

Homework 4 – CS541 – Deep Learning

Joshua Levy

Sreejani Chatterjee

Section 1 – Feed-forward neural network

Hyperparameters used:

```
Hidden_layers_list=[3,4,5]
hidden_numbers_list = [50,30,40]
mini_batch_size_list=[16,50,100]
epsilon_list=[0.1,0.01,0.05,0.001]
epochs_list=[50,75,100]
alpha_list=[0.00001,0.0001,0.00002]
```

Results

Best Hyperparameters:

```
Best Hyper Parameters:
Best_Hidden_Layers: 3
Best_hidden_num: 50
Best_epochs: 50
Best_batch_size: 50
Best_learning_rate: 0.01
Best_alpha: 1e-05
```

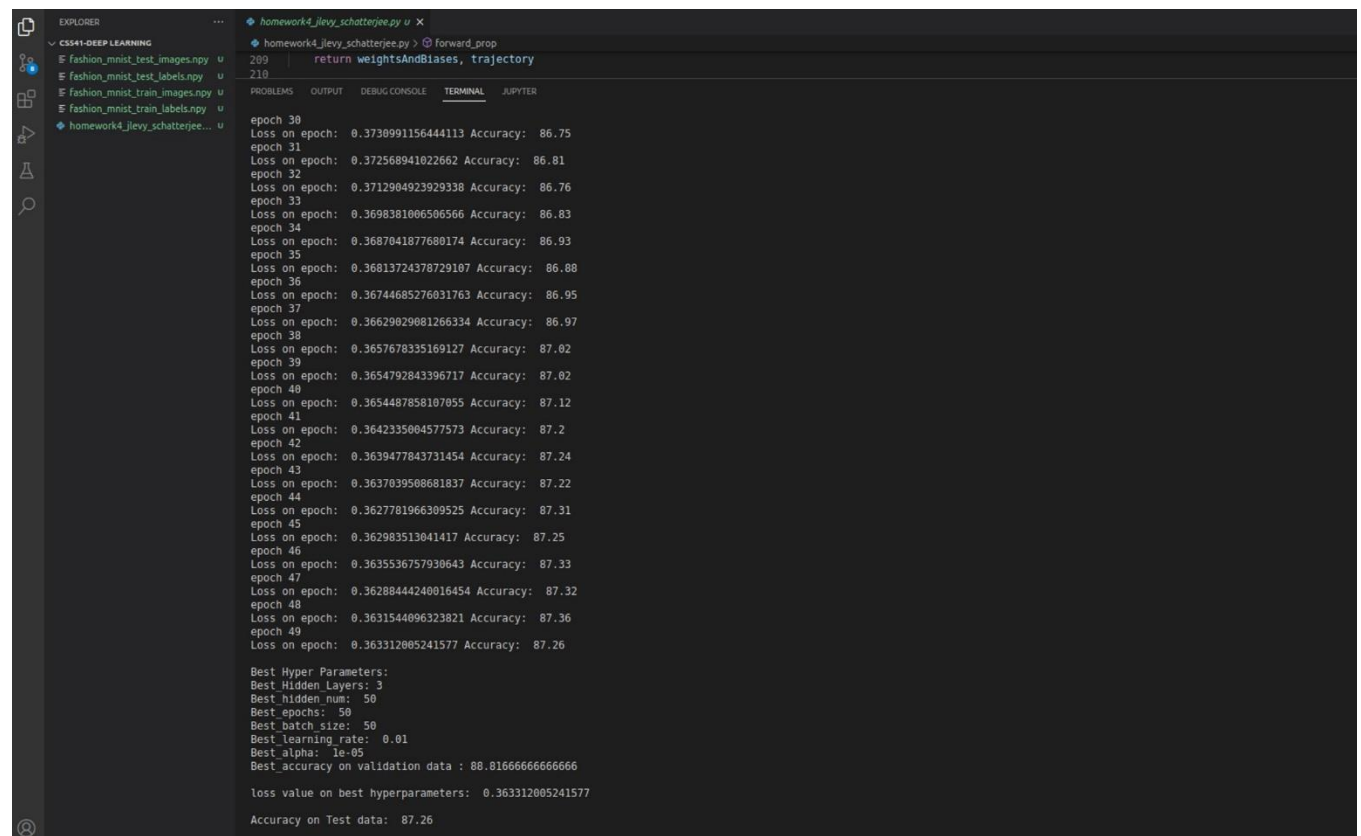
Check Grad:

```
=====scipy.optimize.check_grad=====
8.507616980402489e-07
=====
```

Output Accuracy and Loss:

```
Best_accuracy on validation data : 88.81666666666666
loss value on best hyperparameters: 0.363312005241577
Accuracy on Test data: 87.26
```

