A3_Q1-2024

August 14, 2024

1 Computer Vision 2024 Assignment 3: Deep Learning for Perception Tasks

This assignment contains 2 questions. The first question probes understanding of deep learning for classification. The second question is a more challenging classification experiment on a larger dataset. Answer the questions in separate Python notebooks.

1.1 Question 1: A simple classifier, 20 marks

For this exercise, we provide demo code showing how to train a network on a small dataset called Fashion-MNIST. Please run through the code "tutorial-style" to get a sense of what it is doing. Then use the code alongside lecture notes and other resources to understand how to use pytorch libraries to implement, train and use a neural network.

For the Fashion-MNIST dataset the lables from 0-9 correspond to various clothing classes so you might find it convenient to create a python list as follows:

class_names = ['T-shirt/top', 'Trouser', 'Pullover', 'Dress', 'Coat', 'Sandal', 'Shirt', 'Sneaker', 'Bag', 'Ankle boot']

You will need to answer various questions about the system, how it operates, the results of experiments with it and make modifications to it yourself. You can change the training scheme and the network structure.

Organize your own text and code cell to show the answer of each questions.

Detailed requirements:

Q1.1 (1 point)

Extract 3 images of different types of clothing from the training dataset, print out the size/shape of the training images, and display the three with their corresponding labels.

Q1.2 (2 points)

Run the training code for 10 epochs, for different values of the learning rate. Fill in the table below and plot the loss curves for each experiment:

Lr	Accuracy
1	
0.1	
0.01	

Lr	Accuracy
0.001	

Q1.3 (3 points)

Report the number of epochs when the accuracy reaches 85%. Fill in the table below and plot the loss curve for each experiment:

Lr	Accuracy	Epoch
1		
0.1		
0.01		
0.001		

Q1.4 (2 points)

Compare the results in table 1 and table 2, what is your observation and your understanding of learning rate?

Q1.5 (5 points)

Build a wider network by modifying the code that constructs the network so that the hidden layer(s) contain more perceptrons, and record the accuracy along with the number of trainable parameters in your model. Now modify the oroginal network to be deeper instead of wider (i.e. by adding more hidden layers). Record your accuracy and network size findings. Plot the loss curve for each experiment. Write down your conclusions about changing the network structure?

Structures	Accuracy	Parameters
Base		
Deeper		
Wider		

Q1.6 (2 points)

Calculate the mean of the gradients of the loss to all trainable parameters. Plot the gradients curve for the first 100 training steps. What are your observations? Note that this gradients will be saved with the training weight automatically after you call loss.backwards(). Hint: the mean of the gradients decrease.

For more exlanation of q1.7, you could refer to the following simple instructions: https://colab.research.google.com/drive/1XAsyNegGSvMf3_B6MrsXht7-fHqtJ7OW?usp=sharing

Q1.7 (5 points)

Modify the network structure and training/test to use a small convolutional neural network instead of an MLP. Discuss your findings with rehgard to convergence, accuracy and number of parameters, relative to MLPs.

Hint: Look at the structure of the CNN in the Workshop 3 examples.

NOTE: For consistency with the original data set, we call our validation data "test_data". It is important to keep in mind though that we are using the data for model validation and not for testing the final, trained model (which requires data not used when training the model parameters).

We pass the Dataset as an argument to DataLoader. This wraps an iterable over our dataset and supports automatic batching, sampling, shuffling, and multiprocess data loading. Here we define a batch size of 64, i.e. each element in the dataloader iterable will return a batch of 64 features and labels.

```
Shape of X [N, C, H, W]:  torch.Size([64, 1, 28, 28])
Shape of y:  torch.Size([64]) torch.int64
```
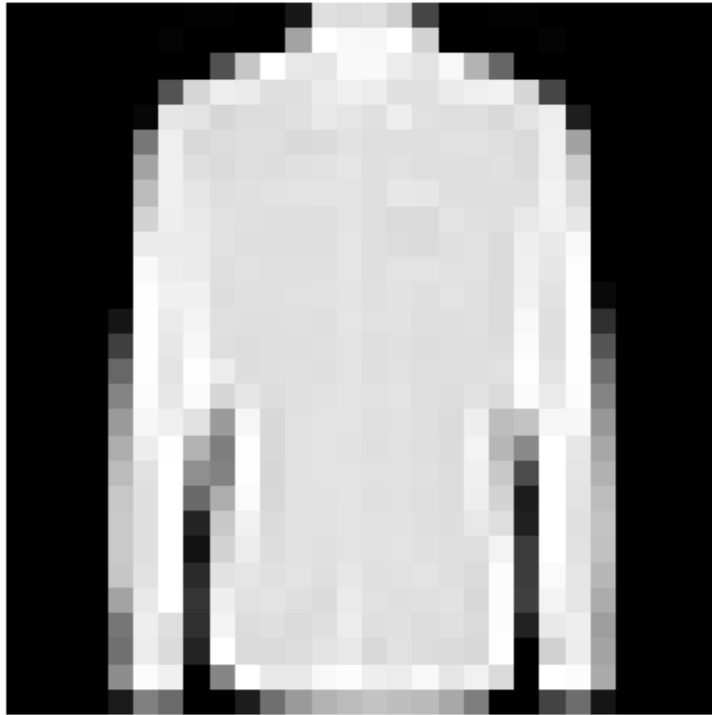
Add in a code cell to inspect the training data, as per Q1.1. Each element of the training_data structure has a greyscale image (which you can use plt.imshow(img[0,:,:]) to display, just like you did in previous assignments.

1.1.1 Question 1.1

Label: Bag
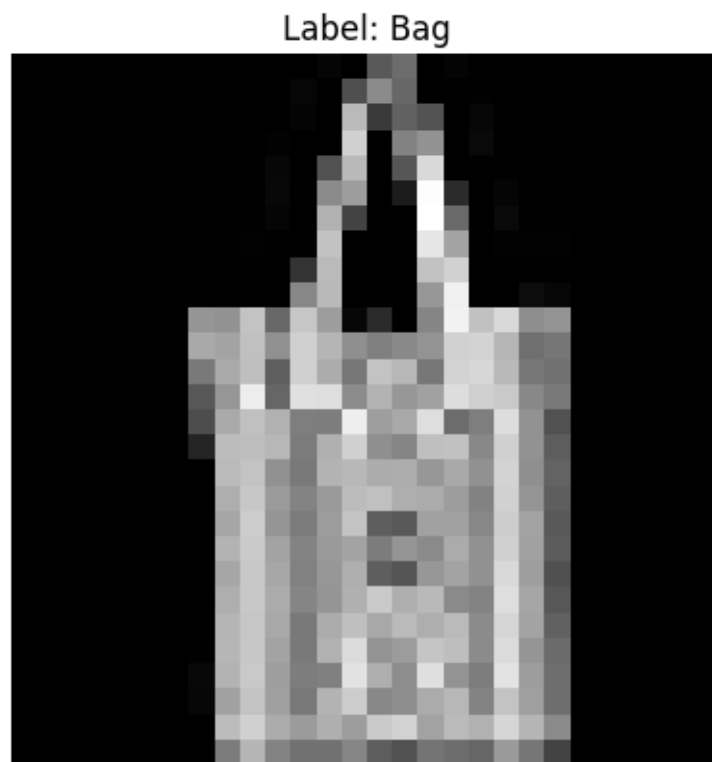
```
Shape: torch.Size([28, 28])
```

Label: Shirt

Shape: torch.Size([28, 28])

Label: Coat
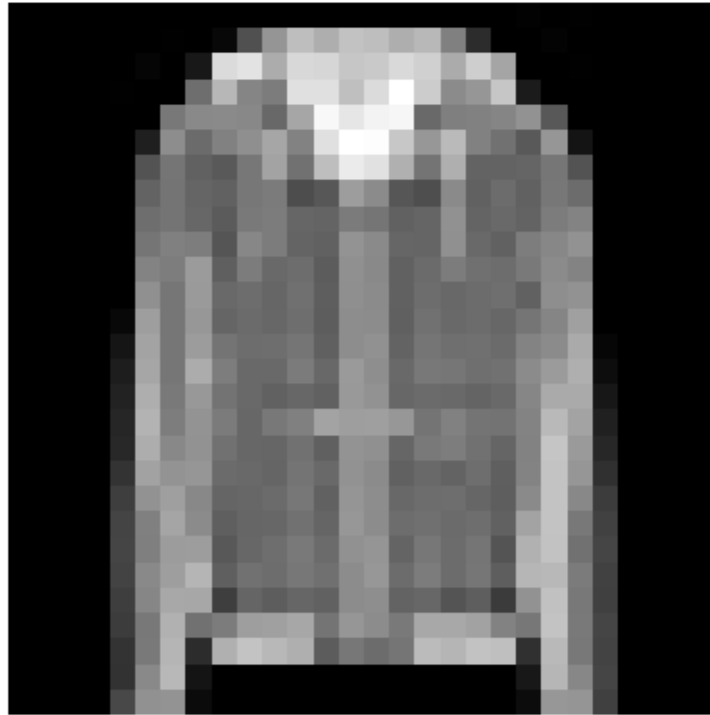
```
Shape: torch.Size([28, 28])
```

To define a neural network in PyTorch, we create a class that inherits from nn.Module. We define the layers of the network in the init function and specify how data will pass through the network in the forward function. To accelerate operations in the neural network, we move it to the GPU if available.

```
Using cuda device
NeuralNetwork(
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (linear_relu_stack): Sequential(
    (0): Linear(in_features=784, out_features=512, bias=True)
    (1): ReLU()
    (2): Linear(in_features=512, out_features=512, bias=True)
    (3): ReLU()
    (4): Linear(in_features=512, out_features=10, bias=True)
  )
)
```

In a single training loop, the model makes predictions on the training dataset (fed to it in batches), and backpropagates the prediction error to adjust the model's parameters.

1.1.2 Question 1.2

```
Training for 10 epochs with learning rate: 1
Epoch 1
-------------------------------
Test Error:
 Accuracy: 10.6%, Avg loss: 2.325459

Epoch 2
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305836

Epoch 3
-------------------------------
Test Error:
 Accuracy: 16.0%, Avg loss: 2.168829

Epoch 4
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 3.489796

Epoch 5
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305837

Epoch 6
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305836

Epoch 7
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305832

Epoch 8
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305843

Epoch 9
-------------------------------
Test Error:
 Accuracy: 10.0%, Avg loss: 2.305843
```

```
Epoch 10
-------------------------------
Test Error:
  Accuracy: 10.0%, Avg loss: 2.305843
```
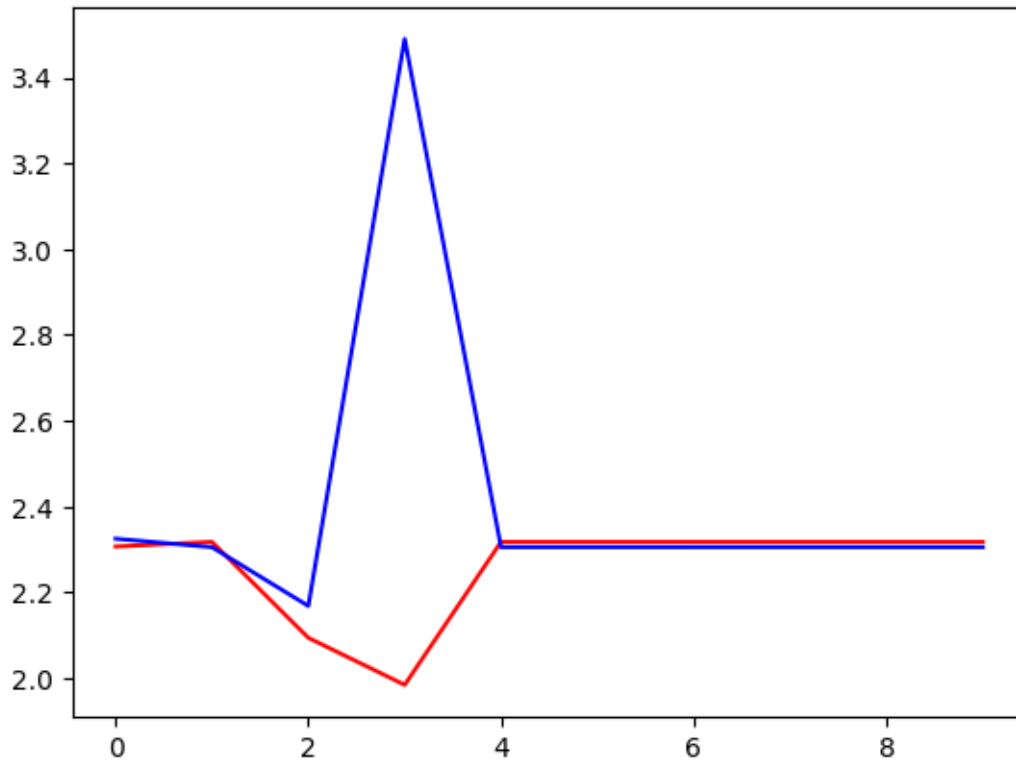


```
Done!

Training for 10 epochs with learning rate: 0.1
Epoch 1
-------------------------------
Test Error:
  Accuracy: 79.1%, Avg loss: 0.548591

Epoch 2
-------------------------------
Test Error:
  Accuracy: 82.7%, Avg loss: 0.468552

Epoch 3
-------------------------------
Test Error:
  Accuracy: 84.5%, Avg loss: 0.422756
```

```
Epoch 4
-------------------------------
Test Error:
 Accuracy: 85.0%, Avg loss: 0.409166

Epoch 5
-------------------------------
Test Error:
 Accuracy: 85.7%, Avg loss: 0.392113

Epoch 6
-------------------------------
Test Error:
 Accuracy: 86.2%, Avg loss: 0.376008

Epoch 7
-------------------------------
Test Error:
 Accuracy: 86.7%, Avg loss: 0.365362

Epoch 8
-------------------------------
Test Error:
 Accuracy: 86.9%, Avg loss: 0.355892

Epoch 9
-------------------------------
Test Error:
 Accuracy: 87.5%, Avg loss: 0.344796

Epoch 10
-------------------------------
Test Error:
 Accuracy: 87.5%, Avg loss: 0.343784
```
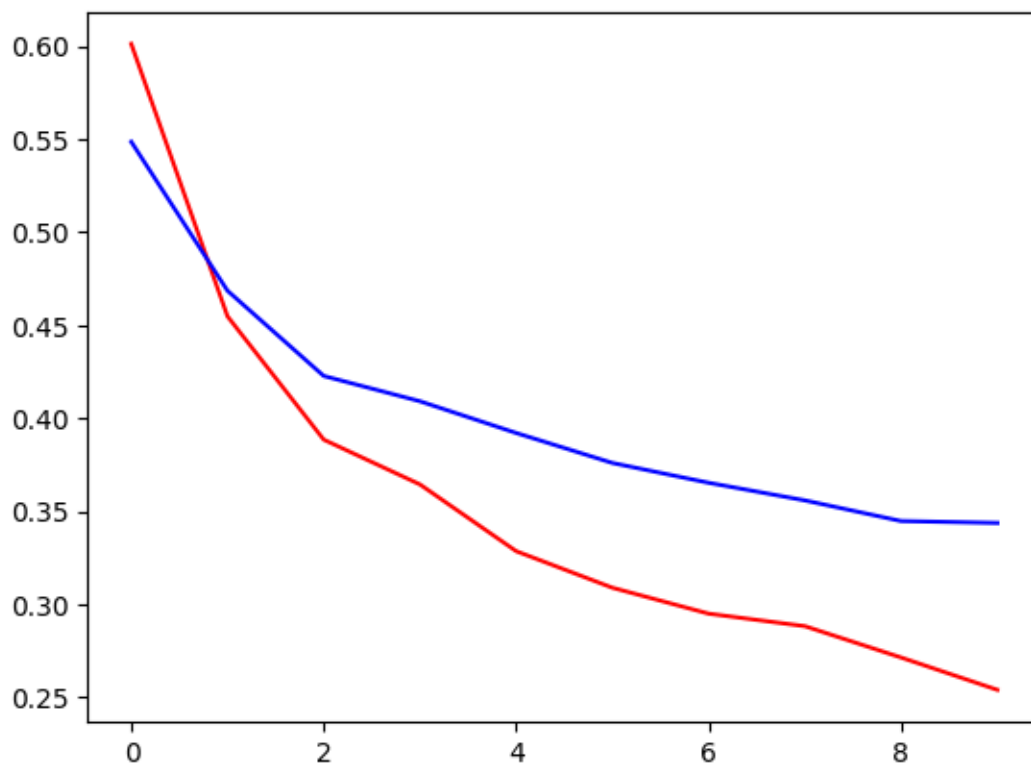
```
Done!

Training for 10 epochs with learning rate: 0.01
Epoch 1
-------------------------------
Test Error:
 Accuracy: 71.0%, Avg loss: 0.799084

Epoch 2
-------------------------------
Test Error:
 Accuracy: 77.9%, Avg loss: 0.636381

Epoch 3
-------------------------------
Test Error:
 Accuracy: 79.8%, Avg loss: 0.573668

Epoch 4
-------------------------------
Test Error:
 Accuracy: 80.6%, Avg loss: 0.542203
```

```
Epoch 5
-------------------------------
Test Error:
 Accuracy: 81.2%, Avg loss: 0.520829

Epoch 6
-------------------------------
Test Error:
 Accuracy: 81.7%, Avg loss: 0.505264

Epoch 7
-------------------------------
Test Error:
 Accuracy: 82.3%, Avg loss: 0.492061

Epoch 8
-------------------------------
Test Error:
 Accuracy: 82.6%, Avg loss: 0.481247

Epoch 9
-------------------------------
Test Error:
 Accuracy: 82.9%, Avg loss: 0.471502

Epoch 10
-------------------------------
Test Error:
 Accuracy: 83.4%, Avg loss: 0.461915
```
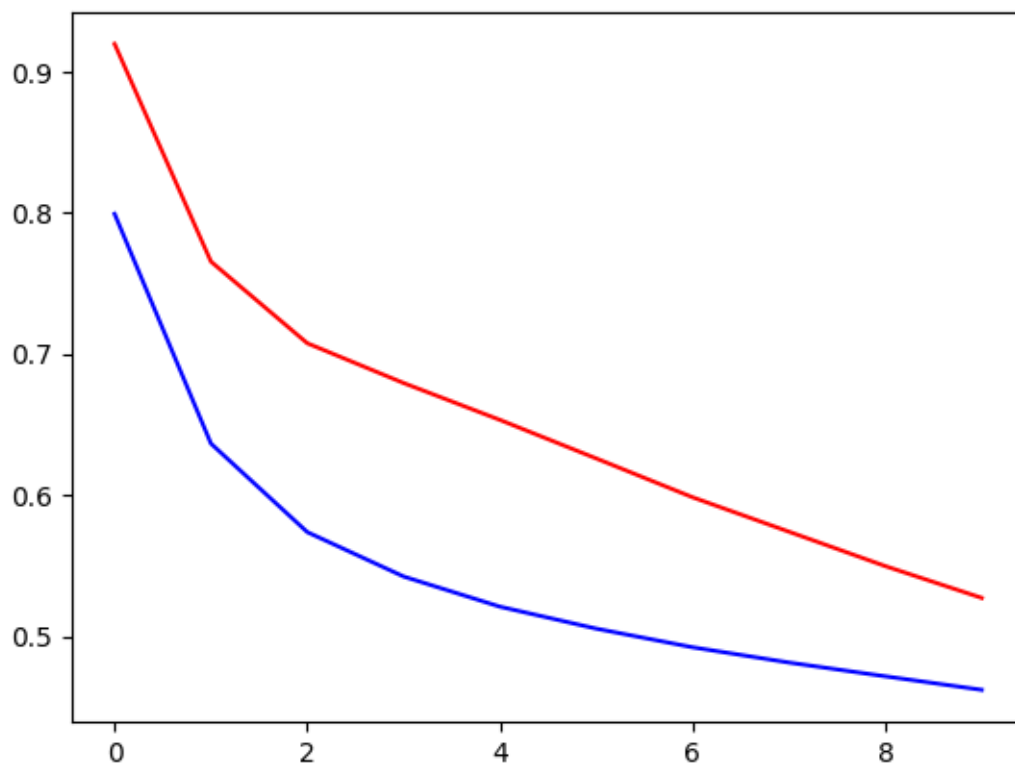
Done!

