

HW3

- You will learn how to train a deep network using PyTorch tool. Please read the following tutorial. You may skip the data parallelism section. Please use Google Colab to perform the experiments.

https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html

- CIFAR10 is a dataset of 60,000 color images of size 32×32 from 10 categories. Please download the PyTorch tutorial code for CIFAR10 to start:

https://pytorch.org/tutorials/_downloads/cifar10_tutorial.py

- When you run the tutorial code, it will download CIFAR10 dataset for you. Please follow the instructions in the following link to install PyTorch: <https://pytorch.org/>

- To learn more, you can also find a tutorial for MNIST here:

https://pytorch.org/tutorials/beginner/blitz/neural_networks_tutorial.html#sphx-glr-beginner-blitz-neural-networks-tutorial-py

and the sample model for MNIST here:

<https://github.com/pytorch/examples/blob/master/mnist/main.py>

and the sample code for Imagenet here:

<https://github.com/pytorch/examples/blob/master/imagenet/main.py>

- For all the following sections, train the model for 50 epochs and plot the curve for loss, training accuracy, and test accuracy evaluated every epoch.

1. Run the tutorial code out of the box and make sure you get reasonable results. You will report these results later, so no report needed here.
2. (15 points) Change the code to have only a single fully connected layer. The model will have a single layer that connects the input to the output. What is the number of parameters? In PyTorch, "nn.Linear" can be used for fully connected layer.
3. (15 points) Change the code to have multiple fully connected layers. Try having a layer from input to 140 neurons and then a layer to 84 neurons, and finally a layer to 10 neurons, one for each category. What happens if you do not use ReLU? Describe why.
4. (15 points) Change the code by adding two convolutional layers along with maxpooling layers before the fully connected layers. This will be similar to the example in the tutorial. Use this model for the following sections.
5. (15 points) Try multiple batch sizes to see the effect and describe the findings. Please use batch size of 1, 4, and 1000. If 1000 does not fit into the memory of your machine, please feel free to reduce it to a largest possible number.
6. (15 points) Try multiple learning rates to see the effect and describe the findings. Please use learning rates of 10, 0.1, 0.01, and 0.0001.
7. (25 points) Please add some data augmentation to avoid overfitting. Note that you need to do this only for the training and not the testing. You may use Line 208 from Imagenet sample code:

<https://github.com/pytorch/examples/blob/master/imagenet/main.py>

"RandomResizedCrop" samples a random patch from the image to train the model on. "RandomHorizontalFlip" flips randomly chosen images horizontally.

8. (15 points) (Extra Credit) Try multiple augmentations (maybe 10 times) at the test time and average the final probability distributions (10 numbers) before evaluation. You may also do Dropout multiple times at the test time and average the results before evaluation.