











Session 1/2 Quiz

* **Due** Mar 7 at 5:29am
* **Points** 260
* **Questions** 15
* **Available** Mar 4 at 2pm - Mar 7 at 5:29am 3 days
* **Time Limit** 30 Minutes

Instructions

1. You have 30 minutes to finish the quiz.
2. Once you answer a question (or skip it), you cannot go back to it, so make sure before you submit an answer.
3. This is an open quiz, so you can search online for answers (you'd be wasting more time though :) ).
4. There are 15 questions in total, some will take less time, some will take more
5. There is negative marking as well.
6. The quiz can only be attempted once (so make sure you have good internet connectivity)
7. The quiz is only available till Thu Mar 7, 2019, 5:30 am

All the best

Attempt History

|  | **Attempt** | **Time** | **Score** |
| --- | --- | --- | --- |
| **LATEST** | [Attempt 1](https://canvas.instructure.com/courses/1546023/quizzes/3675518/history?version=1) | 30 minutes | 138.33 out of 260 |

Correct answers will be available on Mar 7 at 5:30am.

Score for this quiz: **138.33** out of 260

Submitted Mar 7 at 1:15am

This attempt took 30 minutes.

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Incorrect**Question 1

0 / 5 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_choice\_question97122941



Let us assume, we have 100 images in our dataset, and 20 total batches. How many times would be performing a forward prop before we perform 1 backprop?

6020



100



100

exact\_answernone6020

margin of error +/-

5227



20



20

exact\_answernone5227

margin of error +/-

7866



5



5

exact\_answernone7866

margin of error +/-

3880



1



1

exact\_answernone3880

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Incorrect**Question 2

0 / 5 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_choice\_question97122943



Let us assume, we have 100 images in our dataset, and 20 total batches. In one (1) epoch, how many total backprops would we perform?

4057



100



100

exact\_answernone4057

margin of error +/-

8598



20



20

exact\_answernone8598

margin of error +/-

6064



5



5

exact\_answernone6064

margin of error +/-

2851



1



1

exact\_answernone2851

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Incorrect**Question 3

0 / 10 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_choice\_question97122986



Let us assume the network below:

Layer 1 | input(244, 244, 3)  
Layer 2 | conv2D(32, 3, 3, act=ReLU)  
Layer 3 | conv2D(64, 3, 3, act=ReLU)  
Layer 4 | conv2D(128, 5, 5, act=ReLU)  
Layer 5 | conv2D(32, 1, 1, act=ReLU)  
Layer 6 | MaxPooling(2)  
Layer 7 | conv2D(64, 3, 3, act=ReLU)  
...

What is the global receptive field for each kernel/filter in the Layer 7?

5460



11x11



11x11

exact\_answernone5460

margin of error +/-

1878



18x18



18x18

exact\_answernone1878

margin of error +/-

4047



20x20



20x20

exact\_answernone4047

margin of error +/-

4327



16x16



16x16

exact\_answernone4327

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 4

6.67 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123066



Why do we add 1x1 layers? (select all which apply)

5590



To reduce the number of channels



To reduce the number of channels

Your Answer: **(You left this blank)**

exact\_answernone5590

margin of error +/-

4056



Combine a large number of channels into smaller relevant ones



Combine a large number of channels into smaller relevant ones

exact\_answernone4056

margin of error +/-

3136



To increase the number of channels



To increase the number of channels

Your Answer: **(You left this blank)**

exact\_answernone3136

margin of error +/-

7596



It is computationally much efficient as compared to a 3x3 to reduce the number of channels



It is computationally much efficient as compared to a 3x3 to reduce the number of channels

Your Answer: **(You left this blank)**

exact\_answernone7596

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

Question 5

20 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123067



Why do we apply MaxPooling?  (select all which apply)

970



To increase the effective receptive field



To increase the effective receptive field

Your Answer: **(You left this blank)**

exact\_answernone970

margin of error +/-

8552



To reduce the number of channels



To reduce the number of channels

exact\_answernone8552

margin of error +/-

1826



To reduce the number of layers required in a network



To reduce the number of layers required in a network

Your Answer: **(You left this blank)**

exact\_answernone1826

margin of error +/-

1287



To reduce the resolution of the layers



To reduce the resolution of the layers

Your Answer: **(You left this blank)**

exact\_answernone1287

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Incorrect**Question 6

0 / 10 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123072



How many layers should we add to a DNN to make an efficient network for object detection?