Training for 10 epochs with learning rate: 0.001
Epoch 1

Test Error:
 Accuracy: 45.7%, Avg loss: 2.147619

Epoch 2

Test Error:
 Accuracy: 57.9%, Avg loss: 1.872733

Epoch 3

Test Error:
 Accuracy: 62.3%, Avg loss: 1.504780

Epoch 4

Test Error:
 Accuracy: 63.7%, Avg loss: 1.243046

```
Epoch 5
-------------------------------
Test Error:
 Accuracy: 64.7%, Avg loss: 1.081236

Epoch 6
-------------------------------
Test Error:
 Accuracy: 66.1%, Avg loss: 0.976525

Epoch 7
-------------------------------
Test Error:
 Accuracy: 67.3%, Avg loss: 0.904825

Epoch 8
-------------------------------
Test Error:
 Accuracy: 68.8%, Avg loss: 0.853037

Epoch 9
-------------------------------
Test Error:
 Accuracy: 69.9%, Avg loss: 0.813733

Epoch 10
-------------------------------
Test Error:
 Accuracy: 71.5%, Avg loss: 0.782497
```
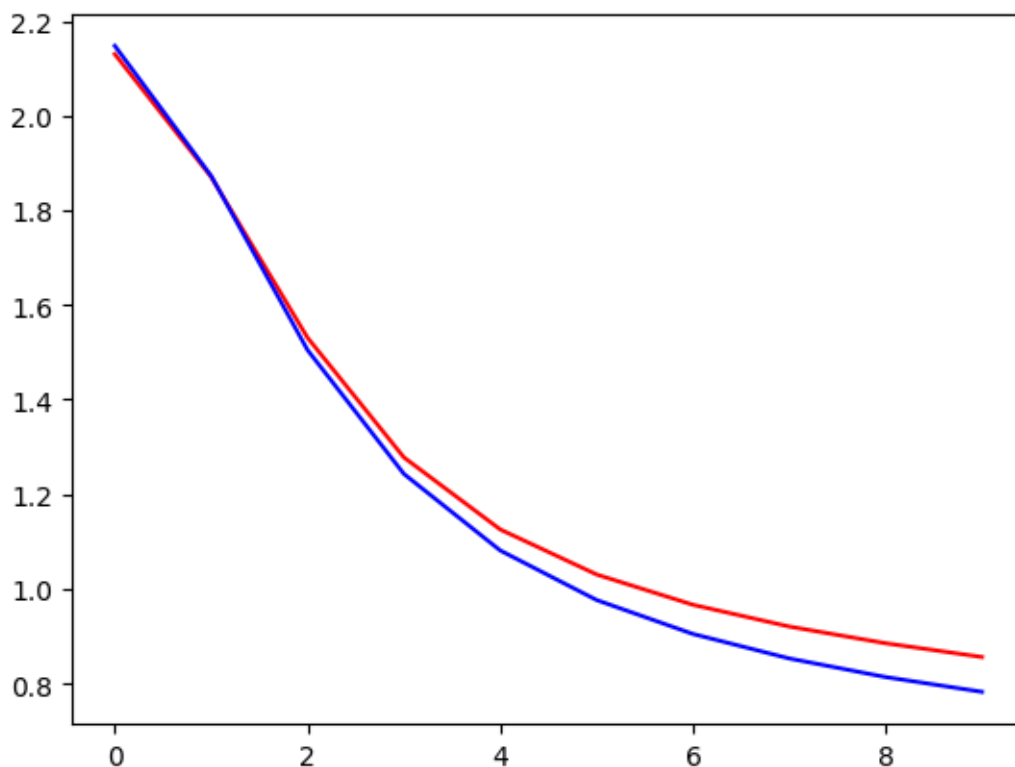
Done!

```
Lr        Accuracy
1         10.0 %
0.1       87.47%
0.01      83.37%
0.001     71.46000000000001%
```

