Programming Assignment 1

You have to solve a classification problem, where your task is to classify handwritten English digits 0-9. So your input is an image containing a digit and the output will be a number representing the digit in the image. There is a popular MNIST hand-written digit classification dataset. The dataset contains a collection of images (i.e., images of hand-written digits) and their corresponding labels (0-9). This dataset is available in PyTorch under the torchvision package. By using this dataset you need to do the following:

a. Design an MLP network having three layers, where the first layer output is of dimension 128, the second layer output is of dimension 32, and the third layer output is 10 (the number of different types of digits). First two layers output will be feeded to the Relu function before feeding into the next layer, and the last layer output will be passed through a Softmax activation function. Use Pytorch's nn.Linear class to implement the layers of the network. You need to vary different hyper parameters such as batch size (8, **16**, 32), the number of epochs (1, 5, **10**, 20), and the learning rate (.0001, **.001**, .01) and see the performance of your trained model in different scenarios. While you vary one parameter, you can keep the other parameters as their default values (marked in bold). Report the best performing combination of hyper parameters and its results.

b. You have to improve your previous network to enhance the performance of the digital classification. At this stage you need to add a convolutional layer followed by a pooling layer and then feed the output to the first MLP layer that you have already developed. For the convolution layer, you need to show the results with three different hyper parameter configurations: number of filters/ output channel = 1 with filter size 2 x 2; number of filters/ output channel = 10 with filter size = 3x3; and number of filters/output channel = 10 with filter size 4x4. Assume the stride s=1 for all the cases. Then add a pooling layer with a filter size of 2x2 and finally feed the output of this layer to the MLP.

c. Extend the above CNN network by adding a second convolution-pooling layer and check whether the performance can be further improved. In this case the first convolution layer will have the number of filters/ output channels = 10 with filter size 2 x 2, and the second convolution layer will have the number of filters/ output channels = 10 with filter size 2 x 2. Use a pooling layer with a filter size of 2x2 after both convolution layers.

Submission

You can submit until 11th February, 11:59 pm.

You need to submit a zip package that contains source codes (for (a), (b), and(c)) and a report pdf file containing your results and hyperparameters for the best performing model for each of the above scenarios (for (a), (b), and(c)). The name of your zip file should be Group ID_Student ID1_Student ID2

Result Representation
Report the accuracy, and confusion metrics of the test dataset for each of the three cases.