6738



As many as we can add, as more layers are always good



As many as we can add, as more layers are always good

exact\_answernone6738

margin of error +/-

2129



As many as required for us to reach the required receptive field



As many as required for us to reach the required receptive field

Your Answer: **(You left this blank)**

exact\_answernone2129

margin of error +/-

4779



Minimum 21 layers



Minimum 21 layers

exact\_answernone4779

margin of error +/-

2875



Does not matter



Does not matter

Your Answer: **(You left this blank)**

exact\_answernone2875

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 7

5 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123116



What among these things are not used anymore? (select all which apply)

9679



Adding as many layers as possible irrespective of the image size



Adding as many layers as possible irrespective of the image size

exact\_answernone9679

margin of error +/-

7834



Sigmoid as activation function in CNNs



Sigmoid as activation function in CNNs

Your Answer: **(You left this blank)**

exact\_answernone7834

margin of error +/-

1705



DropOuts



DropOuts

exact\_answernone1705

margin of error +/-

5110



Fully Connected Layers



Fully Connected Layers

exact\_answernone5110

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

Question 8

10 / 10 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123201



Assume this short DNN below:

Layer 0 | model = Sequential()  
Layer 1 | model.add(Convolution2D(34, 3, 3, activation='relu', input\_shape=(28,28,1)))  
Layer 2 | model.add(Convolution2D(32, 1, activation='relu'))  
Layer 3 | model.add(Convolution2D(32, 3, 3, activation='relu'))....

What is the **total number of parameters** we have added in the above 3 layers? (remember we have by default added biases).   
i.e. Convolution2D(34, 3, 3) has 3x3x34 + 34 = 340 parameters.

4130



724



724

exact\_answernone4130

margin of error +/-

8473



9652



9652

exact\_answernone8473

margin of error +/-

1704



10610



10610

exact\_answernone1704

margin of error +/-

4947



10708



10708

Your Answer: **(You left this blank)**

exact\_answernone4947

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

Question 9

20 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123208



Consider the following network:

input(200, 200, 1)  
conv2D(128, 5, 5)  
MaxPooling(2)  
conv2D(128, 11, 11)  
conv2D(32, 11, 11)  
conv2D(128, 1, 1)  
...

What all statements apply below?

6127



Bad network, we should have used 3x3 kernels for best performance



Bad network, we should have used 3x3 kernels for best performance

Your Answer: **(You left this blank)**

exact\_answernone6127

margin of error +/-

411



Bad network, we should not have used 1x1 to increase the number of channels



Bad network, we should not have used 1x1 to increase the number of channels

Your Answer: **(You left this blank)**

exact\_answernone411

margin of error +/-

2375



Bad network, we should not have performed MaxPooling so early



Bad network, we should not have performed MaxPooling so early

Your Answer: **(You left this blank)**

exact\_answernone2375

margin of error +/-

908



Nothing can be said about this network, as it depends on how long we train it.



Nothing can be said about this network, as it depends on how long we train it.

exact\_answernone908

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 10

15 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123216



(Please select all which apply)

When we select a batch from the dataset, we should make sure that

8244



it represents all the classes equally



it represents all the classes equally

exact\_answernone8244

margin of error +/-

6049



it is utilizing full GPU resources



it is utilizing full GPU resources

Your Answer: **(You left this blank)**

exact\_answernone6049

margin of error +/-

4704



it is never from the "test" dataset



it is never from the "test" dataset

Your Answer: **(You left this blank)**

exact\_answernone4704

margin of error +/-

5354



it is randomly shuffled



it is randomly shuffled

Your Answer: **(You left this blank)**

exact\_answernone5354

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 11

13.33 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123234



Why do we not use Fully Connected (FC) layers these days? (Select all which apply)