1.1.3 Question 1.3

```
Training with learning rate: 1 until reaching 85% accuracy
Epoch 1
-------------------------------
Test Error:
 Accuracy: 31.1%, Avg loss: 1.634982

Epoch 2
-------------------------------
Test Error:
 Accuracy: 31.7%, Avg loss: 1.732914

Epoch 3
-------------------------------
```

```
Test Error:
 Accuracy: 19.8%, Avg loss: 1.780719

Epoch 4
-------------------------------
Test Error:
 Accuracy: 25.1%, Avg loss: 1.733612

Epoch 5
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.723234

Epoch 6
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.718504

Epoch 7
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.717508

Epoch 8
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.709712

Epoch 9
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.737512

Epoch 10
-------------------------------
Test Error:
 Accuracy: 28.7%, Avg loss: 1.639009

Epoch 11
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.711416

Epoch 12
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.710934
```

```
Epoch 13
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.724419

Epoch 14
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.715178

Epoch 15
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.710580

Epoch 16
-------------------------------
Test Error:
 Accuracy: 19.8%, Avg loss: 1.727894

Epoch 17
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.721424

Epoch 18
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.732628

Epoch 19
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.722661

Epoch 20
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.718941

Epoch 21
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717268

Epoch 22
-------------------------------
Test Error:
```

```
 Accuracy: 19.9%, Avg loss: 1.715571

Epoch 23
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.715567

Epoch 24
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.713838

Epoch 25
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.718623

Epoch 26
-------------------------------
Test Error:
 Accuracy: 20.0%, Avg loss: 1.716997

Epoch 27
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.725365

Epoch 28
-------------------------------
Test Error:
 Accuracy: 19.8%, Avg loss: 1.731262

Epoch 29
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.720152

Epoch 30
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.718114

Epoch 31
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717693

Epoch 32
```

```
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717455


Epoch 33
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717385


Epoch 34
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717354


Epoch 35
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717340


Epoch 36
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717333


Epoch 37
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717328


Epoch 38
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717323


Epoch 39
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717320


Epoch 40
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717316


Epoch 41
--------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717313
```

```
Epoch 42
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717510


Epoch 43
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717390


Epoch 44
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717342


Epoch 45
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717320


Epoch 46
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717309


Epoch 47
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717304


Epoch 48
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717301


Epoch 49
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717299


Epoch 50
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717298


Epoch 51
-------------------------------
```

```
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717298

Epoch 52
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717297

Epoch 53
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717297

Epoch 54
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717296

Epoch 55
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717295

Epoch 56
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717371

Epoch 57
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717225

Epoch 58
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717389

Epoch 59
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717333

Epoch 60
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717314
```

```
Epoch 61
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717304

Epoch 62
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717299

Epoch 63
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717297

Epoch 64
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717295

Epoch 65
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717295

Epoch 66
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 67
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 68
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 69
-----------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 70
-----------------------------
Test Error:
```

```
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 71
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 72
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 73
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 74
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 75
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 76
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 77
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 78
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 79
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 80
```

```
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717291

Epoch 81
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717275

Epoch 82
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717284

Epoch 83
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717289

Epoch 84
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717291

Epoch 85
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 86
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 87
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 88
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 89
------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294
```

```
Epoch 90
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 91
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 92
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 93
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 94
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717301

Epoch 95
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717294

Epoch 96
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 97
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 98
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293

Epoch 99
-------------------------------
```

```
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717275

Epoch 100
-------------------------------
Test Error:
 Accuracy: 19.9%, Avg loss: 1.717293
```
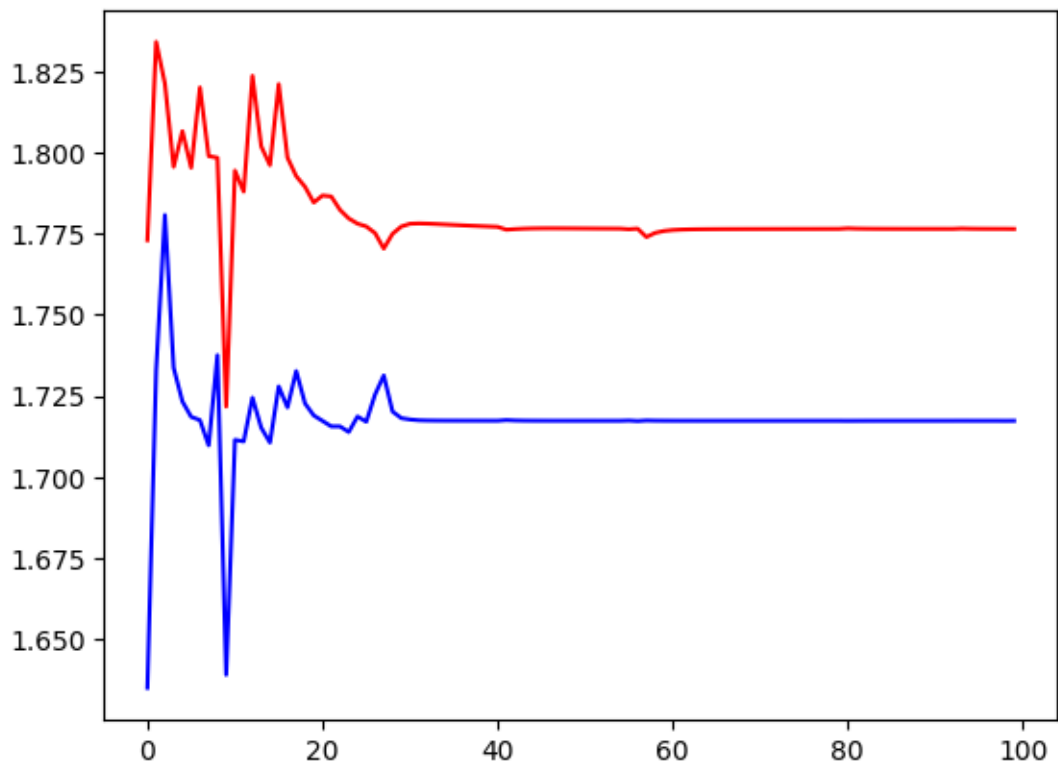


```
Got accuracy 19.9% in 100 epochs
Training with learning rate: 0.1 until reaching 85% accuracy
Epoch 1
-------------------------------
Test Error:
 Accuracy: 79.2%, Avg loss: 0.549380

Epoch 2
-------------------------------
Test Error:
 Accuracy: 81.7%, Avg loss: 0.488078

Epoch 3
-------------------------------
```

```
Test Error:
 Accuracy: 83.9%, Avg loss: 0.432620

Epoch 4
-------------------------------
Test Error:
 Accuracy: 85.0%, Avg loss: 0.406970
```
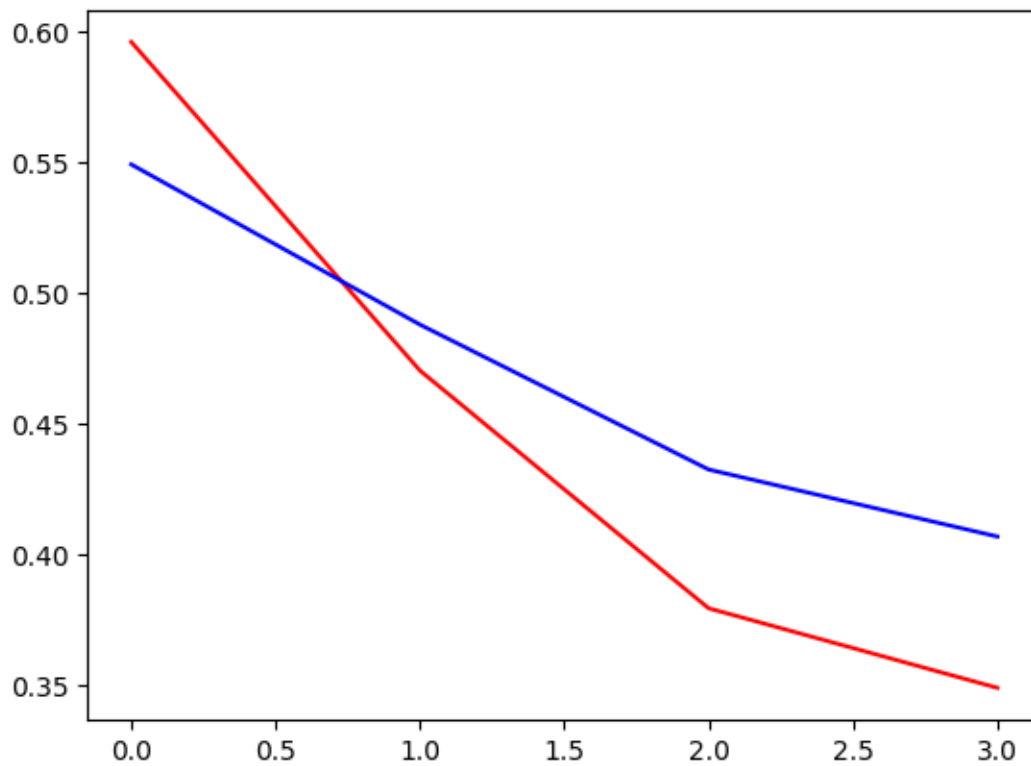


```
Got accuracy 85.0% in 4 epochs
Training with learning rate: 0.01 until reaching 85% accuracy
Epoch 1
-------------------------------
Test Error:
 Accuracy: 71.4%, Avg loss: 0.786560

Epoch 2
-------------------------------
Test Error:
 Accuracy: 78.1%, Avg loss: 0.628442

Epoch 3
-------------------------------
```

```
Test Error:
 Accuracy: 80.1%, Avg loss: 0.564828

Epoch 4
-------------------------------
Test Error:
 Accuracy: 81.0%, Avg loss: 0.534048

Epoch 5
-------------------------------
Test Error:
 Accuracy: 81.4%, Avg loss: 0.514789

Epoch 6
-------------------------------
Test Error:
 Accuracy: 81.9%, Avg loss: 0.499682

Epoch 7
-------------------------------
Test Error:
 Accuracy: 82.4%, Avg loss: 0.488341

Epoch 8
-------------------------------
Test Error:
 Accuracy: 82.8%, Avg loss: 0.478040

Epoch 9
-------------------------------
Test Error:
 Accuracy: 83.1%, Avg loss: 0.468833

Epoch 10
-------------------------------
Test Error:
 Accuracy: 83.4%, Avg loss: 0.460445

Epoch 11
-------------------------------
Test Error:
 Accuracy: 83.7%, Avg loss: 0.452654

Epoch 12
-------------------------------
Test Error:
 Accuracy: 83.9%, Avg loss: 0.445611
```

```
Epoch 13
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.438925

Epoch 14
-------------------------------
Test Error:
 Accuracy: 84.5%, Avg loss: 0.432429

Epoch 15
-------------------------------
Test Error:
 Accuracy: 84.7%, Avg loss: 0.426582

Epoch 16
-------------------------------
Test Error:
 Accuracy: 84.9%, Avg loss: 0.420812

Epoch 17
-------------------------------
Test Error:
 Accuracy: 85.1%, Avg loss: 0.415439
```
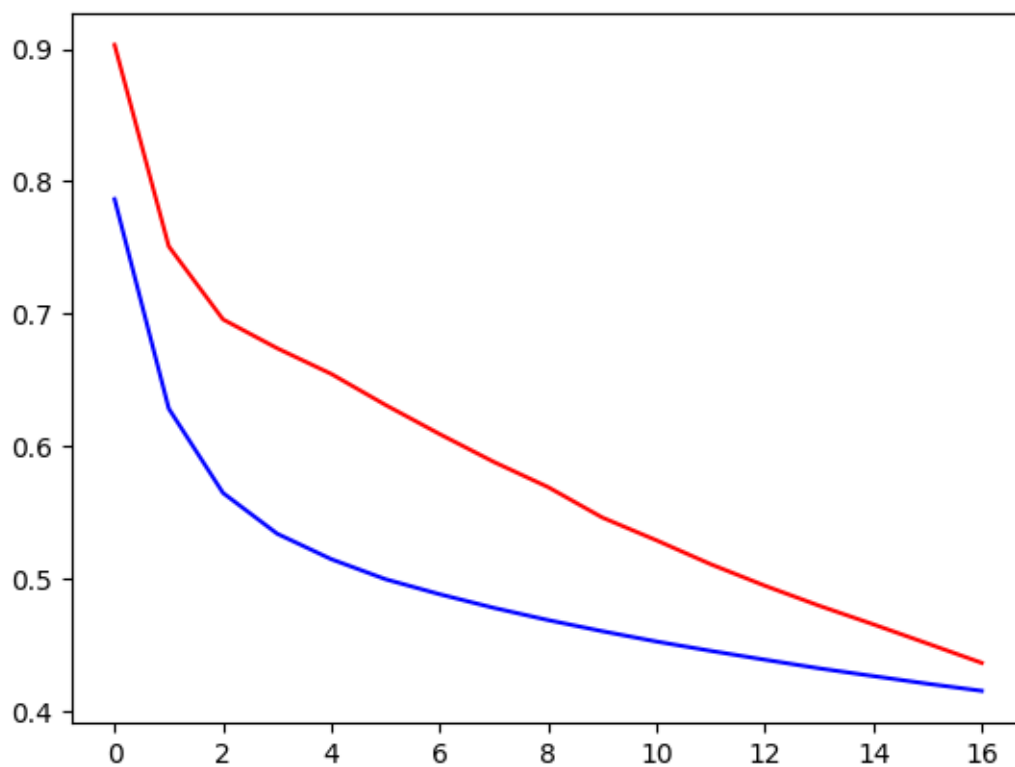
```
Got accuracy 85.1% in 17 epochs
Training with learning rate: 0.001 until reaching 85% accuracy
Epoch 1
-------------------------------
Test Error:
 Accuracy: 44.6%, Avg loss: 2.158727

Epoch 2
-------------------------------
Test Error:
 Accuracy: 56.0%, Avg loss: 1.884165

Epoch 3
-------------------------------
Test Error:
 Accuracy: 60.1%, Avg loss: 1.518632

Epoch 4
-------------------------------
Test Error:
 Accuracy: 63.1%, Avg loss: 1.256925

Epoch 5
-------------------------------
Test Error:
 Accuracy: 64.5%, Avg loss: 1.092697

Epoch 6
-------------------------------
Test Error:
 Accuracy: 65.5%, Avg loss: 0.985632

Epoch 7
-------------------------------
Test Error:
 Accuracy: 66.7%, Avg loss: 0.912364

Epoch 8
-------------------------------
Test Error:
 Accuracy: 67.9%, Avg loss: 0.859665

Epoch 9
-------------------------------
Test Error:
```

```
 Accuracy: 69.0%, Avg loss: 0.820010

Epoch 10
-------------------------------
Test Error:
 Accuracy: 70.2%, Avg loss: 0.788707

Epoch 11
-------------------------------
Test Error:
 Accuracy: 71.7%, Avg loss: 0.762809

Epoch 12
-------------------------------
Test Error:
 Accuracy: 73.0%, Avg loss: 0.740516

Epoch 13
-------------------------------
Test Error:
 Accuracy: 74.0%, Avg loss: 0.720752

Epoch 14
-------------------------------
Test Error:
 Accuracy: 74.7%, Avg loss: 0.702868

Epoch 15
-------------------------------
Test Error:
 Accuracy: 75.5%, Avg loss: 0.686487

Epoch 16
-------------------------------
Test Error:
 Accuracy: 76.2%, Avg loss: 0.671413

Epoch 17
-------------------------------
Test Error:
 Accuracy: 76.8%, Avg loss: 0.657507

Epoch 18
-------------------------------
Test Error:
 Accuracy: 77.3%, Avg loss: 0.644687

Epoch 19
```

```
--------------------------------
Test Error:
 Accuracy: 77.8%, Avg loss: 0.632875

Epoch 20
--------------------------------
Test Error:
 Accuracy: 78.3%, Avg loss: 0.622000

Epoch 21
--------------------------------
Test Error:
 Accuracy: 78.8%, Avg loss: 0.611985

Epoch 22
--------------------------------
Test Error:
 Accuracy: 79.0%, Avg loss: 0.602748

Epoch 23
--------------------------------
Test Error:
 Accuracy: 79.4%, Avg loss: 0.594212

Epoch 24
--------------------------------
Test Error:
 Accuracy: 79.6%, Avg loss: 0.586310

Epoch 25
--------------------------------
Test Error:
 Accuracy: 79.8%, Avg loss: 0.578991

Epoch 26
--------------------------------
Test Error:
 Accuracy: 80.1%, Avg loss: 0.572203

Epoch 27
--------------------------------
Test Error:
 Accuracy: 80.3%, Avg loss: 0.565890

Epoch 28
--------------------------------
Test Error:
 Accuracy: 80.5%, Avg loss: 0.560005
```

```
Epoch 29
-------------------------------
Test Error:
 Accuracy: 80.7%, Avg loss: 0.554508

Epoch 30
-------------------------------
Test Error:
 Accuracy: 80.8%, Avg loss: 0.549356

Epoch 31
-------------------------------
Test Error:
 Accuracy: 81.0%, Avg loss: 0.544522

Epoch 32
-------------------------------
Test Error:
 Accuracy: 81.1%, Avg loss: 0.539977

Epoch 33
-------------------------------
Test Error:
 Accuracy: 81.2%, Avg loss: 0.535708

Epoch 34
-------------------------------
Test Error:
 Accuracy: 81.4%, Avg loss: 0.531698

Epoch 35
-------------------------------
Test Error:
 Accuracy: 81.5%, Avg loss: 0.527926

Epoch 36
-------------------------------
Test Error:
 Accuracy: 81.6%, Avg loss: 0.524375

Epoch 37
-------------------------------
Test Error:
 Accuracy: 81.7%, Avg loss: 0.521024

Epoch 38
-------------------------------
```

```
Test Error:
 Accuracy: 81.8%, Avg loss: 0.517852

Epoch 39
-------------------------------
Test Error:
 Accuracy: 81.8%, Avg loss: 0.514842

Epoch 40
-------------------------------
Test Error:
 Accuracy: 81.9%, Avg loss: 0.511983

Epoch 41
-------------------------------
Test Error:
 Accuracy: 81.9%, Avg loss: 0.509265

Epoch 42
-------------------------------
Test Error:
 Accuracy: 82.1%, Avg loss: 0.506672

Epoch 43
-------------------------------
Test Error:
 Accuracy: 82.1%, Avg loss: 0.504191

Epoch 44
-------------------------------
Test Error:
 Accuracy: 82.2%, Avg loss: 0.501816

Epoch 45
-------------------------------
Test Error:
 Accuracy: 82.2%, Avg loss: 0.499543

Epoch 46
-------------------------------
Test Error:
 Accuracy: 82.3%, Avg loss: 0.497365

Epoch 47
-------------------------------
Test Error:
 Accuracy: 82.3%, Avg loss: 0.495273
```

```
Epoch 48
-------------------------------
Test Error:
 Accuracy: 82.5%, Avg loss: 0.493264


Epoch 49
-------------------------------
Test Error:
 Accuracy: 82.6%, Avg loss: 0.491333


Epoch 50
-------------------------------
Test Error:
 Accuracy: 82.7%, Avg loss: 0.489467


Epoch 51
-------------------------------
Test Error:
 Accuracy: 82.8%, Avg loss: 0.487663


Epoch 52
-------------------------------
Test Error:
 Accuracy: 82.8%, Avg loss: 0.485915


Epoch 53
-------------------------------
Test Error:
 Accuracy: 82.8%, Avg loss: 0.484231


Epoch 54
-------------------------------
Test Error:
 Accuracy: 82.9%, Avg loss: 0.482603


Epoch 55
-------------------------------
Test Error:
 Accuracy: 82.9%, Avg loss: 0.481026


Epoch 56
-------------------------------
Test Error:
 Accuracy: 82.9%, Avg loss: 0.479494


Epoch 57
-------------------------------
Test Error:
```

```
 Accuracy: 83.0%, Avg loss: 0.478003

Epoch 58
-------------------------------
Test Error:
 Accuracy: 83.1%, Avg loss: 0.476550

Epoch 59
-------------------------------
Test Error:
 Accuracy: 83.0%, Avg loss: 0.475135

Epoch 60
-------------------------------
Test Error:
 Accuracy: 83.1%, Avg loss: 0.473756

Epoch 61
-------------------------------
Test Error:
 Accuracy: 83.2%, Avg loss: 0.472409

Epoch 62
-------------------------------
Test Error:
 Accuracy: 83.2%, Avg loss: 0.471092

Epoch 63
-------------------------------
Test Error:
 Accuracy: 83.3%, Avg loss: 0.469807

Epoch 64
-------------------------------
Test Error:
 Accuracy: 83.3%, Avg loss: 0.468551

Epoch 65
-------------------------------
Test Error:
 Accuracy: 83.4%, Avg loss: 0.467321

Epoch 66
-------------------------------
Test Error:
 Accuracy: 83.5%, Avg loss: 0.466120

Epoch 67
```

```
------------------------------
Test Error:
 Accuracy: 83.5%, Avg loss: 0.464936

Epoch 68
------------------------------
Test Error:
 Accuracy: 83.5%, Avg loss: 0.463773

Epoch 69
------------------------------
Test Error:
 Accuracy: 83.6%, Avg loss: 0.462631

Epoch 70
------------------------------
Test Error:
 Accuracy: 83.6%, Avg loss: 0.461511

Epoch 71
------------------------------
Test Error:
 Accuracy: 83.7%, Avg loss: 0.460414

Epoch 72
------------------------------
Test Error:
 Accuracy: 83.7%, Avg loss: 0.459334

Epoch 73
------------------------------
Test Error:
 Accuracy: 83.7%, Avg loss: 0.458273

Epoch 74
------------------------------
Test Error:
 Accuracy: 83.8%, Avg loss: 0.457226

Epoch 75
------------------------------
Test Error:
 Accuracy: 83.8%, Avg loss: 0.456194

Epoch 76
------------------------------
Test Error:
 Accuracy: 83.9%, Avg loss: 0.455172
```

```
Epoch 77
-------------------------------
Test Error:
 Accuracy: 83.9%, Avg loss: 0.454163

Epoch 78
-------------------------------
Test Error:
 Accuracy: 83.9%, Avg loss: 0.453170

Epoch 79
-------------------------------
Test Error:
 Accuracy: 84.0%, Avg loss: 0.452188

Epoch 80
-------------------------------
Test Error:
 Accuracy: 84.0%, Avg loss: 0.451215

Epoch 81
-------------------------------
Test Error:
 Accuracy: 84.0%, Avg loss: 0.450259

Epoch 82
-------------------------------
Test Error:
 Accuracy: 84.1%, Avg loss: 0.449316

Epoch 83
-------------------------------
Test Error:
 Accuracy: 84.1%, Avg loss: 0.448389

Epoch 84
-------------------------------
Test Error:
 Accuracy: 84.1%, Avg loss: 0.447481

Epoch 85
-------------------------------
Test Error:
 Accuracy: 84.1%, Avg loss: 0.446587

Epoch 86
-------------------------------
```

```
Test Error:
 Accuracy: 84.1%, Avg loss: 0.445702

Epoch 87
-------------------------------
Test Error:
 Accuracy: 84.1%, Avg loss: 0.444826

Epoch 88
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.443952

Epoch 89
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.443094

Epoch 90
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.442246

Epoch 91
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.441408

Epoch 92
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.440574

Epoch 93
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.439750

Epoch 94
-------------------------------
Test Error:
 Accuracy: 84.2%, Avg loss: 0.438933

Epoch 95
-------------------------------
Test Error:
 Accuracy: 84.3%, Avg loss: 0.438133
```

```
Epoch 96
-------------------------------
Test Error:
 Accuracy: 84.3%, Avg loss: 0.437340

Epoch 97
-------------------------------
Test Error:
 Accuracy: 84.3%, Avg loss: 0.436551

Epoch 98
-------------------------------
Test Error:
 Accuracy: 84.3%, Avg loss: 0.435774

Epoch 99
-------------------------------
Test Error:
 Accuracy: 84.4%, Avg loss: 0.435004

Epoch 100
-------------------------------
Test Error:
 Accuracy: 84.4%, Avg loss: 0.434238
```
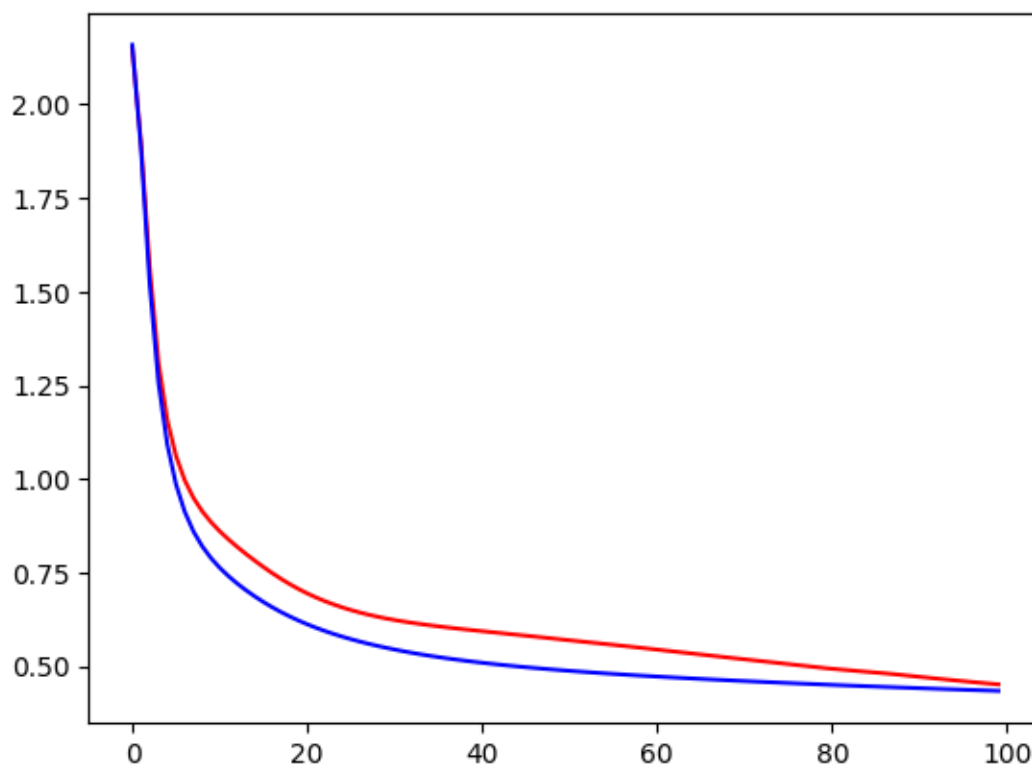
```
Got accuracy 84.4% in 100 epochs
Lr        Accuracy    Epoch
1         19.869999999999997% 100
0.1       85.0 % 4
0.01      85.11% 17
0.001     84.41% 100
```

After some investigation, it woas found that for a learning rate of 1, the accuracy was not able to increase to reach 85%. This can be seen in thee above running of code, where a limit of 100 epochs was chosen and if the accuracy had not reached 85% by this time, then the attempt was abandoned and the accuracy achieved was included in the table as it was found after 100 epochs. This may be the case because the learning rate was too high for the data, and the model is not able to converge. This can be seen in the outputs above for the learning rate of 1, where the loss does not reduce as the model is trained further.

1.1.4 Question 1.4

In comparing the results of the tables generated in Question 1.2 and Question 1.3, it is clear that as the learning rate decreases, the number of epochs required to reach a higher accuracy increases. However, with very high training rates, in this case, a training rate of 1, the model does not converge and the training loss does not reduce with more training of the model. The other training rates are able to successfully achieve an accuracy of over 85%, with the training with a higher learning rate able to achieve this in less epochs than when training at a lower learning rate. This indicates that the number of epochs to achieve a given accuracy and the learning rate are inversely proportional.

