CS 547 Deep Learning: HW 4 Report

Name: Jiashuo Tong, Date Submitted: Oct. 15, 2019

This report is a documentation for CS 547 Homework 4 which includes a total of 4 tasks. For each task the goal is first described, followed by descriptions of the network architecture, training parameters and the training results.

Task 1: Implement and train a ResNet on CIFAR100 dataset. It should achieve 60% test accuracy within a training time of 2 hours.

Neural Network Structure:

Type of Layer		Shape of Layer	Output Shape (batch size,#channel,height,width)
Convolution		in_channels=3, out_channels=32, kernel_size=3, stride=1,padding=1	(batch_size, 32,32,32)
Batch Normalization		num_features=32	(batch_size, 32,32,32)
Dropout		probability=0.5	(batch_size, 32,32,32)
Block × 2	Convolution	in_channels=32, out_channels=32, kernel_size=3, stride=1, padding=1	(batch_size, 32,32,32)
	Batch Norm and ReLU	num_features=32	(batch_size, 32,32,32)
	Convolution	in_channels=32, out_channels=32, kernel_size=3, stride=1, padding=1	(batch_size, 32,32,32)
	Batch Norm	num_features=32	(batch_size, 32,32,32)
Block × 4	Convolution	in_channels=32, out_channels=64, kernel_size=3, stride=2, padding=1	(batch_size, 64,16,16)
	Batch Norm and ReLU	num_features=64	(batch_size, 64,16,16)
	Convolution	in_channels=64, out_channels=64, kernel_size=3, stride=1, padding=1	(batch_size, 64,16,16)
	Batch Norm	num_features=64	(batch_size, 64,16,16)
Block × 4	Convolution	in_channels=64, out_channels=128, kernel_size=3, stride=2, padding=1	(batch_size, 128,8,8)
	Batch Norm and ReLU	num_features=128	(batch_size, 128,8,8)
	Convolution	in_channels=128, out_channels=128, kernel_size=3, stride=1, padding=1	(batch_size, 128,8,8)
	Batch Norm	num_features=128	(batch_size, 128,8,8)
Block × 2	Convolution	in_channels=128, out_channels=256, kernel_size=3, stride=2, padding=1	(batch_size, 256,4,4)
	Batch Norm and ReLU	num_features=256	(batch_size, 256,4,4)
	Convolution	in_channels=256, out_channels=256, kernel_size=3, stride=1, padding=1	(batch_size, 256,4,4)
	Batch Norm	num_features=256	(batch_size, 256,4,4)
Maximum Pooling		kernel_size=4, padding=4	(batch_size, 256,1,1)
Reshape		This layer reshape data from 4D to 2D	(batch_size, 256)
Fully Connected		in_features=256, out_features=100	(batch_size, 100)

Hyperparameters And Optimizer for Training:

batch_size = 100
num_epochs = 20
learning_rate = 1e-3
dropout probability = 0.5
optimizer = Adam

Results:

At the end of the 1st epoch, the model achieves an accuracy of 14.31% on the test set. The accuracy increases with number of epochs trained for the next several epochs and hits an accuracy of over 60% in the 15th epoch. Starting from the 15th epoch, the test accuracy fluctuates around 60% as the training accuracy climbs up, which fulfills the requirement of this task. For more details, please see the training log and the learning curves in the following 2 pages.

Training Log for Task 1

Epoch 1 / 20 is complete in 129.87166833877563 seconds with a training accuracy of 8.648
Accuracy of predictions on the test set is 14.31

==========

Epoch 2 / 20 is complete in 266.6671085357666 seconds with a training accuracy of 19.658
Accuracy of predictions on the test set is 22.45

==========

Epoch 3 / 20 is complete in 403.50407886505127 seconds with a training accuracy of 28.144
Accuracy of predictions on the test set is 32.07

==========

Epoch 4 / 20 is complete in 540.3924939632416 seconds with a training accuracy of 35.942
Accuracy of predictions on the test set is 38.13

==========

Epoch 5 / 20 is complete in 677.3279621601105 seconds with a training accuracy of 41.716
Accuracy of predictions on the test set is 43.41

==========

Epoch 6 / 20 is complete in 814.2343378067017 seconds with a training accuracy of 46.712
Accuracy of predictions on the test set is 44.19

