ECE 57000 Assignment 3 Exercise

Your Name:

Exercise 1: Why use a CNN rather than only fully connected layers? (MNIST)

In this exercise, you will build two models for the **MNIST** dataset: one uses only fully connected layers and another uses a standard CNN layout (convolution layers everywhere except the last layer is fully connected layer). The two models should be built with roughly the same accuracy performance, your task is to compare the number of network parameters (a huge number of parameters can affect training/testing time, memory requirements, overfitting, etc.).

Task 1: Following the structure used in the instructions, you should create

- One network named **OurFC** which should consist with only fully connected layers
  - You should decide how many layers and how many hidden dimensions you want in your network
  - Your final accuracy on the test dataset should lie roughly around 90% ($\pm$2%)
  - There is no need to make the neural network unnecessarily complex, your total training time should no longer than 3 mins
- Another network named **OurCNN** which applys a standard CNN structure
  - Again, you should decide how many layers and how many channels you want for each layer.
  - Your final accuracy also should lie roughly around 90% ($\pm$2%)
  - A standard CNN structure can be composed as **[Conv2d, MaxPooling, ReLU] x num_conv_layers + FC x num_fc_layers**
- Train and test your network on MNIST data as in the instructions

In [ ]:

Task 2: Compare the number of parameters that are used in both your neural networks by printing out the total number of parameters for both of your networks.

In [ ]:

Exercise 2: Train classifier on CIFAR-10 data.

Now, lets move our dataset to color images. CIFAR-10 dataset is another widely used dataset. Here all images have colors, i.e each image has 3 color channels instead of only one channel in MNIST. You need to pay more attention to the dimension of the data when it is walking through your network.

Task 1:

Set up a `train_loader` and `test_loader` for the CIFAR-10 data, and plot a figure:

- 3 x 3 subplot
- each subplot is a randomly chosen image from the test dataset
- label each image with its label

The corresponding names of the classes is given as `classes = ('plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship', 'truck')`

Note: In your transforms, the normalizing constant is given as `transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5))])`

In [ ]:
```
# ----------- <Your code> ---------------
```

Task 2:

Set up a convolutional neural network and have your data trained on it. You have to decide all the details in your network, overall your neural network should meet the following standards:

- You should not use more than three convolutional layers and three fully connected layers
- Accuracy on the test dataset should be roughly 50%

In [ ]:
```
# ----------- <Your code> ---------------
# You should have your test accuracy printed in the output box
```

Task 3:

Plot some misclassified images in your test dataset:

- select three images that are **misclassified** by your neural network
- label each images with true label and predicted label

In [ ]:
```

```

Questions (0 points): Does the mis-classified images somehow also misleading for human eyes? I.e, Does a human also have a high probabability of misclassifying those images?