1.1.5 Question 1.5

Base model

```
Using cuda device
BaseNeuralNetwork(
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (linear_relu_stack): Sequential(
    (0): Linear(in_features=784, out_features=512, bias=True)
    (1): ReLU()
    (2): Linear(in_features=512, out_features=512, bias=True)
    (3): ReLU()
    (4): Linear(in_features=512, out_features=10, bias=True)
  )
)
Training for 10 epochs with learning rate: 0.001
Epoch 1
-------------------------------
loss: 2.284023  [    0/60000]
loss: 2.290835  [ 6400/60000]
loss: 2.265380  [12800/60000]
loss: 2.270940  [19200/60000]
```

```
loss: 2.243598  [25600/60000]
loss: 2.206037  [32000/60000]
loss: 2.221762  [38400/60000]
loss: 2.181980  [44800/60000]
loss: 2.179144  [51200/60000]
loss: 2.147452  [57600/60000]
Test Error:
 Accuracy: 40.1%, Avg loss: 2.142791

Epoch 2
-------------------------------
loss: 2.149441  [    0/60000]
loss: 2.148928  [ 6400/60000]
loss: 2.080436  [12800/60000]
loss: 2.106400  [19200/60000]
loss: 2.039692  [25600/60000]
loss: 1.974602  [32000/60000]
loss: 2.013869  [38400/60000]
loss: 1.931974  [44800/60000]
loss: 1.941559  [51200/60000]
loss: 1.859566  [57600/60000]
Test Error:
 Accuracy: 59.1%, Avg loss: 1.860429

Epoch 3
-------------------------------
loss: 1.893269  [    0/60000]
loss: 1.866668  [ 6400/60000]
loss: 1.738171  [12800/60000]
loss: 1.790199  [19200/60000]
loss: 1.670210  [25600/60000]
loss: 1.624490  [32000/60000]
loss: 1.650172  [38400/60000]
loss: 1.556689  [44800/60000]
loss: 1.581340  [51200/60000]
loss: 1.470580  [57600/60000]
Test Error:
 Accuracy: 63.0%, Avg loss: 1.491480

Epoch 4
-------------------------------
loss: 1.555546  [    0/60000]
loss: 1.528811  [ 6400/60000]
loss: 1.367295  [12800/60000]
loss: 1.445729  [19200/60000]
loss: 1.324501  [25600/60000]
loss: 1.327525  [32000/60000]
loss: 1.337986  [38400/60000]
```

```
loss: 1.272002  [44800/60000]
loss: 1.306239  [51200/60000]
loss: 1.204455  [57600/60000]
Test Error:
 Accuracy: 64.1%, Avg loss: 1.231390


Epoch 5
-------------------------------
loss: 1.303394  [    0/60000]
loss: 1.297739  [ 6400/60000]
loss: 1.118882  [12800/60000]
loss: 1.228072  [19200/60000]
loss: 1.102585  [25600/60000]
loss: 1.134993  [32000/60000]
loss: 1.149452  [38400/60000]
loss: 1.096211  [44800/60000]
loss: 1.135401  [51200/60000]
loss: 1.050895  [57600/60000]
Test Error:
 Accuracy: 64.9%, Avg loss: 1.072112


Epoch 6
-------------------------------
loss: 1.135810  [    0/60000]
loss: 1.153235  [ 6400/60000]
loss: 0.957845  [12800/60000]
loss: 1.094771  [19200/60000]
loss: 0.966929  [25600/60000]
loss: 1.006675  [32000/60000]
loss: 1.032828  [38400/60000]
loss: 0.984467  [44800/60000]
loss: 1.023184  [51200/60000]
loss: 0.954934  [57600/60000]
Test Error:
 Accuracy: 66.1%, Avg loss: 0.969634


Epoch 7
-------------------------------
loss: 1.019201  [    0/60000]
loss: 1.059063  [ 6400/60000]
loss: 0.847920  [12800/60000]
loss: 1.006525  [19200/60000]
loss: 0.881433  [25600/60000]
loss: 0.916493  [32000/60000]
loss: 0.955606  [38400/60000]
loss: 0.911418  [44800/60000]
loss: 0.944783  [51200/60000]
loss: 0.890385  [57600/60000]
```

```
Test Error:
 Accuracy: 67.6%, Avg loss: 0.899625


Epoch 8
-------------------------------
loss: 0.933820  [    0/60000]
loss: 0.993384  [ 6400/60000]
loss: 0.768984  [12800/60000]
loss: 0.944020  [19200/60000]
loss: 0.824278  [25600/60000]
loss: 0.850733  [32000/60000]
loss: 0.900565  [38400/60000]
loss: 0.862198  [44800/60000]
loss: 0.887918  [51200/60000]
loss: 0.843293  [57600/60000]
Test Error:
 Accuracy: 68.7%, Avg loss: 0.848962


Epoch 9
-------------------------------
loss: 0.868239  [    0/60000]
loss: 0.943619  [ 6400/60000]
loss: 0.709608  [12800/60000]
loss: 0.897078  [19200/60000]
loss: 0.783413  [25600/60000]
loss: 0.801366  [32000/60000]
loss: 0.858411  [38400/60000]
loss: 0.827493  [44800/60000]
loss: 0.844995  [51200/60000]
loss: 0.806761  [57600/60000]
Test Error:
 Accuracy: 70.0%, Avg loss: 0.810308


Epoch 10
-------------------------------
loss: 0.815688  [    0/60000]
loss: 0.903373  [ 6400/60000]
loss: 0.663068  [12800/60000]
loss: 0.860377  [19200/60000]
loss: 0.752300  [25600/60000]
loss: 0.763325  [32000/60000]
loss: 0.823931  [38400/60000]
loss: 0.801608  [44800/60000]
loss: 0.811533  [51200/60000]
loss: 0.776827  [57600/60000]
Test Error:
 Accuracy: 71.5%, Avg loss: 0.779326
```
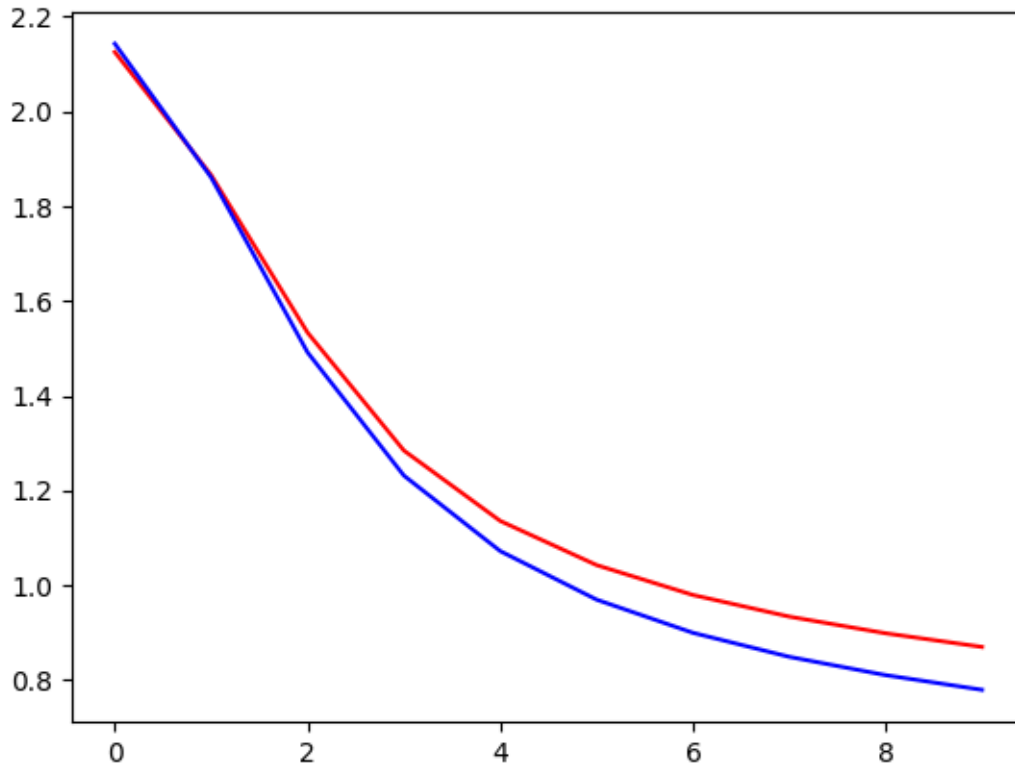
```
Model accuracy is: 0.715
Model has 669706 parameters
```

Wider model

```
Using cuda device
WiderNeuralNetwork(
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (linear_relu_stack): Sequential(
    (0): Linear(in_features=784, out_features=1024, bias=True)
    (1): ReLU()
    (2): Linear(in_features=1024, out_features=1024, bias=True)
    (3): ReLU()
    (4): Linear(in_features=1024, out_features=10, bias=True)
  )
)
Training for 10 epochs with learning rate: 0.001
Epoch 1
-------------------------------
loss: 2.294228  [    0/60000]
loss: 2.282108  [ 6400/60000]
loss: 2.250143  [12800/60000]
loss: 2.246608  [19200/60000]
```

```
loss: 2.225414  [25600/60000]
loss: 2.173778  [32000/60000]
loss: 2.183649  [38400/60000]
loss: 2.134665  [44800/60000]
loss: 2.140521  [51200/60000]
loss: 2.086738  [57600/60000]
Test Error:
 Accuracy: 50.7%, Avg loss: 2.084289

Epoch 2
-------------------------------
loss: 2.098476  [    0/60000]
loss: 2.086004  [ 6400/60000]
loss: 2.005016  [12800/60000]
loss: 2.032051  [19200/60000]
loss: 1.948710  [25600/60000]
loss: 1.881223  [32000/60000]
loss: 1.903079  [38400/60000]
loss: 1.806126  [44800/60000]
loss: 1.830235  [51200/60000]
loss: 1.723341  [57600/60000]
Test Error:
 Accuracy: 58.9%, Avg loss: 1.732546

Epoch 3
-------------------------------
loss: 1.776526  [    0/60000]
loss: 1.744135  [ 6400/60000]
loss: 1.603045  [12800/60000]
loss: 1.661202  [19200/60000]
loss: 1.522103  [25600/60000]
loss: 1.495981  [32000/60000]
loss: 1.502928  [38400/60000]
loss: 1.412930  [44800/60000]
loss: 1.452959  [51200/60000]
loss: 1.330293  [57600/60000]
Test Error:
 Accuracy: 63.3%, Avg loss: 1.360404

Epoch 4
-------------------------------
loss: 1.430295  [    0/60000]
loss: 1.409116  [ 6400/60000]
loss: 1.241252  [12800/60000]
loss: 1.340417  [19200/60000]
loss: 1.198205  [25600/60000]
loss: 1.216638  [32000/60000]
loss: 1.225267  [38400/60000]
```

```
loss: 1.158432  [44800/60000]
loss: 1.203346  [51200/60000]
loss: 1.100107  [57600/60000]
Test Error:
 Accuracy: 65.0%, Avg loss: 1.128813


Epoch 5
-------------------------------
loss: 1.195606  [    0/60000]
loss: 1.200710  [ 6400/60000]
loss: 1.013940  [12800/60000]
loss: 1.153067  [19200/60000]
loss: 1.008445  [25600/60000]
loss: 1.044112  [32000/60000]
loss: 1.069519  [38400/60000]
loss: 1.011380  [44800/60000]
loss: 1.057445  [51200/60000]
loss: 0.971611  [57600/60000]
Test Error:
 Accuracy: 66.3%, Avg loss: 0.992829


Epoch 6
-------------------------------
loss: 1.045995  [    0/60000]
loss: 1.076968  [ 6400/60000]
loss: 0.872110  [12800/60000]
loss: 1.040059  [19200/60000]
loss: 0.899170  [25600/60000]
loss: 0.932844  [32000/60000]
loss: 0.976365  [38400/60000]
loss: 0.923609  [44800/60000]
loss: 0.965178  [51200/60000]
loss: 0.892231  [57600/60000]
Test Error:
 Accuracy: 67.9%, Avg loss: 0.906997


Epoch 7
-------------------------------
loss: 0.943602  [    0/60000]
loss: 0.997531  [ 6400/60000]
loss: 0.777510  [12800/60000]
loss: 0.965532  [19200/60000]
loss: 0.831394  [25600/60000]
loss: 0.856613  [32000/60000]
loss: 0.914690  [38400/60000]
loss: 0.867946  [44800/60000]
loss: 0.902151  [51200/60000]
loss: 0.838382  [57600/60000]
```

```
Test Error:
 Accuracy: 69.0%, Avg loss: 0.848338


Epoch 8
-----------------------------
loss: 0.868571  [    0/60000]
loss: 0.940954  [ 6400/60000]
loss: 0.710481  [12800/60000]
loss: 0.912355  [19200/60000]
loss: 0.785823  [25600/60000]
loss: 0.801800  [32000/60000]
loss: 0.869935  [38400/60000]
loss: 0.830706  [44800/60000]
loss: 0.856643  [51200/60000]
loss: 0.798754  [57600/60000]
Test Error:
 Accuracy: 70.7%, Avg loss: 0.805432


Epoch 9
-----------------------------
loss: 0.810451  [    0/60000]
loss: 0.896896  [ 6400/60000]
loss: 0.659963  [12800/60000]
loss: 0.872424  [19200/60000]
loss: 0.752532  [25600/60000]
loss: 0.760932  [32000/60000]
loss: 0.834525  [38400/60000]
loss: 0.803977  [44800/60000]
loss: 0.821726  [51200/60000]
loss: 0.767446  [57600/60000]
Test Error:
 Accuracy: 71.9%, Avg loss: 0.771980


Epoch 10
-----------------------------
loss: 0.763295  [    0/60000]
loss: 0.860174  [ 6400/60000]
loss: 0.620027  [12800/60000]
loss: 0.840867  [19200/60000]
loss: 0.726688  [25600/60000]
loss: 0.729305  [32000/60000]
loss: 0.804698  [38400/60000]
loss: 0.783319  [44800/60000]
loss: 0.793652  [51200/60000]
loss: 0.741340  [57600/60000]
Test Error:
 Accuracy: 73.2%, Avg loss: 0.744473
```
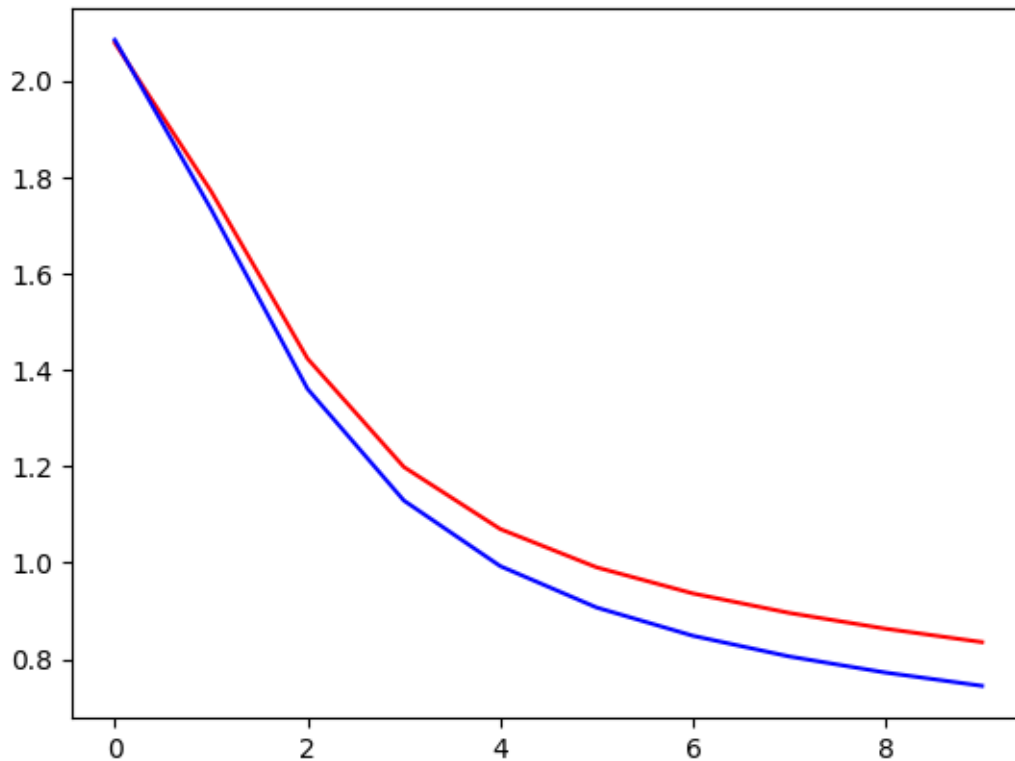
```
Model accuracy is: 0.7323
Model has 1863690 parameters
```