2309



FC layers looses all the spatial information especially required in vision domain



FC layers looses all the spatial information especially required in vision domain

Your Answer: **(You left this blank)**

exact\_answernone2309

margin of error +/-

1917



We need different kind of activation functions for FC layers which are not efficient



We need different kind of activation functions for FC layers which are not efficient

exact\_answernone1917

margin of error +/-

253



If we use FC we force our network to use only a specific size of input image, but we want a network which can process images of all sizes



If we use FC we force our network to use only a specific size of input image, but we want a network which can process images of all sizes

Your Answer: **(You left this blank)**

exact\_answernone253

margin of error +/-

330



FC layers add a lot more parameters than Convolution layers and hence are slower to process



FC layers add a lot more parameters than Convolution layers and hence are slower to process

exact\_answernone330

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 12

15 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123278



Adding a bias in a network is useful.

Select the ones which apply:

246



it is useful for simple, fully connected networks, where the problem we are solving is also linear/simple



it is useful for simple, fully connected networks, where the problem we are solving is also linear/simple

Your Answer: **(You left this blank)**

exact\_answernone246

margin of error +/-

3753



it is not useful, and people are today cautious about removing them



it is not useful, and people are today cautious about removing them

exact\_answernone3753

margin of error +/-

642



it is not helpful, but sticking with us because of it's historicity



it is not helpful, but sticking with us because of it's historicity

Your Answer: **(You left this blank)**

exact\_answernone642

margin of error +/-

9357



it is not useful, as in vision domain we have too many parameters, and 1 bias variable cannot help resolve non-linear models we have



it is not useful, as in vision domain we have too many parameters, and 1 bias variable cannot help resolve non-linear models we have

Your Answer: **(You left this blank)**

exact\_answernone9357

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Partial**Question 13

13.33 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123285



Select all which are true below:

We use squared loss because:

687



it always gives us positive loss/cost values



it always gives us positive loss/cost values

Your Answer: **(You left this blank)**

exact\_answernone687

margin of error +/-

6874



we want to punish larger values more (especially when our network has just been initialized)



we want to punish larger values more (especially when our network has just been initialized)

Your Answer: **(You left this blank)**

exact\_answernone6874

margin of error +/-

2159



it is a continuous function



it is a continuous function

exact\_answernone2159

margin of error +/-

1583



there are no other loss functions



there are no other loss functions

exact\_answernone1583

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

Question 14

20 / 20 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123321



Would you add MaxPooling just before final prediction layers? (Select all which apply)

2301



yes, because it helps reduce the total number of parameters



yes, because it helps reduce the total number of parameters

exact\_answernone2301

margin of error +/-

8501



no, because it will reduce the number of parameters



no, because it will reduce the number of parameters

exact\_answernone8501

margin of error +/-

1360



yes, as it helps reduce the resolution of the channels



yes, as it helps reduce the resolution of the channels

exact\_answernone1360

margin of error +/-

7494



no, because it will reduce the rich information in the last layers required for accurate predictions



no, because it will reduce the rich information in the last layers required for accurate predictions

Your Answer: **(You left this blank)**

exact\_answernone7494

margin of error +/-

*Move To... This element is a more accessible alternative to drag & drop reordering. Press Enter or Space to move this question.*

**Incorrect**Question 15

0 / 40 pts

[*Edit this QuestionDelete this Question*](https://canvas.instructure.com/courses/1546023/quizzes/3675518?headless=1)

0multiple\_answers\_question97123333



Select all which are true below:

5687



We always use 3x3 kernels (with the exception of 1x1 sometimes)



We always use 3x3 kernels (with the exception of 1x1 sometimes)

