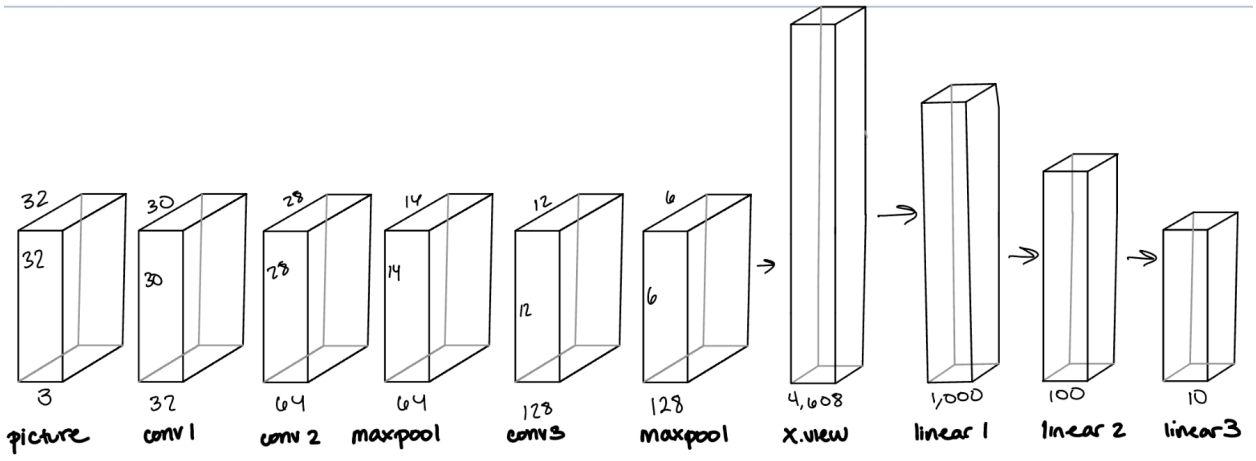
Part 1

My model uses 3 convolutional layers, 2 maxpools and 3 linear layers. All 3 convolutional layers use a kernel size of 3. I also used batch normalization on some of the layers. My activation function was relu and my loss function was cross entropy. I used **x** epochs, a batch size of 300 and a learning rate of 0.001. The model combines the different layers in the following order:

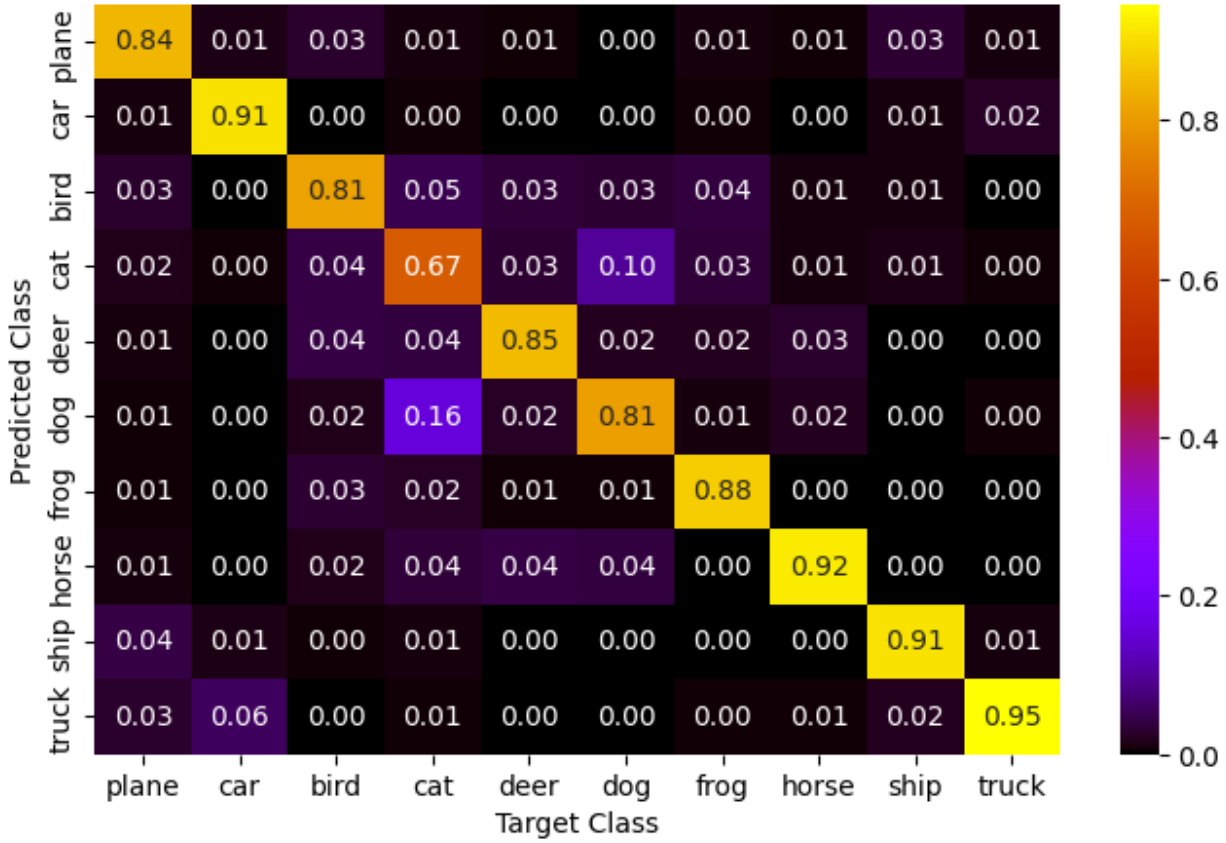
1. Relu + Batch Normalization + Convolution (30 x 30 x 32)
2. Relu + Convolution (28 x 28 x 64)
3. Maxpool (14 x 14 x 64)
4. Relu + Batch Normalization + Convolution (12 x 12 x 128)
5. Maxpool (6 x 6 x 128)
6. Reshaping using x.view (4,608)
7. Relu + Linear (1,000)
8. Relu + Linear (100)
9. Linear (10)



Part 2

Train Set: Accuracy: 43888/50000 (87.8%)

Test Set: Accuracy: 8549/10000 (85.5%)



Part 3

Test Set: Accuracy: 8517/10000 (85.2%)