Deeper model

```
Using cuda device
DeeperNeuralNetwork(
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (linear_relu_stack): Sequential(
    (0): Linear(in_features=784, out_features=512, bias=True)
    (1): ReLU()
    (2): Linear(in_features=512, out_features=512, bias=True)
    (3): ReLU()
    (4): Linear(in_features=512, out_features=512, bias=True)
    (5): ReLU()
    (6): Linear(in_features=512, out_features=512, bias=True)
    (7): ReLU()
    (8): Linear(in_features=512, out_features=10, bias=True)
  )
)
Training for 10 epochs with learning rate: 0.001
Epoch 1
-------------------------------
```

```
loss: 2.308864  [    0/60000]
loss: 2.305411  [ 6400/60000]
loss: 2.304214  [12800/60000]
loss: 2.298974  [19200/60000]
loss: 2.305872  [25600/60000]
loss: 2.302269  [32000/60000]
loss: 2.297973  [38400/60000]
loss: 2.301029  [44800/60000]
loss: 2.302034  [51200/60000]
loss: 2.294982  [57600/60000]
Test Error:
 Accuracy: 10.9%, Avg loss: 2.298675


Epoch 2
-------------------------------
loss: 2.303227  [    0/60000]
loss: 2.300767  [ 6400/60000]
loss: 2.298394  [12800/60000]
loss: 2.294687  [19200/60000]
loss: 2.300640  [25600/60000]
loss: 2.295948  [32000/60000]
loss: 2.293504  [38400/60000]
loss: 2.294982  [44800/60000]
loss: 2.296152  [51200/60000]
loss: 2.289225  [57600/60000]
Test Error:
 Accuracy: 18.3%, Avg loss: 2.292671


Epoch 3
-------------------------------
loss: 2.297134  [    0/60000]
loss: 2.295576  [ 6400/60000]
loss: 2.291838  [12800/60000]
loss: 2.289579  [19200/60000]
loss: 2.294342  [25600/60000]
loss: 2.288192  [32000/60000]
loss: 2.287678  [38400/60000]
loss: 2.287238  [44800/60000]
loss: 2.288347  [51200/60000]
loss: 2.281404  [57600/60000]
Test Error:
 Accuracy: 32.1%, Avg loss: 2.284543


Epoch 4
-------------------------------
loss: 2.288868  [    0/60000]
loss: 2.288630  [ 6400/60000]
loss: 2.282728  [12800/60000]
```

```
loss: 2.282163  [19200/60000]
loss: 2.285621  [25600/60000]
loss: 2.276428  [32000/60000]
loss: 2.278496  [38400/60000]
loss: 2.275411  [44800/60000]
loss: 2.276236  [51200/60000]
loss: 2.268576  [57600/60000]
Test Error:
 Accuracy: 37.5%, Avg loss: 2.271610


Epoch 5
-------------------------------
loss: 2.275932  [    0/60000]
loss: 2.277426  [ 6400/60000]
loss: 2.267933  [12800/60000]
loss: 2.269371  [19200/60000]
loss: 2.271265  [25600/60000]
loss: 2.256174  [32000/60000]
loss: 2.262405  [38400/60000]
loss: 2.254585  [44800/60000]
loss: 2.254651  [51200/60000]
loss: 2.244685  [57600/60000]
Test Error:
 Accuracy: 36.9%, Avg loss: 2.248066


Epoch 6
-------------------------------
loss: 2.252879  [    0/60000]
loss: 2.256568  [ 6400/60000]
loss: 2.240787  [12800/60000]
loss: 2.244942  [19200/60000]
loss: 2.243276  [25600/60000]
loss: 2.217571  [32000/60000]
loss: 2.230319  [38400/60000]
loss: 2.213525  [44800/60000]
loss: 2.211298  [51200/60000]
loss: 2.195366  [57600/60000]
Test Error:
 Accuracy: 35.2%, Avg loss: 2.200115


Epoch 7
-------------------------------
loss: 2.207201  [    0/60000]
loss: 2.212687  [ 6400/60000]
loss: 2.184900  [12800/60000]
loss: 2.192368  [19200/60000]
loss: 2.182794  [25600/60000]
loss: 2.135795  [32000/60000]
```

```
loss: 2.159876  [38400/60000]
loss: 2.124000  [44800/60000]
loss: 2.116314  [51200/60000]
loss: 2.087010  [57600/60000]
Test Error:
 Accuracy: 34.5%, Avg loss: 2.095847


Epoch 8
-------------------------------
loss: 2.109983  [    0/60000]
loss: 2.113041  [ 6400/60000]
loss: 2.060632  [12800/60000]
loss: 2.073324  [19200/60000]
loss: 2.042734  [25600/60000]
loss: 1.972249  [32000/60000]
loss: 2.004576  [38400/60000]
loss: 1.938589  [44800/60000]
loss: 1.924908  [51200/60000]
loss: 1.869776  [57600/60000]
Test Error:
 Accuracy: 45.9%, Avg loss: 1.881745


Epoch 9
-------------------------------
loss: 1.918547  [    0/60000]
loss: 1.898183  [ 6400/60000]
loss: 1.799175  [12800/60000]
loss: 1.813821  [19200/60000]
loss: 1.739024  [25600/60000]
loss: 1.689274  [32000/60000]
loss: 1.699228  [38400/60000]
loss: 1.618214  [44800/60000]
loss: 1.625353  [51200/60000]
loss: 1.537084  [57600/60000]
Test Error:
 Accuracy: 44.0%, Avg loss: 1.548805


Epoch 10
-------------------------------
loss: 1.628103  [    0/60000]
loss: 1.583613  [ 6400/60000]
loss: 1.451048  [12800/60000]
loss: 1.486290  [19200/60000]
loss: 1.428170  [25600/60000]
loss: 1.412671  [32000/60000]
loss: 1.420581  [38400/60000]
loss: 1.359449  [44800/60000]
loss: 1.386141  [51200/60000]
```

```
loss: 1.309399  [57600/60000]
Test Error:
 Accuracy: 49.7%, Avg loss: 1.317051
```
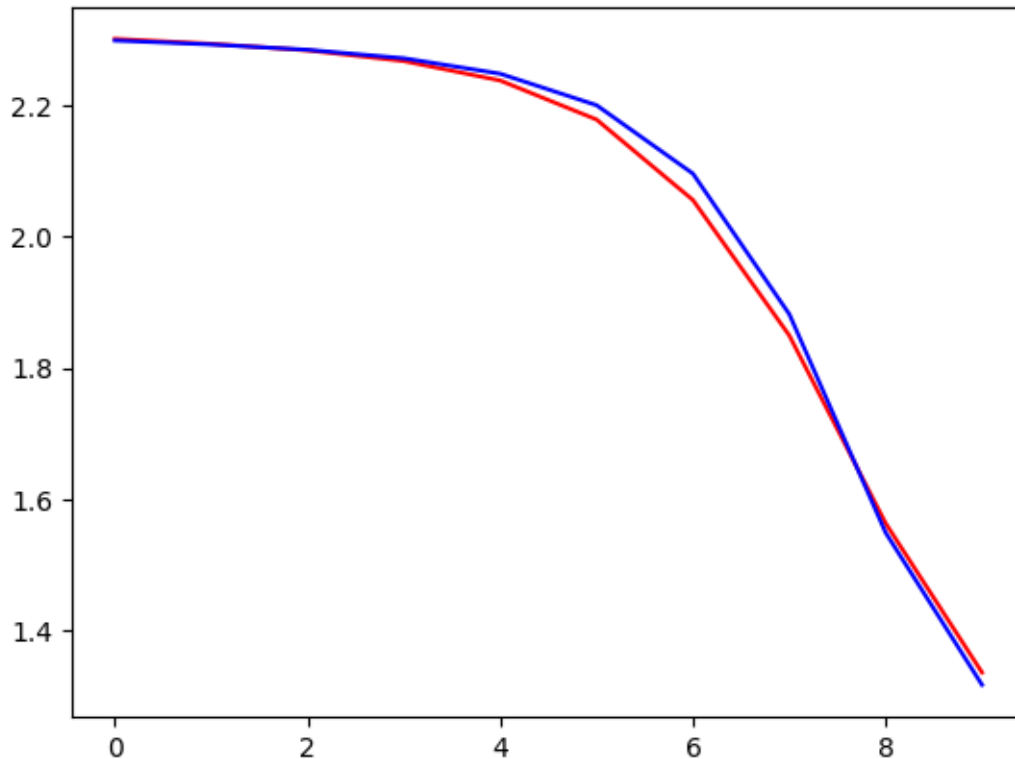


```
Model accuracy is: 0.497
Model has 1195018 parameters

Structures Accuracy    Parameters
base        71.5       % 669706
wider       73.22999999999999% 1863690
deeper      49.7       % 1195018
```

From the results above, the wider network structure was able to achieve a slightly better accuracy than the base model, but the deeper model actually had an accuracy reduction. It is possible that the network needs more training as it now has more parameters, and this should be tested to see if that is the case. We can test training the same deeper model but with more epochs to see if the accuracy improves.