==========

Epoch 7 / 20 is complete in 951.1294023990631 seconds with a training accuracy of 50.546
Accuracy of predictions on the test set is 47.49

==========

Epoch 8 / 20 is complete in 1087.9535694122314 seconds with a training accuracy of 53.72
Accuracy of predictions on the test set is 51.82

==========

Epoch 9 / 20 is complete in 1224.8243753910065 seconds with a training accuracy of 56.508
Accuracy of predictions on the test set is 52.98

==========

Epoch 10 / 20 is complete in 1361.6341841220856 seconds with a training accuracy of 59.26
Accuracy of predictions on the test set is 54.19

==========

Epoch 11 / 20 is complete in 1498.4422030448914 seconds with a training accuracy of 61.036
Accuracy of predictions on the test set is 55.94

==========

Epoch 12 / 20 is complete in 1635.2787289619446 seconds with a training accuracy of 64.98
Accuracy of predictions on the test set is 57.04

==========

Epoch 13 / 20 is complete in 1772.1652250289917 seconds with a training accuracy of 66.534
Accuracy of predictions on the test set is 56.76

==========

Epoch 14 / 20 is complete in 1909.0270631313324 seconds with a training accuracy of 67.952
Accuracy of predictions on the test set is 58.97

==========

Epoch 15 / 20 is complete in 2045.8729979991913 seconds with a training accuracy of 69.788

Accuracy of predictions on the test set is 60.47

==========

Epoch 16 / 20 is complete in 2182.7688806056976 seconds with a training accuracy of 70.808

Accuracy of predictions on the test set is 59.94

==========

Epoch 17 / 20 is complete in 2319.6531207561493 seconds with a training accuracy of 72.88

Accuracy of predictions on the test set is 60.12

==========

Epoch 18 / 20 is complete in 2456.4883399009705 seconds with a training accuracy of 73.966

Accuracy of predictions on the test set is 59.69

==========

Epoch 19 / 20 is complete in 2593.3545763492584 seconds with a training accuracy of 75.35

Accuracy of predictions on the test set is 60.84

==========

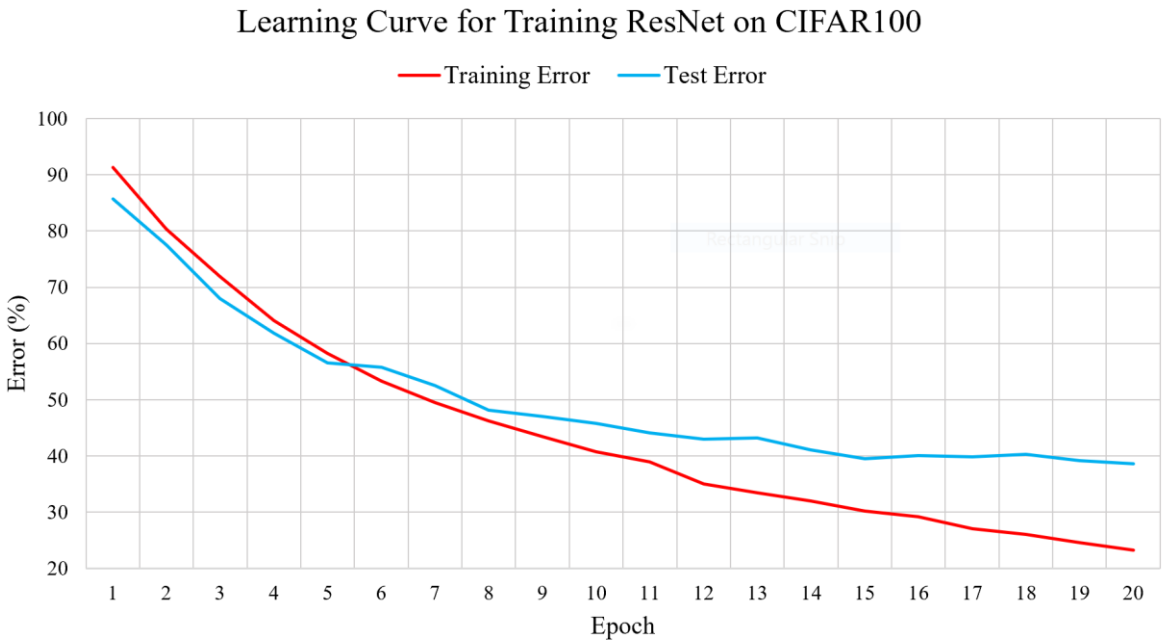
Epoch 20 / 20 is complete in 2796.740917444229 seconds with a training accuracy of 76.684

