

3B_Convolutional_Neural_Networks_Assignment

May 12, 2021

1 PyTorch Assignment: Convolutional Neural Network (CNN)

Duke Community Standard: By typing your name below, you are certifying that you have adhered to the Duke Community Standard in completing this assignment.

Name:

1.0.1 Convolutional Neural Network

Adapt the CNN example for MNIST digit classification from Notebook 3A. Feel free to play around with the model architecture and see how the training time/performance changes, but to begin, try the following:

Image ->

convolution (32 3x3 filters) -> nonlinearity (ReLU) ->

convolution (32 3x3 filters) -> nonlinearity (ReLU) -> (2x2 max pool) ->

convolution (64 3x3 filters) -> nonlinearity (ReLU) ->

convolution (64 3x3 filters) -> nonlinearity (ReLU) -> (2x2 max pool) -> flatten -> fully connected (256 hidden units) -> nonlinearity (ReLU) ->

fully connected (10 hidden units) -> softmax

Note: The CNN model might take a while to train. Depending on your machine, you might expect this to take up to half an hour. If you see your validation performance start to plateau, you can kill the training.

[]: `### YOUR CODE HERE ###`

1.0.2 Short answer

1. How does the CNN compare in accuracy with yesterday's logistic regression and MLP models? How about training time?

[Your answer here]

2. How many trainable parameters are there in the CNN you built for this assignment?

Note: The total of trainable parameters counts each element in a tensor. For example, a weight matrix that is 10x5 has 50 trainable parameters.

[Your answer here]

3. When would you use a CNN versus a logistic regression model or an MLP?

[Your answer here]