Last 20 Epochs:



The screenshot shows a Jupyter Notebook interface with a file explorer on the left and a terminal window on the right. The terminal window displays the output of a training process, showing the last 20 epochs of training results. The output includes the loss on epoch, accuracy, and the best hyperparameters found.

```
epoch 30
Loss on epoch: 0.3730991156444113 Accuracy: 86.75
epoch 31
Loss on epoch: 0.372568941022662 Accuracy: 86.81
epoch 32
Loss on epoch: 0.3712904923929338 Accuracy: 86.76
epoch 33
Loss on epoch: 0.3698381006506566 Accuracy: 86.83
epoch 34
Loss on epoch: 0.3687041877680174 Accuracy: 86.93
epoch 35
Loss on epoch: 0.36813724378729107 Accuracy: 86.88
epoch 36
Loss on epoch: 0.36744685276031763 Accuracy: 86.95
epoch 37
Loss on epoch: 0.36629029881266334 Accuracy: 86.97
epoch 38
Loss on epoch: 0.3657678335169127 Accuracy: 87.02
epoch 39
Loss on epoch: 0.3654792843396717 Accuracy: 87.02
epoch 40
Loss on epoch: 0.3654487858107055 Accuracy: 87.12
epoch 41
Loss on epoch: 0.3642335004577573 Accuracy: 87.2
epoch 42
Loss on epoch: 0.3639477843731454 Accuracy: 87.24
epoch 43
Loss on epoch: 0.3637039508681837 Accuracy: 87.22
epoch 44
Loss on epoch: 0.3627701966309525 Accuracy: 87.31
epoch 45
Loss on epoch: 0.362983513041417 Accuracy: 87.25
epoch 46
Loss on epoch: 0.3635536757938643 Accuracy: 87.33
epoch 47
Loss on epoch: 0.36288444240016454 Accuracy: 87.32
epoch 48
Loss on epoch: 0.3631544096323821 Accuracy: 87.36
epoch 49
Loss on epoch: 0.363312005241577 Accuracy: 87.26

Best Hyper Parameters:
Best Hidden Layers: 3
Best hidden num: 50
Best epochs: 50
Best batch size: 50
Best learning rate: 0.01
Best alpha: 1e-05
Best accuracy on validation data : 88.81666666666666
loss value on best hyperparameters: 0.363312005241577
Accuracy on Test data: 87.26
```

```
epoch 30
Loss on epoch: 0.3730991156444113 Accuracy: 86.75
epoch 31
Loss on epoch: 0.372568941022662 Accuracy: 86.81
epoch 32
Loss on epoch: 0.3712904923929338 Accuracy: 86.76
epoch 33
Loss on epoch: 0.3698381006506566 Accuracy: 86.83
epoch 34
Loss on epoch: 0.3687041877680174 Accuracy: 86.93
epoch 35
Loss on epoch: 0.36813724378729107 Accuracy: 86.88
epoch 36
Loss on epoch: 0.36744685276031763 Accuracy: 86.95
epoch 37
Loss on epoch: 0.36629029081266334 Accuracy: 86.97
epoch 38
Loss on epoch: 0.3657678335169127 Accuracy: 87.02
epoch 39
Loss on epoch: 0.3654792843396717 Accuracy: 87.02
epoch 40
Loss on epoch: 0.3654487858107055 Accuracy: 87.12
epoch 41
Loss on epoch: 0.3642335004577573 Accuracy: 87.2
epoch 42
Loss on epoch: 0.3639477843731454 Accuracy: 87.24
epoch 43
Loss on epoch: 0.3637039508681837 Accuracy: 87.22
epoch 44
Loss on epoch: 0.3627781966309525 Accuracy: 87.31
epoch 45
Loss on epoch: 0.362983513041417 Accuracy: 87.25
epoch 46
Loss on epoch: 0.3635536757930643 Accuracy: 87.33
epoch 47
Loss on epoch: 0.36288444240016454 Accuracy: 87.32
epoch 48
Loss on epoch: 0.3631544096323821 Accuracy: 87.36
epoch 49
Loss on epoch: 0.363312005241577 Accuracy: 87.26
```

Best Hyper Parameters:

Best_Hidden_Layers: 3

Best_hidden_num: 50

Best_epochs: 50

Best_batch_size: 50

Best_learning_rate: 0.01

Best_alpha: 1e-05

Best_accuracy on validation data : 88.81666666666666

loss value on best hyperparameters: 0.363312005241577

Accuracy on Test data: 87.26

Section 02 – Mountains and valleys

Visualization of SGD trajectory:

Using 5000 Examples

```
plotSGDPath(trainX[:, 0:5000], trainY[:, 0:5000], trajectory, H_Layers, hidden_num)
```