Accuracy of predictions on the test set is 61.36

==========

Model is saved. Total time consumed is 2805.260555744171 seconds

Learning Curve for Task 1:



Task 2: Implement and train a ResNet on Tiny ImageNet dataset. It should achieve 50% test accuracy within a training time of 2 hours.

Neural Network Structure:

Type of Layer		Shape of Layer	Output Shape (batch size,#channel,height,width)
Convolution		in_channels=3, out_channels=32, kernel_size=3, stride=1,padding=1	(batch_size, 3,64,64)
Batch Normalization		num_features=32	(batch_size, 32,64,64)
Dropout		probability=0.5	(batch_size, 32,64,64)
Block × 2	Convolution	in_channels=32, out_channels=32, kernel_size=3, stride=1, padding=1	(batch_size, 32,64,64)
	Batch Norm and ReLU	num_features=32	(batch_size, 32,64,64)
	Convolution	in_channels=32, out_channels=32, kernel_size=3, stride=1, padding=1	(batch_size, 32,64,64)
	Batch Norm	num_features=32	(batch_size, 32,64,64)
Block × 4	Convolution	in_channels=32, out_channels=64, kernel_size=3, stride=2, padding=1	(batch_size, 64,32,32)
	Batch Norm and ReLU	num_features=64	(batch_size, 64,32,32)
	Convolution	in_channels=64, out_channels=64, kernel_size=3, stride=1, padding=1	(batch_size, 64,32,32)
	Batch Norm	num_features=64	(batch_size, 64,32,32)
Block × 4	Convolution	in_channels=64, out_channels=128, kernel_size=3, stride=2, padding=1	(batch_size, 128,16,16)
	Batch Norm and ReLU	num_features=128	(batch_size, 128,16,16)
	Convolution	in_channels=128, out_channels=128, kernel_size=3, stride=1, padding=1	(batch_size, 128,16,16)
	Batch Norm	num_features=128	(batch_size, 128,16,16)
Block × 2	Convolution	in_channels=128, out_channels=256, kernel_size=3, stride=2, padding=1	(batch_size, 256,8,8)
	Batch Norm and ReLU	num_features=256	(batch_size, 256,8,8)
	Convolution	in_channels=256, out_channels=256, kernel_size=3, stride=1, padding=1	(batch_size, 256,8,8)
	Batch Norm	num_features=256	(batch_size, 256,8,8)
Maximum Pooling		kernel_size=4, padding=4	(batch_size, 256,2,2)
Reshape		This layer reshape data from 4D to 2D	(batch_size, 1024)
Fully Connected		in_features=1024, out_features=200	(batch_size, 200)

Hyperparameters And Optimizer for Training:

batch_size = 150
num_epochs = 20
learning_rate = 1e-3
dropout probability = 0.4
optimizer = Adam

Results:

At the end of the 1st epoch, the model achieves an accuracy of 11.31% on the test set. The accuracy increases with number of epochs trained for the next several epochs and hits an accuracy of over 50% in the 12th epoch within 9756 seconds (approximately 3.05 hours). After reaching 50%, the test accuracy stays around 51% as the training accuracy climbs up, which fulfills the requirement of this task. For more details, please see the training log and the learning curves in the following 2 pages.

