Assignment 2

Task 1:

文本

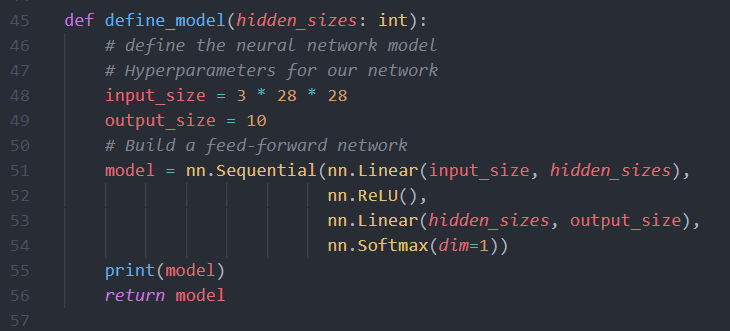
描述已自动生成

文本

描述已自动生成

The transformation that I choose for validation set and test set are ToTensor() and Normailize(). ToTensor() is required and Normailize() generally speeds up the learning process and leads to faster convergence. For training set, in addition to ToTensor() and Normalize(), I also add RandomHorizontalFlip and ColorJitter because I want neural network to be able recognize letter even it is horizontally flipped or the brightness is different. To split data set into training, validation, test sets, train\_test\_split is the helper function I use. After removing the corrupted images, the training set will have 15000 images, validation set has 1000 images and test set has 2724 images. A small batch size will result in a longer training time, so I choose 32 in my case which suits best on my pc.

Task 2:



The model has only one hidden layer with Relu as activation function and has a softmax function after the output layer.

文本

描述已自动生成

The loss function is defined as CrossEntropyLoss. The optimizer is SGD.

文本

描述已自动生成

For each epoch, training loss and training accuracy has been calculated.

文本

描述已自动生成

The validation accuracy and validation loss are calculated as well.

文本

描述已自动生成

The early stop is implemented as well. By default, 2 consecutive increasing in validation loss will stop the training process.

文本

描述已自动生成

The five learning rates I choose are 0.1, 0.05, 0.03, 0.01, 0.005. The training process takes a bit longer time, so I’ve saved the print messages in Task 2 print result.txt file. It turns out that 0.1 works the best. The plot is shown below.

The plot has been saved as plot\_result\_task2.png. And for this model, the test loss is 1.5706202945345125 and the test accuracy is 89.46%. (Result can be different if training multiple times)

图表, 折线图

描述已自动生成

The best training loss and validation loss are 1.5394 and 1.5562

Task 3:

文本

描述已自动生成

Training model with layer size = 100, 500, and 1000 separately. I use 0.1 as learning rate. The validation losses are 1.5982, 1.5863, 1.5828

The best model has 1000 units in hidden layer

The test loss is 1.5748151756068158. The test accuracy is 89.24%

The validation losses are extremely close to each other with hidden unit = 1000 works slightly better than 500 which works slightly better than 100. Thus, the more the number of units are, the lower the loss is. The printed messages are stored in Task 3 print result.txt file.

Task 4:

文本

描述已自动生成

The model is defined above

文本

描述已自动生成

The learning rate I use is 0.1 since it works the best in task 2. The validation loss and training loss are shown below. The plot has been saved as plot\_result\_task4.png.

图表, 折线图

描述已自动生成

It early stops at epoch 6.

The best training Loss is 1.5708 and the best validation Loss: 1.5587 which happens at epoch4.

The test loss is 1.5820134526307363. The test accuracy is 88.40%

Compare with model that has only 1 layer with 1000 hidden units, the validation loss of two layers neural network is slightly lower. However, the test results are quite closed. The print messages are stored in Task 4 print result.txt file.

Task 5:

文本

描述已自动生成

The model is defined above. The final training loss and validation loss for this model is 1.5654 and 1.5530. The training loss and validation loss in task 2 are 1.5394 and 1.5562. As we can see the training loss is higher when applying dropout. That's because the dropout prevent model from overfitting, so the training loss is expected to be higher. The print messages are stored in Task 5 print result.txt file.

图表, 折线图

描述已自动生成

The test loss is 1.581805121355995. The test accuracy is 87.89%

The plot has been saved as plot\_result\_task5.png.