Instructions

Instructions:

1. You have 45 minutes to attempt the S2-Solution.
2. Make sure you have played around with the COLAB FILE shared earlier. Here is the link again .
3. Once you start the solution, you cannot go back and re-attempt it
4. You will not find answers online, so please make sure you are ready for the quiz
5. For Multiple Answer Questions, ALL the answers must be correct to score any point
6. You will be training a model "during" this submission so make sure you are on your laptop.
7. Only 1 question will be shown at once
8. Once answered, question will be locked

Please make sure that you have good internet connection, else you will lose you data. There is only 1 attempt available for this quiz.

This quiz was locked Apr 5 at 6:30am.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	28 minutes	171.67 out of 200

Score for this quiz: **171.67** out of 200
Submitted Apr 3 at 11:34am
This attempt took 28 minutes.

Question 1 — 10 / 10 pts

What is torch?

Correct!

- [x] An open source machine learning framework that accelerates the path from research prototyping to production deployment.
- [ ] is a fictional superhero appearing in American comic books published by Marvel Comics.
- [ ] a portable battery-powered electric lamp.

Question 2 — 0 / 10 pts

What is the purpose of adding padding=1?

You Answered

- [x] To create equal size output after convolution with any kernel

Correct!

- [x] To add 2 additional pixels in x and y rows for convolution
- [ ] To provide cushioning to the channels before kernel hits with a great force

You Answered

- [x] To increase the kernel size by 2px in x and y columns

Question 3 — 5 / 10 pts

What is that -1 in output shape when we call summary(model, input_size=(1, 28, 28))?

Correct!

- [x] It refers to the batch size

Correct Answer

- [ ] It refers to the dimension "outside" what might be available of input_size
- [ ] it refers to the z-axis
- [ ] It refers to the z-axis of the kernels

Question 4 — 10 / 10 pts

What is CUDA?

Correct!

- [x] CUDA is a parallel computing platform and application programming interface model created by Nvidia. It allows software developers and software engineers to use a CUDA-enabled graphics processing unit for general purpose processing – an approach termed GPGPU
- [ ] CUDA is a garbage collector
- [ ] An end-to-end open source machine learning platform.
- [ ] Something without which my journey in ML would be useless! :(

Question 5 — 6.67 / 10 pts

What is a Tensor?

Correct Answer

- [ ] A tensor is a container which can house data in N dimensions.
- [ ] A tensor is a matrix

Correct!

- [x] Tensor is NOT a matrix, as matrices are specifically 2D, where as Tensors can be nD

Correct!

- [x] Is an algebraic object that describes a linear mapping from one set of algebraic objects to another

Question 6 — 0 / 10 pts

What is 0.1307 and 0.3081 in transforms.Normalize?

Correct Answer

- [ ] That's mean and std of the complete dataset
- [ ] I don't know, and I don't care!

You Answered

- [x] that's std and mean of the dataset
- [ ] That's mean and std of the training set

Question 7 — 10 / 10 pts

What is the use of torch.no_grad()?

Correct!

- [x] To perform inference, but without training

Correct!

- [x] To make sure test data does not "leak" into the model

Correct!

- [x] To perform inference without gradient calculation
- [ ] To tell us that knowing just this function won't help us get graduation degree

Question 8 — 70 / 70 pts

What the hell is wrong with this model? Generally in 1 epoch we should be able to get 95%+, but here we do not? Explain according to you what is wrong with the model. 0 Points if you miss the main point.

Your Answer:

We are using two activation functions in the last layer namely the ReLU and log_softmax. Here the ReLU is a Linear function and for high value gives output as high values and the softmax function increases it exponentially. Hence decreasing the accuracy to just 38% instead of 95%+.

Ideally, we would want to have values between 0 and 1 in out output layer hence using a sigmoid function which has output in range [0,1] can we be used for better results or directly using the soft_max function without ReLU.

עבודה טובה!

Question 9 — 60 / 60 pts

Only 1 change is required in this model such that it gets up to 97% within 1 epoch!

What is that 1 change?

Your Answer:

The removal of ReLU function from the last later and directly passing the convoluted layer to log_softmax can increase the accuracy to 98% within 1 epoch.

עבודה טובה!

Quiz Score: **171.67** out of 200