Training Log for Task 2:

Epoch 1 / 20 is complete in 792.2873017787933 seconds with a training accuracy of 5.7141429285357335
Accuracy of predictions on the test set is 11.064676616915424

==========

Epoch 2 / 20 is complete in 1611.2190911769867 seconds with a training accuracy of 18.3208395802099
Accuracy of predictions on the test set is 23.70149253731343

==========

Epoch 3 / 20 is complete in 2425.3032574653625 seconds with a training accuracy of 26.997501249375315
Accuracy of predictions on the test set is 29.99004975124378

==========

Epoch 4 / 20 is complete in 3235.2870235443115 seconds with a training accuracy of 33.38630684657671
Accuracy of predictions on the test set is 34.437810945273625

==========

Epoch 5 / 20 is complete in 4043.911552667618 seconds with a training accuracy of 38.32083958020989
Accuracy of predictions on the test set is 36.557213930348254

==========

Epoch 6 / 20 is complete in 4852.095100402832 seconds with a training accuracy of 42.31284357821089
Accuracy of predictions on the test set is 40.39800995024876

==========

Epoch 7 / 20 is complete in 5659.654972314835 seconds with a training accuracy of 46.93653173413293
Accuracy of predictions on the test set is 42.85572139303482

==========

Epoch 8 / 20 is complete in 6468.467147111893 seconds with a training accuracy of 49.69915042478761
Accuracy of predictions on the test set is 43.48258706467662

==========

Epoch 9 / 20 is complete in 7276.839107990265 seconds with a training accuracy of 52.035982008995504
Accuracy of predictions on the test set is 44.67661691542288

==========

Epoch 10 / 20 is complete in 8088.18318104744 seconds with a training accuracy of 54.35982008995502
Accuracy of predictions on the test set is 46.42786069651742

==========

Epoch 11 / 20 is complete in 8897.84295463562 seconds with a training accuracy of 56.53473263368316
Accuracy of predictions on the test set is 45.70149253731344

==========

Epoch 12 / 20 is complete in 9755.76317191124 seconds with a training accuracy of 64.26086956521739
Accuracy of predictions on the test set is 50.56716417910447

==========

Epoch 13 / 20 is complete in 10567.027132749557 seconds with a training accuracy of 65.59120439780109
Accuracy of predictions on the test set is 50.82587064676617

==========

Epoch 14 / 20 is complete in 11377.19905424118 seconds with a training accuracy of 66.28685657171414
Accuracy of predictions on the test set is 50.95522388059702

==========

Epoch 15 / 20 is complete in 12185.719141244888 seconds with a training accuracy of 67.07146426786606

Accuracy of predictions on the test set is 50.9452736318408

==========

Epoch 16 / 20 is complete in 12994.959188222885 seconds with a training accuracy of 67.68415792103949
Accuracy of predictions on the test set is 50.82587064676617

==========

Epoch 17 / 20 is complete in 13804.963065624237 seconds with a training accuracy of 68.13493253373314
Accuracy of predictions on the test set is 50.74626865671642

==========

Epoch 18 / 20 is complete in 14614.57498550415 seconds with a training accuracy of 68.69265367316342
Accuracy of predictions on the test set is 50.81592039800996

==========

Epoch 19 / 20 is complete in 15424.815161705017 seconds with a training accuracy of 69.06146926536732
Accuracy of predictions on the test set is 51.094527363184085