Deeper model with more epochs

```
Training for 50 epochs with learning rate: 0.001
Epoch 1
-------------------------------
```

```
loss: 2.306122  [    0/60000]
loss: 2.302324  [ 6400/60000]
loss: 2.298337  [12800/60000]
loss: 2.298876  [19200/60000]
loss: 2.298498  [25600/60000]
loss: 2.297678  [32000/60000]
loss: 2.302513  [38400/60000]
loss: 2.299364  [44800/60000]
loss: 2.301500  [51200/60000]
loss: 2.297606  [57600/60000]
Test Error:
 Accuracy: 14.5%, Avg loss: 2.296031


Epoch 2
-------------------------------
loss: 2.299687  [    0/60000]
loss: 2.296878  [ 6400/60000]
loss: 2.291625  [12800/60000]
loss: 2.293523  [19200/60000]
loss: 2.292868  [25600/60000]
loss: 2.290528  [32000/60000]
loss: 2.296480  [38400/60000]
loss: 2.292460  [44800/60000]
loss: 2.294390  [51200/60000]
loss: 2.290358  [57600/60000]
Test Error:
 Accuracy: 22.5%, Avg loss: 2.288619


Epoch 3
-------------------------------
loss: 2.292092  [    0/60000]
loss: 2.290196  [ 6400/60000]
loss: 2.283271  [12800/60000]
loss: 2.286692  [19200/60000]
loss: 2.285498  [25600/60000]
loss: 2.280947  [32000/60000]
loss: 2.288521  [38400/60000]
loss: 2.282939  [44800/60000]
loss: 2.284654  [51200/60000]
loss: 2.280222  [57600/60000]
Test Error:
 Accuracy: 32.7%, Avg loss: 2.278043


Epoch 4
-------------------------------
loss: 2.281410  [    0/60000]
loss: 2.280589  [ 6400/60000]
loss: 2.271055  [12800/60000]
```

```
loss: 2.276355  [19200/60000]
loss: 2.274150  [25600/60000]
loss: 2.265912  [32000/60000]
loss: 2.275970  [38400/60000]
loss: 2.267357  [44800/60000]
loss: 2.268623  [51200/60000]
loss: 2.263568  [57600/60000]
Test Error:
 Accuracy: 34.1%, Avg loss: 2.260629


Epoch 5
-------------------------------
loss: 2.264105  [    0/60000]
loss: 2.264619  [ 6400/60000]
loss: 2.250670  [12800/60000]
loss: 2.258639  [19200/60000]
loss: 2.254504  [25600/60000]
loss: 2.239305  [32000/60000]
loss: 2.253822  [38400/60000]
loss: 2.239129  [44800/60000]
loss: 2.239439  [51200/60000]
loss: 2.231955  [57600/60000]
Test Error:
 Accuracy: 30.5%, Avg loss: 2.228109


Epoch 6
-------------------------------
loss: 2.232272  [    0/60000]
loss: 2.234218  [ 6400/60000]
loss: 2.211771  [12800/60000]
loss: 2.223506  [19200/60000]
loss: 2.215447  [25600/60000]
loss: 2.186414  [32000/60000]
loss: 2.209352  [38400/60000]
loss: 2.181425  [44800/60000]
loss: 2.178585  [51200/60000]
loss: 2.165749  [57600/60000]
Test Error:
 Accuracy: 29.6%, Avg loss: 2.160384


Epoch 7
-------------------------------
loss: 2.167090  [    0/60000]
loss: 2.168939  [ 6400/60000]
loss: 2.128812  [12800/60000]
loss: 2.146444  [19200/60000]
loss: 2.128064  [25600/60000]
loss: 2.074512  [32000/60000]
```

```
loss: 2.109255  [38400/60000]
loss: 2.053948  [44800/60000]
loss: 2.045877  [51200/60000]
loss: 2.022367  [57600/60000]
Test Error:
 Accuracy: 38.8%, Avg loss: 2.013637


Epoch 8
-------------------------------
loss: 2.029922  [    0/60000]
loss: 2.022650  [ 6400/60000]
loss: 1.945439  [12800/60000]
loss: 1.971844  [19200/60000]
loss: 1.924621  [25600/60000]
loss: 1.856316  [32000/60000]
loss: 1.891286  [38400/60000]
loss: 1.798473  [44800/60000]
loss: 1.799883  [51200/60000]
loss: 1.751824  [57600/60000]
Test Error:
 Accuracy: 42.6%, Avg loss: 1.734861


Epoch 9
-------------------------------
loss: 1.783302  [    0/60000]
loss: 1.750264  [ 6400/60000]
loss: 1.619721  [12800/60000]
loss: 1.664525  [19200/60000]
loss: 1.597832  [25600/60000]
loss: 1.553546  [32000/60000]
loss: 1.577017  [38400/60000]
loss: 1.488329  [44800/60000]
loss: 1.517901  [51200/60000]
loss: 1.445644  [57600/60000]
Test Error:
 Accuracy: 43.3%, Avg loss: 1.445689


Epoch 10
-------------------------------
loss: 1.523915  [    0/60000]
loss: 1.487930  [ 6400/60000]
loss: 1.342559  [12800/60000]
loss: 1.412391  [19200/60000]
loss: 1.356689  [25600/60000]
loss: 1.345698  [32000/60000]
loss: 1.359540  [38400/60000]
loss: 1.297171  [44800/60000]
loss: 1.324786  [51200/60000]
```

```
loss: 1.262741  [57600/60000]
Test Error:
 Accuracy: 51.4%, Avg loss: 1.267291


Epoch 11
-------------------------------
loss: 1.352653  [    0/60000]
loss: 1.327062  [ 6400/60000]
loss: 1.166540  [12800/60000]
loss: 1.262613  [19200/60000]
loss: 1.197061  [25600/60000]
loss: 1.214414  [32000/60000]
loss: 1.237640  [38400/60000]
loss: 1.180329  [44800/60000]
loss: 1.205886  [51200/60000]
loss: 1.154961  [57600/60000]
Test Error:
 Accuracy: 55.5%, Avg loss: 1.155002


Epoch 12
-------------------------------
loss: 1.242293  [    0/60000]
loss: 1.228773  [ 6400/60000]
loss: 1.043730  [12800/60000]
loss: 1.163547  [19200/60000]
loss: 1.081263  [25600/60000]
loss: 1.117114  [32000/60000]
loss: 1.155961  [38400/60000]
loss: 1.097171  [44800/60000]
loss: 1.116146  [51200/60000]
loss: 1.078036  [57600/60000]
Test Error:
 Accuracy: 57.8%, Avg loss: 1.073490


Epoch 13
-------------------------------
loss: 1.154626  [    0/60000]
loss: 1.159131  [ 6400/60000]
loss: 0.951020  [12800/60000]
loss: 1.089053  [19200/60000]
loss: 0.991610  [25600/60000]
loss: 1.035482  [32000/60000]
loss: 1.090563  [38400/60000]
loss: 1.031257  [44800/60000]
loss: 1.039475  [51200/60000]
loss: 1.016046  [57600/60000]
Test Error:
 Accuracy: 61.3%, Avg loss: 1.007990
```

```
Epoch 14
-------------------------------
loss: 1.077340  [    0/60000]
loss: 1.101772  [ 6400/60000]
loss: 0.878248  [12800/60000]
loss: 1.026296  [19200/60000]
loss: 0.923819  [25600/60000]
loss: 0.960708  [32000/60000]
loss: 1.033873  [38400/60000]
loss: 0.977396  [44800/60000]
loss: 0.970152  [51200/60000]
loss: 0.966169  [57600/60000]
Test Error:
 Accuracy: 64.3%, Avg loss: 0.953208

Epoch 15
-------------------------------
loss: 1.007594  [    0/60000]
loss: 1.049870  [ 6400/60000]
loss: 0.819603  [12800/60000]
loss: 0.971342  [19200/60000]
loss: 0.875097  [25600/60000]
loss: 0.893968  [32000/60000]
loss: 0.986760  [38400/60000]
loss: 0.935234  [44800/60000]
loss: 0.910296  [51200/60000]
loss: 0.928349  [57600/60000]
Test Error:
 Accuracy: 66.2%, Avg loss: 0.909262

Epoch 16
-------------------------------
loss: 0.948586  [    0/60000]
loss: 1.004184  [ 6400/60000]
loss: 0.772376  [12800/60000]
loss: 0.924821  [19200/60000]
loss: 0.840641  [25600/60000]
loss: 0.838737  [32000/60000]
loss: 0.949472  [38400/60000]
loss: 0.903787  [44800/60000]
loss: 0.862694  [51200/60000]
loss: 0.901117  [57600/60000]
Test Error:
 Accuracy: 67.4%, Avg loss: 0.875718

Epoch 17
-------------------------------
```

```
loss: 0.901775  [    0/60000]
loss: 0.966237  [ 6400/60000]
loss: 0.734105  [12800/60000]
loss: 0.888489  [19200/60000]
loss: 0.814510  [25600/60000]
loss: 0.795238  [32000/60000]
loss: 0.920913  [38400/60000]
loss: 0.879435  [44800/60000]
loss: 0.827484  [51200/60000]
loss: 0.881058  [57600/60000]
Test Error:
 Accuracy: 68.4%, Avg loss: 0.850340


Epoch 18
-------------------------------
loss: 0.864785  [    0/60000]
loss: 0.936361  [ 6400/60000]
loss: 0.702215  [12800/60000]
loss: 0.861315  [19200/60000]
loss: 0.792852  [25600/60000]
loss: 0.760342  [32000/60000]
loss: 0.898786  [38400/60000]
loss: 0.859002  [44800/60000]
loss: 0.801129  [51200/60000]
loss: 0.864530  [57600/60000]
Test Error:
 Accuracy: 69.3%, Avg loss: 0.830026


Epoch 19
-------------------------------
loss: 0.834160  [    0/60000]
loss: 0.912166  [ 6400/60000]
loss: 0.674668  [12800/60000]
loss: 0.840026  [19200/60000]
loss: 0.773553  [25600/60000]
loss: 0.730873  [32000/60000]
loss: 0.880459  [38400/60000]
loss: 0.840918  [44800/60000]
loss: 0.780594  [51200/60000]
loss: 0.849166  [57600/60000]
Test Error:
 Accuracy: 70.2%, Avg loss: 0.812432


Epoch 20
-------------------------------
loss: 0.807236  [    0/60000]
loss: 0.891442  [ 6400/60000]
loss: 0.650463  [12800/60000]
```

```
loss: 0.822276  [19200/60000]
loss: 0.755365  [25600/60000]
loss: 0.704507  [32000/60000]
loss: 0.864304  [38400/60000]
loss: 0.824860  [44800/60000]
loss: 0.763973  [51200/60000]
loss: 0.833845  [57600/60000]
Test Error:
 Accuracy: 71.1%, Avg loss: 0.796196


Epoch 21
-------------------------------
loss: 0.782547  [    0/60000]
loss: 0.872811  [ 6400/60000]
loss: 0.628453  [12800/60000]
loss: 0.806294  [19200/60000]
loss: 0.738005  [25600/60000]
loss: 0.680306  [32000/60000]
loss: 0.849068  [38400/60000]
loss: 0.810279  [44800/60000]
loss: 0.749507  [51200/60000]
loss: 0.818249  [57600/60000]
Test Error:
 Accuracy: 71.7%, Avg loss: 0.780572


Epoch 22
-------------------------------
loss: 0.759161  [    0/60000]
loss: 0.855387  [ 6400/60000]
loss: 0.607955  [12800/60000]
loss: 0.791616  [19200/60000]
loss: 0.721758  [25600/60000]
loss: 0.657569  [32000/60000]
loss: 0.834254  [38400/60000]
loss: 0.796822  [44800/60000]
loss: 0.736467  [51200/60000]
loss: 0.801973  [57600/60000]
Test Error:
 Accuracy: 72.4%, Avg loss: 0.765100


Epoch 23
-------------------------------
loss: 0.736595  [    0/60000]
loss: 0.838498  [ 6400/60000]
loss: 0.588450  [12800/60000]
loss: 0.777942  [19200/60000]
loss: 0.706268  [25600/60000]
loss: 0.635870  [32000/60000]
```

```
loss: 0.819342  [38400/60000]
loss: 0.784271  [44800/60000]
loss: 0.724215  [51200/60000]
loss: 0.784922  [57600/60000]
Test Error:
 Accuracy: 73.1%, Avg loss: 0.749538


Epoch 24
-------------------------------
loss: 0.714381  [    0/60000]
loss: 0.822027  [ 6400/60000]
loss: 0.569857  [12800/60000]
loss: 0.764798  [19200/60000]
loss: 0.691474  [25600/60000]
loss: 0.615320  [32000/60000]
loss: 0.804180  [38400/60000]
loss: 0.772915  [44800/60000]
loss: 0.712534  [51200/60000]
loss: 0.767355  [57600/60000]
Test Error:
 Accuracy: 73.5%, Avg loss: 0.733994


Epoch 25
-------------------------------
loss: 0.692518  [    0/60000]
loss: 0.806192  [ 6400/60000]
loss: 0.552165  [12800/60000]
loss: 0.752313  [19200/60000]
loss: 0.677679  [25600/60000]
loss: 0.596100  [32000/60000]
loss: 0.788636  [38400/60000]
loss: 0.762620  [44800/60000]
loss: 0.701335  [51200/60000]
loss: 0.750021  [57600/60000]
Test Error:
 Accuracy: 74.0%, Avg loss: 0.718616


Epoch 26
-------------------------------
loss: 0.671017  [    0/60000]
loss: 0.790676  [ 6400/60000]
loss: 0.535506  [12800/60000]
loss: 0.740331  [19200/60000]
loss: 0.664943  [25600/60000]
loss: 0.578322  [32000/60000]
loss: 0.772950  [38400/60000]
loss: 0.753199  [44800/60000]
loss: 0.691031  [51200/60000]
```

```
loss: 0.733065  [57600/60000]
Test Error:
 Accuracy: 74.7%, Avg loss: 0.703705


Epoch 27
-------------------------------
loss: 0.649934  [    0/60000]
loss: 0.775592  [ 6400/60000]
loss: 0.520100  [12800/60000]
loss: 0.728721  [19200/60000]
loss: 0.653201  [25600/60000]
loss: 0.562532  [32000/60000]
loss: 0.757200  [38400/60000]
loss: 0.744808  [44800/60000]
loss: 0.681416  [51200/60000]
loss: 0.717162  [57600/60000]
Test Error:
 Accuracy: 75.3%, Avg loss: 0.689523


Epoch 28
-------------------------------
loss: 0.629660  [    0/60000]
loss: 0.761243  [ 6400/60000]
loss: 0.506144  [12800/60000]
loss: 0.717210  [19200/60000]
loss: 0.642926  [25600/60000]
loss: 0.548945  [32000/60000]
loss: 0.741293  [38400/60000]
loss: 0.737261  [44800/60000]
loss: 0.672617  [51200/60000]
loss: 0.702457  [57600/60000]
Test Error:
 Accuracy: 75.8%, Avg loss: 0.676333


Epoch 29
-------------------------------
loss: 0.610807  [    0/60000]
loss: 0.747840  [ 6400/60000]
loss: 0.493583  [12800/60000]
loss: 0.705898  [19200/60000]
loss: 0.633891  [25600/60000]
loss: 0.537287  [32000/60000]
loss: 0.725714  [38400/60000]
loss: 0.730368  [44800/60000]
loss: 0.664713  [51200/60000]
loss: 0.689077  [57600/60000]
Test Error:
 Accuracy: 76.0%, Avg loss: 0.664105
```

```
Epoch 30
-------------------------------
loss: 0.593529  [    0/60000]
loss: 0.735634  [ 6400/60000]
loss: 0.482178  [12800/60000]
loss: 0.694763  [19200/60000]
loss: 0.626102  [25600/60000]
loss: 0.527600  [32000/60000]
loss: 0.710600  [38400/60000]
loss: 0.724381  [44800/60000]
loss: 0.657442  [51200/60000]
loss: 0.676704  [57600/60000]
Test Error:
 Accuracy: 76.4%, Avg loss: 0.652898

Epoch 31
-------------------------------
loss: 0.578105  [    0/60000]
loss: 0.724276  [ 6400/60000]
loss: 0.472014  [12800/60000]
loss: 0.683844  [19200/60000]
loss: 0.619983  [25600/60000]
loss: 0.519652  [32000/60000]
loss: 0.695616  [38400/60000]
loss: 0.718741  [44800/60000]
loss: 0.650779  [51200/60000]
loss: 0.665157  [57600/60000]
Test Error:
 Accuracy: 76.8%, Avg loss: 0.642701

Epoch 32
-------------------------------
loss: 0.563749  [    0/60000]
loss: 0.713810  [ 6400/60000]
loss: 0.462811  [12800/60000]
loss: 0.673465  [19200/60000]
loss: 0.614912  [25600/60000]
loss: 0.513277  [32000/60000]
loss: 0.681183  [38400/60000]
loss: 0.713297  [44800/60000]
loss: 0.644355  [51200/60000]
loss: 0.654710  [57600/60000]
Test Error:
 Accuracy: 77.3%, Avg loss: 0.633407

Epoch 33
-------------------------------
```

```
loss: 0.550839  [    0/60000]
loss: 0.704265  [ 6400/60000]
loss: 0.454300  [12800/60000]
loss: 0.663350  [19200/60000]
loss: 0.610819  [25600/60000]
loss: 0.508150  [32000/60000]
loss: 0.667375  [38400/60000]
loss: 0.708490  [44800/60000]
loss: 0.638419  [51200/60000]
loss: 0.645172  [57600/60000]
Test Error:
 Accuracy: 77.7%, Avg loss: 0.624922


Epoch 34
-------------------------------
loss: 0.538917  [    0/60000]
loss: 0.695499  [ 6400/60000]
loss: 0.446397  [12800/60000]
loss: 0.653513  [19200/60000]
loss: 0.607476  [25600/60000]
loss: 0.503948  [32000/60000]
loss: 0.654201  [38400/60000]
loss: 0.704106  [44800/60000]
loss: 0.633003  [51200/60000]
loss: 0.636184  [57600/60000]
Test Error:
 Accuracy: 78.0%, Avg loss: 0.617137


Epoch 35
-------------------------------
loss: 0.528078  [    0/60000]
loss: 0.687496  [ 6400/60000]
loss: 0.439091  [12800/60000]
loss: 0.644008  [19200/60000]
loss: 0.604467  [25600/60000]
loss: 0.500318  [32000/60000]
loss: 0.641796  [38400/60000]
loss: 0.700179  [44800/60000]
loss: 0.628062  [51200/60000]
loss: 0.627720  [57600/60000]
Test Error:
 Accuracy: 78.3%, Avg loss: 0.609993


Epoch 36
-------------------------------
loss: 0.518074  [    0/60000]
loss: 0.680085  [ 6400/60000]
loss: 0.432359  [12800/60000]
```

```
loss: 0.634908  [19200/60000]
loss: 0.601645  [25600/60000]
loss: 0.497187  [32000/60000]
loss: 0.630043  [38400/60000]
loss: 0.696826  [44800/60000]
loss: 0.623565  [51200/60000]
loss: 0.619728  [57600/60000]
Test Error:
 Accuracy: 78.5%, Avg loss: 0.603351


Epoch 37
-------------------------------
loss: 0.508850  [    0/60000]
loss: 0.673258  [ 6400/60000]
loss: 0.426279  [12800/60000]
loss: 0.626269  [19200/60000]
loss: 0.599125  [25600/60000]
loss: 0.494407  [32000/60000]
loss: 0.618831  [38400/60000]
loss: 0.693839  [44800/60000]
loss: 0.619322  [51200/60000]
loss: 0.612080  [57600/60000]
Test Error:
 Accuracy: 78.8%, Avg loss: 0.597170


Epoch 38
-------------------------------
loss: 0.499945  [    0/60000]
loss: 0.666881  [ 6400/60000]
loss: 0.420747  [12800/60000]
loss: 0.618004  [19200/60000]
loss: 0.596611  [25600/60000]
loss: 0.491839  [32000/60000]
loss: 0.608018  [38400/60000]
loss: 0.691193  [44800/60000]
loss: 0.615308  [51200/60000]
loss: 0.604801  [57600/60000]
Test Error:
 Accuracy: 79.0%, Avg loss: 0.591356


Epoch 39
-------------------------------
loss: 0.491446  [    0/60000]
loss: 0.660997  [ 6400/60000]
loss: 0.415847  [12800/60000]
loss: 0.610229  [19200/60000]
loss: 0.594264  [25600/60000]
loss: 0.489359  [32000/60000]
```

```
loss: 0.597889  [38400/60000]
loss: 0.688754  [44800/60000]
loss: 0.611524  [51200/60000]
loss: 0.597871  [57600/60000]
Test Error:
 Accuracy: 79.2%, Avg loss: 0.585928


Epoch 40
-------------------------------
loss: 0.483303  [    0/60000]
loss: 0.655494  [ 6400/60000]
loss: 0.411256  [12800/60000]
loss: 0.602872  [19200/60000]
loss: 0.591780  [25600/60000]
loss: 0.486836  [32000/60000]
loss: 0.588195  [38400/60000]
loss: 0.686732  [44800/60000]
loss: 0.608236  [51200/60000]
loss: 0.591368  [57600/60000]
Test Error:
 Accuracy: 79.4%, Avg loss: 0.580833


Epoch 41
-------------------------------
loss: 0.475491  [    0/60000]
loss: 0.650330  [ 6400/60000]
loss: 0.407075  [12800/60000]
loss: 0.595985  [19200/60000]
loss: 0.589398  [25600/60000]
loss: 0.484296  [32000/60000]
loss: 0.579070  [38400/60000]
loss: 0.685230  [44800/60000]
loss: 0.605256  [51200/60000]
loss: 0.585142  [57600/60000]
Test Error:
 Accuracy: 79.6%, Avg loss: 0.576029


Epoch 42
-------------------------------
loss: 0.468047  [    0/60000]
loss: 0.645521  [ 6400/60000]
loss: 0.403215  [12800/60000]
loss: 0.589711  [19200/60000]
loss: 0.586836  [25600/60000]
loss: 0.481520  [32000/60000]
loss: 0.570493  [38400/60000]
loss: 0.684234  [44800/60000]
loss: 0.602743  [51200/60000]
```

```
loss: 0.578953  [57600/60000]
Test Error:
 Accuracy: 79.8%, Avg loss: 0.571501


Epoch 43
-------------------------------
loss: 0.461098  [    0/60000]
loss: 0.641064  [ 6400/60000]
loss: 0.399692  [12800/60000]
loss: 0.583801  [19200/60000]
loss: 0.583525  [25600/60000]
loss: 0.478781  [32000/60000]
loss: 0.562658  [38400/60000]
loss: 0.683396  [44800/60000]
loss: 0.600530  [51200/60000]
loss: 0.573478  [57600/60000]
Test Error:
 Accuracy: 80.0%, Avg loss: 0.567236


Epoch 44
-------------------------------
loss: 0.454356  [    0/60000]
loss: 0.636822  [ 6400/60000]
loss: 0.396659  [12800/60000]
loss: 0.577901  [19200/60000]
loss: 0.580143  [25600/60000]
loss: 0.476183  [32000/60000]
loss: 0.554837  [38400/60000]
loss: 0.682870  [44800/60000]
loss: 0.598588  [51200/60000]
loss: 0.568286  [57600/60000]
Test Error:
 Accuracy: 80.1%, Avg loss: 0.563164


Epoch 45
-------------------------------
loss: 0.447800  [    0/60000]
loss: 0.632550  [ 6400/60000]
loss: 0.393864  [12800/60000]
loss: 0.572657  [19200/60000]
loss: 0.576576  [25600/60000]
loss: 0.473569  [32000/60000]
loss: 0.547286  [38400/60000]
loss: 0.682229  [44800/60000]
loss: 0.596418  [51200/60000]
loss: 0.563273  [57600/60000]
Test Error:
 Accuracy: 80.2%, Avg loss: 0.559284
```

```
Epoch 46
-------------------------------
loss: 0.441505  [    0/60000]
loss: 0.628285  [ 6400/60000]
loss: 0.391154  [12800/60000]
loss: 0.567797  [19200/60000]
loss: 0.572780  [25600/60000]
loss: 0.470926  [32000/60000]
loss: 0.540025  [38400/60000]
loss: 0.681747  [44800/60000]
loss: 0.594529  [51200/60000]
loss: 0.558585  [57600/60000]
Test Error:
 Accuracy: 80.3%, Avg loss: 0.555566

Epoch 47
-------------------------------
loss: 0.435469  [    0/60000]
loss: 0.624192  [ 6400/60000]
loss: 0.388531  [12800/60000]
loss: 0.563237  [19200/60000]
loss: 0.568618  [25600/60000]
loss: 0.468432  [32000/60000]
loss: 0.532986  [38400/60000]
loss: 0.681074  [44800/60000]
loss: 0.592820  [51200/60000]
loss: 0.553898  [57600/60000]
Test Error:
 Accuracy: 80.4%, Avg loss: 0.552023

Epoch 48
-------------------------------
loss: 0.429845  [    0/60000]
loss: 0.620076  [ 6400/60000]
loss: 0.386105  [12800/60000]
loss: 0.558812  [19200/60000]
loss: 0.564571  [25600/60000]
loss: 0.465895  [32000/60000]
loss: 0.526224  [38400/60000]
loss: 0.680504  [44800/60000]
loss: 0.591107  [51200/60000]
loss: 0.549362  [57600/60000]
Test Error:
 Accuracy: 80.5%, Avg loss: 0.548645

Epoch 49
-------------------------------
```

```
loss: 0.424296  [    0/60000]
loss: 0.616190  [ 6400/60000]
loss: 0.383679  [12800/60000]
loss: 0.554370  [19200/60000]
loss: 0.560642  [25600/60000]
loss: 0.463336  [32000/60000]
loss: 0.519838  [38400/60000]
loss: 0.680078  [44800/60000]
loss: 0.589292  [51200/60000]
loss: 0.545156  [57600/60000]
Test Error:
 Accuracy: 80.5%, Avg loss: 0.545368

Epoch 50
-------------------------------
loss: 0.418958  [    0/60000]
loss: 0.612306  [ 6400/60000]
loss: 0.381352  [12800/60000]
loss: 0.550164  [19200/60000]
loss: 0.556548  [25600/60000]
loss: 0.460808  [32000/60000]
loss: 0.513494  [38400/60000]
loss: 0.679798  [44800/60000]
loss: 0.587409  [51200/60000]
loss: 0.541127  [57600/60000]
Test Error:
 Accuracy: 80.6%, Avg loss: 0.542195
```
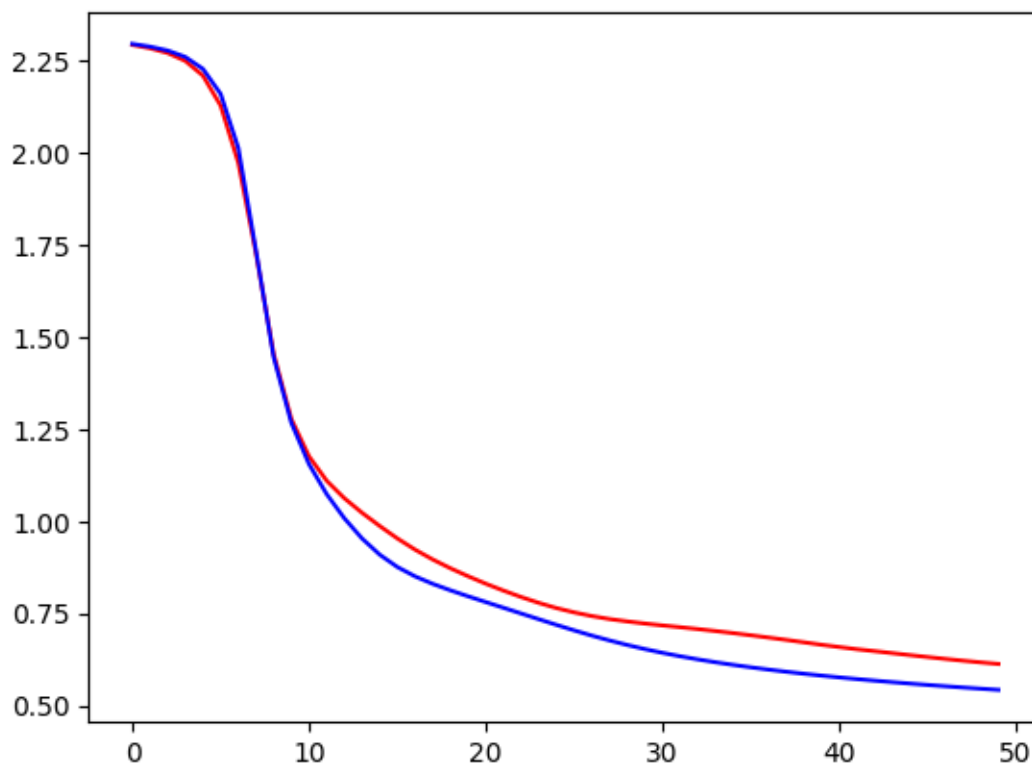
```
Deeper model accuracy is: 0.8062
