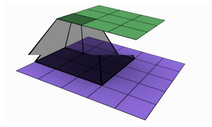
Correct! ✓ Hidden Layer has 100 weights

Correct! ✓ total 7260 weights are used

Question 6

0 / 20 pts

In this Image:



Correct answer If we draw lines to show the connections, we will end up drawing 91 lines

Correct! ✓ If we flatten both Input and output, we would need an FC layer with 225 weights

- If we flatten both input and output, we would need an FC layer with 45 weights

You Answered ✓ If we draw lines to show the connections, we will end up drawing 9 lines

Question 7

0 / 25 pts

In the Image below, the 3 blue boxes represent 3 FC (first two have same 4096 neurons) (don't consider biases):

You Answered ✓ total 123642856 parameters are used in the fc layers

Correct answer total 123633664 parameters are used in the fc layers

- total 104857448 parameters are used in the FC layers

Question 8

10 / 10 pts

It is a good idea to use ReLU as the activation function for the logits to softmax

Correct! ✓ No! Are you kidding! Never!

- Yes, always!

Question 9

10 / 10 pts

Why Softmax is not probability, but likelihood!

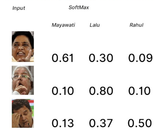
Correct! ✓ Because it is the measure of the features it has actually found!

Correct! ✓ Because everything which sums up to 1 is not probability.

Question 10

35 / 35 pts

Assume that we are using **Negative-Log_Likelihood**. Then in the Image below:



Total loss is

Correct! ✓ 1.41058

- 0.41261
- 3

Question 11

0 / 10 pts

In the BatchNormalization notes, you read "Indirectly you have sort of already used it!". What do you think it means?

- BN is built into PyTorch, so when we worked on Assignment 2, we were indirectly using it.
- When we train a model, weights get normalized during backpropagation, so we indirectly used it.

You Answered ✓ Since we used it Indirectly, we Indirectly used it!

Correct! ✓ When we applied formalization to our images, that was very similar to what we do in batch normalization

Question 12

20 / 20 pts

Select all which are true (context dropout):

- Since we drop weights when we use DropOut, after training we can delete the weights which were dropped.
- During DropOut always a fixed set of weights are dropped out.

Correct! ✓ If we actually have used dropout of 0.5 before the final layer, the training accuracy of a very well trained model will not cross 50% (assume it was holiday Nothinbdag problem).

Correct! ✓ It is not recommended to use Dropout before the last prediction layer

- We need to use large values of dropout, like 0.5~0.9

Correct! ✓ DropOut is applied only during training. During test/validation, it is automatically removed.

Correct! ✓ In DropOut, we need to divide the input to a layer by 2 if dropout of 0.5 was used while training it.

Quiz score: **141** out of 180
This quiz score has been manually adjusted by +1.0 points.

‹ Previous Next ›