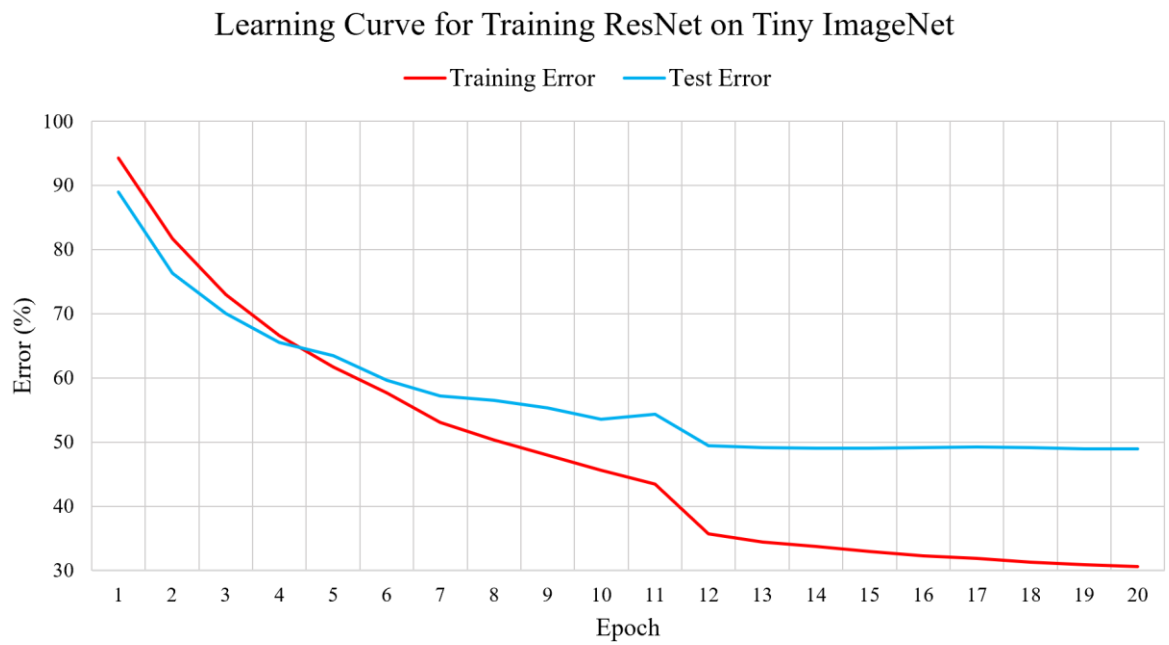
==========

Epoch 20 / 20 is complete in 16233.33943605423 seconds with a training accuracy of 69.42128935532234
Accuracy of predictions on the test set is 51.08457711442786

==========

Model is saved. Total time consumed is 16264.718199253082 seconds

Learning Curve for Task 2:



Task 3: Load and fine tune a pretrained ResNet (ResNet-18) on CIFAR100 dataset. It should achieve 70% test accuracy within a training time of 4 hours.

Data Transformation:
Since ResNet18 is assumes an image size of 224 by 224 and CIFAR100 are images with shape 32 by 32, we need to resize the CIFAR100 data to fit the input size of ResNet18. We also augment the data through randomly applying horizontal flip on the images. Finally, we normalize the input data by precalculated means and standard deviations. All the data transformation techniques are implemented using pytorch functions in torchvision.transforms.

Neural Network Structure:
We keep all of the ResNet-18 structure except that we change the parameter "out_features" for the fully connected layer from 1000 to 100, using the following command

```
in_features = model.fc.in_features
model.fc = nn.Linear(in_features, 100)
```

Hyperparameters And Optimizer for Training:
batch_size = 100
num_epochs = 15
learning_rate = 1e-3
optimizer = Adam

Results:
At the end of the 1st epoch, the model achieves an accuracy of 59.1% on the test set. The accuracy increases with number of epochs trained for the next several epochs and hits an accuracy of 70% in the 5th epoch. After exceeding 70%, both test accuracy climb a little bit and stay quite stable at around 73%, which fulfills the requirement of this task. For more details, please see the training log and the learning curves in following 2 pages.

Training Log for Task 3:

Epoch 1 / 15 is complete in 507.401208400726 seconds with a training accuracy of 50.204
Accuracy of predictions on the test set is 59.1

==========

Epoch 2 / 15 is complete in 1061.6860074996948 seconds with a training accuracy of 67.3
Accuracy of predictions on the test set is 63.45

==========

Epoch 3 / 15 is complete in 1615.4935669898987 seconds with a training accuracy of 74.168
Accuracy of predictions on the test set is 67.7

==========

Epoch 4 / 15 is complete in 2168.8452801704407 seconds with a training accuracy of 79.554
Accuracy of predictions on the test set is 69.18

==========

Epoch 5 / 15 is complete in 2721.9333691596985 seconds with a training accuracy of 83.384
Accuracy of predictions on the test set is 70.56

==========

Epoch 6 / 15 is complete in 3274.920712709427 seconds with a training accuracy of 87.29
Accuracy of predictions on the test set is 70.02

==========

Epoch 7 / 15 is complete in 3827.945460319519 seconds with a training accuracy of 89.092
Accuracy of predictions on the test set is 70.44

==========

Epoch 8 / 15 is complete in 4380.865374088287 seconds with a training accuracy of 91.124
Accuracy of predictions on the test set is 70.8

==========

Epoch 9 / 15 is complete in 4933.733560323715 seconds with a training accuracy of 92.226
Accuracy of predictions on the test set is 71.38

==========

Epoch 10 / 15 is complete in 5486.46541595459 seconds with a training accuracy of 93.354
Accuracy of predictions on the test set is 70.13