Training for 50 epochs with learning rate: 0.001
Epoch 1
-------------------------------
loss: 2.303873  [    0/60000]
loss: 2.292670  [ 6400/60000]
loss: 2.281146  [12800/60000]
loss: 2.271633  [19200/60000]
loss: 2.256431  [25600/60000]
loss: 2.224319  [32000/60000]
loss: 2.234491  [38400/60000]
loss: 2.198293  [44800/60000]
loss: 2.193406  [51200/60000]
loss: 2.169432  [57600/60000]
Test Error:
 Accuracy: 46.5%, Avg loss: 2.163906

Epoch 2
-------------------------------
loss: 2.171407  [    0/60000]
loss: 2.161674  [ 6400/60000]
loss: 2.115687  [12800/60000]
```

```
loss: 2.130499  [19200/60000]
loss: 2.078270  [25600/60000]
loss: 2.022399  [32000/60000]
loss: 2.050585  [38400/60000]
loss: 1.973611  [44800/60000]
loss: 1.977142  [51200/60000]
loss: 1.915118  [57600/60000]
Test Error:
 Accuracy: 59.2%, Avg loss: 1.910656


Epoch 3
-------------------------------
loss: 1.940034  [    0/60000]
loss: 1.907181  [ 6400/60000]
loss: 1.806028  [12800/60000]
loss: 1.842938  [19200/60000]
loss: 1.730339  [25600/60000]
loss: 1.684275  [32000/60000]
loss: 1.706461  [38400/60000]
loss: 1.606690  [44800/60000]
loss: 1.627955  [51200/60000]
loss: 1.527157  [57600/60000]
Test Error:
 Accuracy: 60.8%, Avg loss: 1.541754


Epoch 4
-------------------------------
loss: 1.609476  [    0/60000]
loss: 1.562255  [ 6400/60000]
loss: 1.427720  [12800/60000]
loss: 1.492257  [19200/60000]
loss: 1.370719  [25600/60000]
loss: 1.365842  [32000/60000]
loss: 1.379410  [38400/60000]
loss: 1.304386  [44800/60000]
loss: 1.335701  [51200/60000]
loss: 1.235417  [57600/60000]
Test Error:
 Accuracy: 63.0%, Avg loss: 1.262464


Epoch 5
-------------------------------
loss: 1.344014  [    0/60000]
loss: 1.310780  [ 6400/60000]
loss: 1.161917  [12800/60000]
loss: 1.259528  [19200/60000]
loss: 1.133491  [25600/60000]
loss: 1.157936  [32000/60000]
```

```
loss: 1.178953  [38400/60000]
loss: 1.119257  [44800/60000]
loss: 1.154473  [51200/60000]
loss: 1.070377  [57600/60000]
Test Error:
 Accuracy: 64.6%, Avg loss: 1.091809


Epoch 6
-------------------------------
loss: 1.167542  [    0/60000]
loss: 1.154522  [ 6400/60000]
loss: 0.990726  [12800/60000]
loss: 1.117895  [19200/60000]
loss: 0.988400  [25600/60000]
loss: 1.022370  [32000/60000]
loss: 1.058441  [38400/60000]
loss: 1.005863  [44800/60000]
loss: 1.039830  [51200/60000]
loss: 0.970029  [57600/60000]
Test Error:
 Accuracy: 65.9%, Avg loss: 0.983853


Epoch 7
-------------------------------
loss: 1.047385  [    0/60000]
loss: 1.055334  [ 6400/60000]
loss: 0.875890  [12800/60000]
loss: 1.024784  [19200/60000]
loss: 0.898206  [25600/60000]
loss: 0.928615  [32000/60000]
loss: 0.981039  [38400/60000]
loss: 0.933747  [44800/60000]
loss: 0.961509  [51200/60000]
loss: 0.903657  [57600/60000]
Test Error:
 Accuracy: 67.2%, Avg loss: 0.911069


Epoch 8
-------------------------------
loss: 0.960051  [    0/60000]
loss: 0.987547  [ 6400/60000]
loss: 0.794683  [12800/60000]
loss: 0.959304  [19200/60000]
loss: 0.838598  [25600/60000]
loss: 0.860607  [32000/60000]
loss: 0.927812  [38400/60000]
loss: 0.886178  [44800/60000]
loss: 0.905227  [51200/60000]
```

```
loss: 0.856294  [57600/60000]
Test Error:
 Accuracy: 68.2%, Avg loss: 0.858987


Epoch 9
-------------------------------
loss: 0.893278  [    0/60000]
loss: 0.936911  [ 6400/60000]
loss: 0.734503  [12800/60000]
loss: 0.910322  [19200/60000]
loss: 0.796240  [25600/60000]
loss: 0.809208  [32000/60000]
loss: 0.887725  [38400/60000]
loss: 0.853156  [44800/60000]
loss: 0.862515  [51200/60000]
loss: 0.820264  [57600/60000]
Test Error:
 Accuracy: 69.4%, Avg loss: 0.819476


Epoch 10
-------------------------------
loss: 0.839695  [    0/60000]
loss: 0.896534  [ 6400/60000]
loss: 0.687726  [12800/60000]
loss: 0.872095  [19200/60000]
loss: 0.764598  [25600/60000]
loss: 0.769333  [32000/60000]
loss: 0.855130  [38400/60000]
loss: 0.828629  [44800/60000]
loss: 0.828728  [51200/60000]
loss: 0.791319  [57600/60000]
Test Error:
 Accuracy: 70.8%, Avg loss: 0.787981


Epoch 11
-------------------------------
loss: 0.794974  [    0/60000]
loss: 0.862666  [ 6400/60000]
loss: 0.649913  [12800/60000]
loss: 0.841434  [19200/60000]
loss: 0.739528  [25600/60000]
loss: 0.737574  [32000/60000]
loss: 0.827203  [38400/60000]
loss: 0.809241  [44800/60000]
loss: 0.801141  [51200/60000]
loss: 0.766895  [57600/60000]
Test Error:
 Accuracy: 72.0%, Avg loss: 0.761764
```

```
Epoch 12
-------------------------------
loss: 0.756405  [    0/60000]
loss: 0.833272  [ 6400/60000]
loss: 0.618174  [12800/60000]
loss: 0.816291  [19200/60000]
loss: 0.718671  [25600/60000]
loss: 0.711576  [32000/60000]
loss: 0.802181  [38400/60000]
loss: 0.793036  [44800/60000]
loss: 0.777932  [51200/60000]
loss: 0.745514  [57600/60000]
Test Error:
 Accuracy: 73.3%, Avg loss: 0.739154

Epoch 13
-------------------------------
loss: 0.722454  [    0/60000]
loss: 0.807023  [ 6400/60000]
loss: 0.590794  [12800/60000]
loss: 0.795077  [19200/60000]
loss: 0.700669  [25600/60000]
loss: 0.689929  [32000/60000]
loss: 0.779207  [38400/60000]
loss: 0.778795  [44800/60000]
loss: 0.757979  [51200/60000]
loss: 0.726265  [57600/60000]
Test Error:
 Accuracy: 74.2%, Avg loss: 0.719100

Epoch 14
-------------------------------
loss: 0.692237  [    0/60000]
loss: 0.783131  [ 6400/60000]
loss: 0.566655  [12800/60000]
loss: 0.776837  [19200/60000]
loss: 0.684880  [25600/60000]
loss: 0.671718  [32000/60000]
loss: 0.757862  [38400/60000]
loss: 0.765968  [44800/60000]
loss: 0.740439  [51200/60000]
loss: 0.708884  [57600/60000]
Test Error:
 Accuracy: 74.9%, Avg loss: 0.700972

Epoch 15
-------------------------------
```

```
loss: 0.664933  [    0/60000]
loss: 0.761112  [ 6400/60000]
loss: 0.545297  [12800/60000]
loss: 0.760738  [19200/60000]
loss: 0.670889  [25600/60000]
loss: 0.656096  [32000/60000]
loss: 0.737731  [38400/60000]
loss: 0.754223  [44800/60000]
loss: 0.724762  [51200/60000]
loss: 0.692910  [57600/60000]
Test Error:
 Accuracy: 75.8%, Avg loss: 0.684374


Epoch 16
-------------------------------
loss: 0.640301  [    0/60000]
loss: 0.740805  [ 6400/60000]
loss: 0.526202  [12800/60000]
loss: 0.746212  [19200/60000]
loss: 0.658458  [25600/60000]
loss: 0.642361  [32000/60000]
loss: 0.718656  [38400/60000]
loss: 0.743339  [44800/60000]
loss: 0.710708  [51200/60000]
loss: 0.678007  [57600/60000]
Test Error:
 Accuracy: 76.6%, Avg loss: 0.669057


Epoch 17
-------------------------------
loss: 0.617968  [    0/60000]
loss: 0.722022  [ 6400/60000]
loss: 0.509044  [12800/60000]
loss: 0.733012  [19200/60000]
loss: 0.647218  [25600/60000]
loss: 0.630369  [32000/60000]
loss: 0.700783  [38400/60000]
loss: 0.733483  [44800/60000]
loss: 0.698288  [51200/60000]
loss: 0.663941  [57600/60000]
Test Error:
 Accuracy: 77.2%, Avg loss: 0.654930


Epoch 18
-------------------------------
loss: 0.597719  [    0/60000]
loss: 0.704612  [ 6400/60000]
loss: 0.493512  [12800/60000]
```

```
loss: 0.720886  [19200/60000]
loss: 0.637213  [25600/60000]
loss: 0.619732  [32000/60000]
loss: 0.684095  [38400/60000]
loss: 0.724577  [44800/60000]
loss: 0.687314  [51200/60000]
loss: 0.650693  [57600/60000]
Test Error:
 Accuracy: 77.8%, Avg loss: 0.641852


Epoch 19
-------------------------------
loss: 0.579285  [    0/60000]
loss: 0.688543  [ 6400/60000]
loss: 0.479381  [12800/60000]
loss: 0.709693  [19200/60000]
loss: 0.628206  [25600/60000]
loss: 0.610176  [32000/60000]
loss: 0.668577  [38400/60000]
loss: 0.716611  [44800/60000]
loss: 0.677698  [51200/60000]
loss: 0.638301  [57600/60000]
Test Error:
 Accuracy: 78.4%, Avg loss: 0.629749


Epoch 20
-------------------------------
loss: 0.562499  [    0/60000]
loss: 0.673659  [ 6400/60000]
loss: 0.466497  [12800/60000]
loss: 0.699337  [19200/60000]
loss: 0.619986  [25600/60000]
loss: 0.601518  [32000/60000]
loss: 0.654112  [38400/60000]
loss: 0.709602  [44800/60000]
loss: 0.669386  [51200/60000]
loss: 0.626625  [57600/60000]
Test Error:
 Accuracy: 78.8%, Avg loss: 0.618569


Epoch 21
-------------------------------
loss: 0.547125  [    0/60000]
loss: 0.659857  [ 6400/60000]
loss: 0.454773  [12800/60000]
loss: 0.689606  [19200/60000]
loss: 0.612464  [25600/60000]
loss: 0.593709  [32000/60000]
```

```
loss: 0.640730  [38400/60000]
loss: 0.703518  [44800/60000]
loss: 0.662297  [51200/60000]
loss: 0.615599  [57600/60000]
Test Error:
 Accuracy: 79.1%, Avg loss: 0.608258


Epoch 22
-------------------------------
loss: 0.533147  [    0/60000]
loss: 0.647070  [ 6400/60000]
loss: 0.444088  [12800/60000]
loss: 0.680515  [19200/60000]
loss: 0.605524  [25600/60000]
loss: 0.586550  [32000/60000]
loss: 0.628334  [38400/60000]
loss: 0.698401  [44800/60000]
loss: 0.656289  [51200/60000]
loss: 0.605133  [57600/60000]
Test Error:
 Accuracy: 79.3%, Avg loss: 0.598758


Epoch 23
-------------------------------
loss: 0.520366  [    0/60000]
loss: 0.635235  [ 6400/60000]
loss: 0.434305  [12800/60000]
loss: 0.671958  [19200/60000]
loss: 0.599038  [25600/60000]
loss: 0.579916  [32000/60000]
loss: 0.616958  [38400/60000]
loss: 0.694093  [44800/60000]
loss: 0.651202  [51200/60000]
loss: 0.595192  [57600/60000]
Test Error:
 Accuracy: 79.7%, Avg loss: 0.590007


Epoch 24
-------------------------------
loss: 0.508569  [    0/60000]
loss: 0.624258  [ 6400/60000]
loss: 0.425347  [12800/60000]
loss: 0.663801  [19200/60000]
loss: 0.592908  [25600/60000]
loss: 0.573849  [32000/60000]
loss: 0.606550  [38400/60000]
loss: 0.690574  [44800/60000]
loss: 0.646927  [51200/60000]
```

```
loss: 0.585800  [57600/60000]
Test Error:
 Accuracy: 79.9%, Avg loss: 0.581953


Epoch 25
-------------------------------
loss: 0.497583  [    0/60000]
loss: 0.614092  [ 6400/60000]
loss: 0.417150  [12800/60000]
loss: 0.656061  [19200/60000]
loss: 0.586992  [25600/60000]
loss: 0.568170  [32000/60000]
loss: 0.596982  [38400/60000]
loss: 0.687713  [44800/60000]
loss: 0.643338  [51200/60000]
loss: 0.576842  [57600/60000]
Test Error:
 Accuracy: 80.3%, Avg loss: 0.574541


Epoch 26
-------------------------------
loss: 0.487424  [    0/60000]
loss: 0.604684  [ 6400/60000]
loss: 0.409593  [12800/60000]
loss: 0.648730  [19200/60000]
loss: 0.581309  [25600/60000]
loss: 0.562854  [32000/60000]
loss: 0.588233  [38400/60000]
loss: 0.685531  [44800/60000]
loss: 0.640369  [51200/60000]
loss: 0.568245  [57600/60000]
Test Error:
 Accuracy: 80.6%, Avg loss: 0.567716


Epoch 27
-------------------------------
loss: 0.477984  [    0/60000]
loss: 0.595962  [ 6400/60000]
loss: 0.402559  [12800/60000]
loss: 0.641799  [19200/60000]
loss: 0.575764  [25600/60000]
loss: 0.557817  [32000/60000]
loss: 0.580209  [38400/60000]
loss: 0.683882  [44800/60000]
loss: 0.637708  [51200/60000]
loss: 0.559995  [57600/60000]
Test Error:
 Accuracy: 80.8%, Avg loss: 0.561416
```

```
Epoch 28
-------------------------------
loss: 0.469170  [    0/60000]
loss: 0.587876  [ 6400/60000]
loss: 0.396000  [12800/60000]
loss: 0.635168  [19200/60000]
loss: 0.570320  [25600/60000]
loss: 0.553035  [32000/60000]
loss: 0.572849  [38400/60000]
loss: 0.682612  [44800/60000]
loss: 0.635281  [51200/60000]
loss: 0.552107  [57600/60000]
Test Error:
 Accuracy: 80.9%, Avg loss: 0.555589


Epoch 29
-------------------------------
loss: 0.460853  [    0/60000]
loss: 0.580326  [ 6400/60000]
loss: 0.389951  [12800/60000]
loss: 0.628756  [19200/60000]
loss: 0.564975  [25600/60000]
loss: 0.548358  [32000/60000]
loss: 0.566060  [38400/60000]
loss: 0.681749  [44800/60000]
loss: 0.633066  [51200/60000]
loss: 0.544506  [57600/60000]
Test Error:
 Accuracy: 81.0%, Avg loss: 0.550182


Epoch 30
-------------------------------
loss: 0.453028  [    0/60000]
loss: 0.573295  [ 6400/60000]
loss: 0.384339  [12800/60000]
loss: 0.622593  [19200/60000]
loss: 0.559784  [25600/60000]
loss: 0.543809  [32000/60000]
loss: 0.559774  [38400/60000]
loss: 0.681126  [44800/60000]
loss: 0.630988  [51200/60000]
loss: 0.537200  [57600/60000]
Test Error:
 Accuracy: 81.1%, Avg loss: 0.545153


Epoch 31
-------------------------------
```

```
loss: 0.445685  [    0/60000]
loss: 0.566751  [ 6400/60000]
loss: 0.379096  [12800/60000]
loss: 0.616717  [19200/60000]
loss: 0.554693  [25600/60000]
loss: 0.539456  [32000/60000]
loss: 0.553990  [38400/60000]
loss: 0.680672  [44800/60000]
loss: 0.629022  [51200/60000]
loss: 0.530216  [57600/60000]
Test Error:
 Accuracy: 81.2%, Avg loss: 0.540475


Epoch 32
-------------------------------
loss: 0.438745  [    0/60000]
loss: 0.560635  [ 6400/60000]
loss: 0.374175  [12800/60000]
loss: 0.611052  [19200/60000]
loss: 0.549627  [25600/60000]
loss: 0.535169  [32000/60000]
loss: 0.548627  [38400/60000]
loss: 0.680392  [44800/60000]
loss: 0.627125  [51200/60000]
loss: 0.523505  [57600/60000]
Test Error:
 Accuracy: 81.3%, Avg loss: 0.536104


Epoch 33
-------------------------------
loss: 0.432170  [    0/60000]
loss: 0.554922  [ 6400/60000]
loss: 0.369555  [12800/60000]
loss: 0.605577  [19200/60000]
loss: 0.544604  [25600/60000]
loss: 0.530948  [32000/60000]
loss: 0.543611  [38400/60000]
loss: 0.680160  [44800/60000]
loss: 0.625264  [51200/60000]
loss: 0.517067  [57600/60000]
Test Error:
 Accuracy: 81.4%, Avg loss: 0.532016


Epoch 34
-------------------------------
loss: 0.425933  [    0/60000]
loss: 0.549605  [ 6400/60000]
loss: 0.365209  [12800/60000]
```

```
loss: 0.600231  [19200/60000]
loss: 0.539660  [25600/60000]
loss: 0.526772  [32000/60000]
loss: 0.538898  [38400/60000]
loss: 0.679980  [44800/60000]
loss: 0.623436  [51200/60000]
loss: 0.510915  [57600/60000]
Test Error:
 Accuracy: 81.5%, Avg loss: 0.528183


Epoch 35
-------------------------------
loss: 0.419950  [    0/60000]
loss: 0.544604  [ 6400/60000]
loss: 0.361110  [12800/60000]
loss: 0.595054  [19200/60000]
loss: 0.534813  [25600/60000]
loss: 0.522702  [32000/60000]
loss: 0.534482  [38400/60000]
loss: 0.679793  [44800/60000]
loss: 0.621551  [51200/60000]
loss: 0.505032  [57600/60000]
Test Error:
 Accuracy: 81.7%, Avg loss: 0.524572


Epoch 36
-------------------------------
loss: 0.414200  [    0/60000]
loss: 0.539942  [ 6400/60000]
loss: 0.357218  [12800/60000]
loss: 0.590063  [19200/60000]
loss: 0.530064  [25600/60000]
loss: 0.518742  [32000/60000]
loss: 0.530371  [38400/60000]
loss: 0.679609  [44800/60000]
loss: 0.619621  [51200/60000]
loss: 0.499419  [57600/60000]
Test Error:
 Accuracy: 81.8%, Avg loss: 0.521174


Epoch 37
-------------------------------
loss: 0.408698  [    0/60000]
loss: 0.535579  [ 6400/60000]
loss: 0.353549  [12800/60000]
loss: 0.585233  [19200/60000]
loss: 0.525392  [25600/60000]
loss: 0.514843  [32000/60000]
```

```
loss: 0.526397  [38400/60000]
loss: 0.679338  [44800/60000]
loss: 0.617659  [51200/60000]
loss: 0.494079  [57600/60000]
Test Error:
 Accuracy: 82.0%, Avg loss: 0.517968


Epoch 38
-------------------------------
loss: 0.403414  [    0/60000]
loss: 0.531500  [ 6400/60000]
loss: 0.350066  [12800/60000]
loss: 0.580510  [19200/60000]
loss: 0.520799  [25600/60000]
loss: 0.511017  [32000/60000]
loss: 0.522680  [38400/60000]
loss: 0.678944  [44800/60000]
loss: 0.615709  [51200/60000]
loss: 0.488982  [57600/60000]
Test Error:
 Accuracy: 82.0%, Avg loss: 0.514935


Epoch 39
-------------------------------
loss: 0.398317  [    0/60000]
loss: 0.527636  [ 6400/60000]
loss: 0.346736  [12800/60000]
loss: 0.575945  [19200/60000]
loss: 0.516323  [25600/60000]
loss: 0.507284  [32000/60000]
loss: 0.519138  [38400/60000]
loss: 0.678487  [44800/60000]
loss: 0.613748  [51200/60000]
loss: 0.484179  [57600/60000]
Test Error:
 Accuracy: 82.1%, Avg loss: 0.512062


Epoch 40
-------------------------------
loss: 0.393425  [    0/60000]
loss: 0.523959  [ 6400/60000]
loss: 0.343545  [12800/60000]
loss: 0.571548  [19200/60000]
loss: 0.511949  [25600/60000]
loss: 0.503585  [32000/60000]
loss: 0.515781  [38400/60000]
loss: 0.677981  [44800/60000]
loss: 0.611740  [51200/60000]
```

```
loss: 0.479623  [57600/60000]
Test Error:
 Accuracy: 82.2%, Avg loss: 0.509328


Epoch 41
-------------------------------
loss: 0.388678  [    0/60000]
loss: 0.520578  [ 6400/60000]
loss: 0.340496  [12800/60000]
loss: 0.567305  [19200/60000]
loss: 0.507701  [25600/60000]
loss: 0.499994  [32000/60000]
loss: 0.512593  [38400/60000]
loss: 0.677378  [44800/60000]
loss: 0.609739  [51200/60000]
loss: 0.475284  [57600/60000]
Test Error:
 Accuracy: 82.2%, Avg loss: 0.506727


Epoch 42
-------------------------------
loss: 0.384083  [    0/60000]
loss: 0.517398  [ 6400/60000]
loss: 0.337565  [12800/60000]
loss: 0.563235  [19200/60000]
loss: 0.503554  [25600/60000]
loss: 0.496431  [32000/60000]
loss: 0.509514  [38400/60000]
loss: 0.676646  [44800/60000]
loss: 0.607716  [51200/60000]
loss: 0.471181  [57600/60000]
Test Error:
 Accuracy: 82.2%, Avg loss: 0.504249


Epoch 43
-------------------------------
loss: 0.379638  [    0/60000]
loss: 0.514382  [ 6400/60000]
loss: 0.334762  [12800/60000]
loss: 0.559277  [19200/60000]
loss: 0.499467  [25600/60000]
loss: 0.493009  [32000/60000]
loss: 0.506529  [38400/60000]
loss: 0.675821  [44800/60000]
loss: 0.605699  [51200/60000]
loss: 0.467293  [57600/60000]
Test Error:
 Accuracy: 82.3%, Avg loss: 0.501885
```

```
Epoch 44
-------------------------------
loss: 0.375344  [    0/60000]
loss: 0.511530  [ 6400/60000]
loss: 0.332111  [12800/60000]
loss: 0.555430  [19200/60000]
loss: 0.495519  [25600/60000]
loss: 0.489661  [32000/60000]
loss: 0.503663  [38400/60000]
loss: 0.674823  [44800/60000]
loss: 0.603631  [51200/60000]
loss: 0.463592  [57600/60000]
Test Error:
 Accuracy: 82.4%, Avg loss: 0.499625

Epoch 45
-------------------------------
loss: 0.371173  [    0/60000]
loss: 0.508784  [ 6400/60000]
loss: 0.329526  [12800/60000]
loss: 0.551710  [19200/60000]
loss: 0.491672  [25600/60000]
loss: 0.486396  [32000/60000]
loss: 0.500909  [38400/60000]
loss: 0.673726  [44800/60000]
loss: 0.601543  [51200/60000]
loss: 0.460097  [57600/60000]
Test Error:
 Accuracy: 82.4%, Avg loss: 0.497460

Epoch 46
-------------------------------
loss: 0.367093  [    0/60000]
loss: 0.506171  [ 6400/60000]
loss: 0.327002  [12800/60000]
loss: 0.548180  [19200/60000]
loss: 0.487933  [25600/60000]
loss: 0.483258  [32000/60000]
loss: 0.498213  [38400/60000]
loss: 0.672559  [44800/60000]
loss: 0.599544  [51200/60000]
loss: 0.456784  [57600/60000]
Test Error:
 Accuracy: 82.6%, Avg loss: 0.495381

Epoch 47
-------------------------------
```

```
loss: 0.363121  [    0/60000]
loss: 0.503684  [ 6400/60000]
loss: 0.324557  [12800/60000]
loss: 0.544768  [19200/60000]
loss: 0.484304  [25600/60000]
loss: 0.480249  [32000/60000]
loss: 0.495595  [38400/60000]
loss: 0.671296  [44800/60000]
loss: 0.597470  [51200/60000]
loss: 0.453647  [57600/60000]
Test Error:
 Accuracy: 82.7%, Avg loss: 0.493383


Epoch 48
-------------------------------
loss: 0.359290  [    0/60000]
loss: 0.501274  [ 6400/60000]
loss: 0.322188  [12800/60000]
loss: 0.541475  [19200/60000]
loss: 0.480774  [25600/60000]
loss: 0.477294  [32000/60000]
loss: 0.493041  [38400/60000]
loss: 0.669907  [44800/60000]
loss: 0.595427  [51200/60000]
loss: 0.450703  [57600/60000]
Test Error:
 Accuracy: 82.7%, Avg loss: 0.491458


Epoch 49
-------------------------------
loss: 0.355578  [    0/60000]
loss: 0.498993  [ 6400/60000]
loss: 0.319940  [12800/60000]
loss: 0.538331  [19200/60000]
loss: 0.477325  [25600/60000]
loss: 0.474424  [32000/60000]
loss: 0.490540  [38400/60000]
loss: 0.668574  [44800/60000]
loss: 0.593414  [51200/60000]
loss: 0.447911  [57600/60000]
Test Error:
 Accuracy: 82.7%, Avg loss: 0.489601


Epoch 50
-------------------------------
loss: 0.351914  [    0/60000]
loss: 0.496841  [ 6400/60000]
loss: 0.317769  [12800/60000]
```

```
loss: 0.535295   [19200/60000]
loss: 0.473950   [25600/60000]
loss: 0.471680   [32000/60000]
loss: 0.488097   [38400/60000]
loss: 0.667117   [44800/60000]
loss: 0.591421   [51200/60000]
loss: 0.445266   [57600/60000]
Test Error:
 Accuracy: 82.8%, Avg loss: 0.487809
```
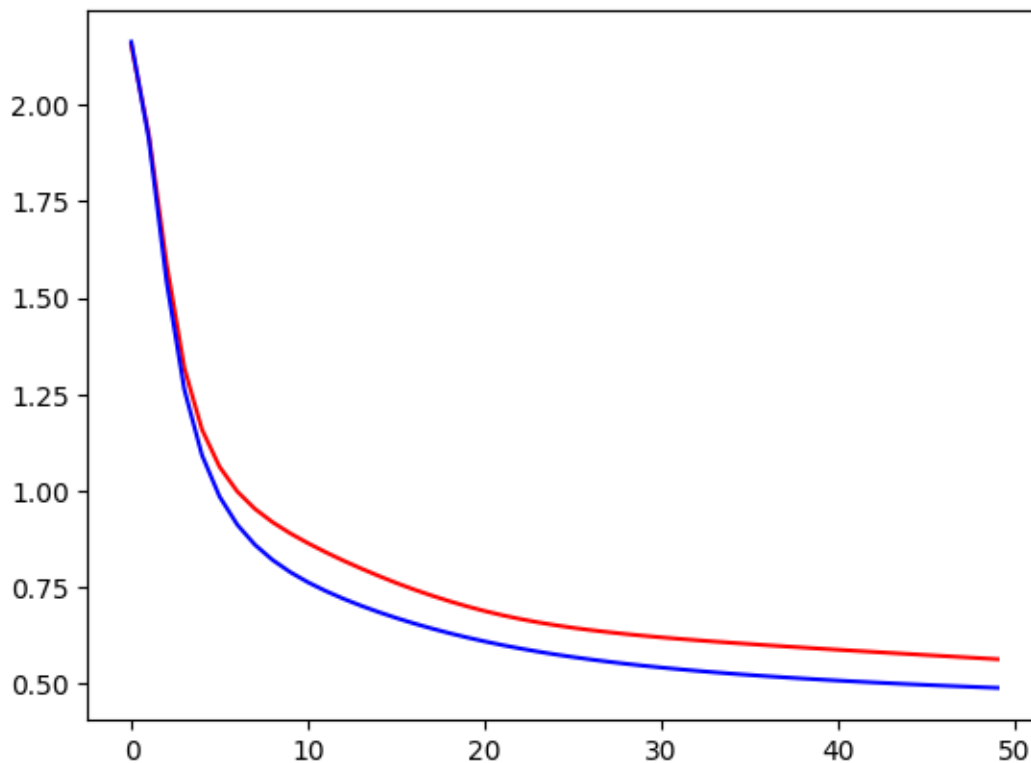


```
Base model accuracy is: 0.8275
```