Your Answer: **(You left this blank)**

exact\_answernone5687

margin of error +/-

3831



Total number of kernels is dependent on the total number of input channels



Total number of kernels is dependent on the total number of input channels

exact\_answernone3831

margin of error +/-

8566



Our interception point (location where we add MaxPooling and 1x1 layers), are different for each dataset, and depends on the size of the image as well as the features we want to extract



Our interception point (location where we add MaxPooling and 1x1 layers), are different for each dataset, and depends on the size of the image as well as the features we want to extract

exact\_answernone8566

margin of error +/-

951



We use 1x1 to increase the total number of layers



We use 1x1 to increase the total number of layers

exact\_answernone951

margin of error +/-

7138



A model with a larger number of parameters will always have higher accuracy than a fewer number of parameters



A model with a larger number of parameters will always have higher accuracy than a fewer number of parameters

Your Answer: **(You left this blank)**

exact\_answernone7138

margin of error +/-

2477



If GPU resources allow, we would rather prefer to make our batch size equal to total number of images for backpropagation



If GPU resources allow, we would rather prefer to make our batch size equal to total number of images for backpropagation

Your Answer: **(You left this blank)**

exact\_answernone2477

margin of error +/-

4160



1x1 is mostly used to increase the number of channels



1x1 is mostly used to increase the number of channels

exact\_answernone4160

margin of error +/-

3930



Softmax does not improves the accuracy of the network, but just creates large separation in prediction values, and hence might be misleading



Softmax does not improves the accuracy of the network, but just creates large separation in prediction values, and hence might be misleading

exact\_answernone3930

margin of error +/-

7793



We always add 2x2 MaxPooling, else we would loose too much of the information



We always add 2x2 MaxPooling, else we would loose too much of the information

Your Answer: **(You left this blank)**

exact\_answernone7793

margin of error +/-

4407



A model with large number of kernels will always have higher accuracy than a fewer number of kernels



A model with large number of kernels will always have higher accuracy than a fewer number of kernels

Your Answer: **(You left this blank)**

exact\_answernone4407

margin of error +/-

3605



CNN cannot be used for audio or text related problems



CNN cannot be used for audio or text related problems

exact\_answernone3605

margin of error +/-

8471



It is possible to get exact same accuracy everytime we train a network from scratch



It is possible to get exact same accuracy everytime we train a network from scratch

exact\_answernone8471

margin of error +/-

9970



Total number of output channels is dependent on the total number of input channels



Total number of output channels is dependent on the total number of input channels

Your Answer: **(You left this blank)**

exact\_answernone9970

margin of error +/-

9180



It is guaranteed to reach global minima for out DNN if we select right number of layers and kernels



It is guaranteed to reach global minima for out DNN if we select right number of layers and kernels

exact\_answernone9180

margin of error +/-

9929



We mostly use ReLU as activation function, because it is efficient, and accelerated. In all other activation functions, increase in computation requirements does not justify minimal increase in accuracy



We mostly use ReLU as activation function, because it is efficient, and accelerated. In all other activation functions, increase in computation requirements does not justify minimal increase in accuracy

Your Answer: **(You left this blank)**

exact\_answernone9929

margin of error +/-

9718



Softmax improves the accuracy of a network



Softmax improves the accuracy of a network

Your Answer: **(You left this blank)**

exact\_answernone9718

margin of error +/-

9881



We add as many layers as possible, GPU resources being the only constraint



We add as many layers as possible, GPU resources being the only constraint

Your Answer: **(You left this blank)**

exact\_answernone9881

margin of error +/-

8933



Accuracy of a model is dependent on the total number of layers in a model



Accuracy of a model is dependent on the total number of layers in a model

exact\_answernone8933

margin of error +/-

37



Total number of channels in the kernel is dependent on the total number of output channels



Total number of channels in the kernel is dependent on the total number of output channels

exact\_answernone37

margin of error +/-

7039



We initialize all our kernels/filters randomly



We initialize all our kernels/filters randomly

Your Answer: **(You left this blank)**

exact\_answernone7039

margin of error +/-

Quiz Score: **138.33** out of 260