After incresasing the number of epochs used to train the model, it can be seen that the accuracy for the deeper model has continued to increase and the loss was able to continue decreasing as more epochs were trained. From the above graphs comparing the loss as the deeper model and the base model are trained, it seems that the loss of the deeper model can continue to decrease, while the loss of the base model has flattened out. This may indicate that the deeper model can be trained further to continue improving the accuracy, while the base model is not able to be trained further.

1.1.6 Question 1.6
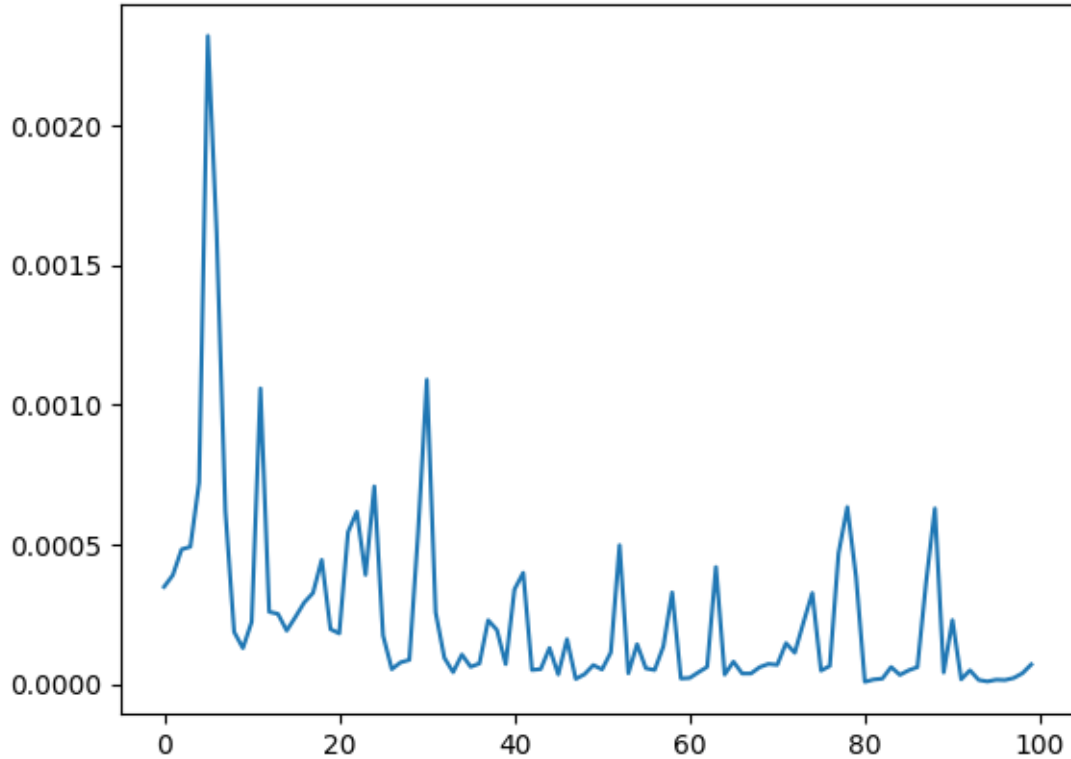
As shown above, the mean of the gradients of loss has been calculated for the first 100 steps in the training. This has then been plotted on a graph. It can be seen in the graph that over the training for the first 100 steps, the gradients curve decreases.

1.1.7 Question 1.7

```
Finished Training
Accuracy of the network on the test images: 87.29 %

Accuracy for class: T-shirt/top is 81.7 %
Accuracy for class: Trouser is 98.2 %
Accuracy for class: Pullover is 60.9 %
Accuracy for class: Dress is 88.5 %
Accuracy for class: Coat  is 84.5 %
Accuracy for class: Sandal is 96.3 %
Accuracy for class: Shirt is 73.0 %
Accuracy for class: Sneaker is 96.6 %
Accuracy for class: Bag   is 97.5 %
Accuracy for class: Ankle boot is 95.7 %

Structures Accuracy   Parameters
Base MLP   71.5      % 669706
```

```
Wider MLP  73.22999999999999% 1863690
Deeper MLP 49.7      % 1195018
CNN        87.29     % 46982
```

From the results above, the convolutional neural network was able to achieve results with better accuracy to the MLP that was trained with the same learning rate. A result of 87% accuracy as able to be achieved with the CNN compared to only 72% for the base MLP, 73% for the wider MLP and 50% for the deeper MLP when training for just 10 epochs.

Because of this, it is clear that the CNN is able to converge in less epochs than the MLP model, achieving a higher accuracy with less training time. When comparing the number of parameters across models, it is also clear that this CNN has far less parameters, than any of the MLP models tested. With less than one tenth the number of parameters in the base MLP, and an even smaller percentage of the number of parameters in the wider or deeper MLPs tested.

We can also see that the accuracies found for classifying the images into classes was very high for the CNN model, often approaching 100%, especially for the trousers, sneakers, sandals, bag, and ankle boot. But the accuracy for the pullover was quite low in comparison, only 61%. This may indicate that more training is needed for that particular class of image or that there are less features available in pullover images to effectively match on.

Given this, it is likely that a CNN is a better method to classify these images compared to using a MLP to do the same task.