==========

Epoch 11 / 15 is complete in 6039.065474033356 seconds with a training accuracy of 94.208
Accuracy of predictions on the test set is 69.54

==========

Epoch 12 / 15 is complete in 6592.497533559799 seconds with a training accuracy of 96.704
Accuracy of predictions on the test set is 73.79

==========

Epoch 13 / 15 is complete in 7145.521236658096 seconds with a training accuracy of 97.49
Accuracy of predictions on the test set is 73.24

==========

Epoch 14 / 15 is complete in 7698.345458745956 seconds with a training accuracy of 97.198
Accuracy of predictions on the test set is 72.06

==========

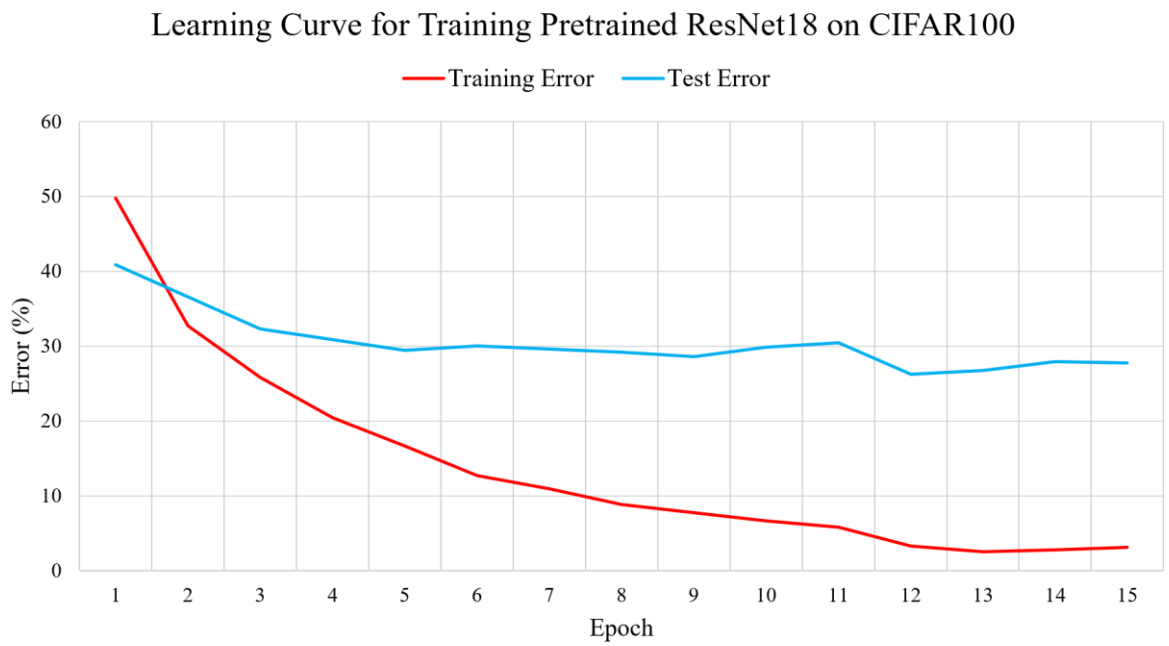
Epoch 15 / 15 is complete in 8251.293516159058 seconds with a training accuracy of 96.87

Accuracy of predictions on the test set is 72.23

==========

Model is saved. Total time consumed is 8300.546686649323 seconds

Learning Curve for Task 3:



Task 4: Train a ResNet on CIFAR100 dataset using Synchronous SGD method. It should achieve 60% test accuracy within a training time of 2 hours.

Neural Network Structure
The network structure is the same as the structure for Task 1.

Parallel Computing Detals:
For the purpose of practicing data parallelism, a total of two computing nodes are used for this task. The model is initialized in CPU and sent to each of the computing nodes. During the training, gradients are generated by both computing nodes, sent to the CPU, and taken average by the all_reduce function. The parameters for models in different computing nodes should keep the same throughout the training process, since the same average of gradients is returned to the computing nodes for each parameter update.

Hyperparameters For Training:
batch_size = 100
num_epochs = 30
learning_rate = 1e-3
dropout probability = 0.5
optimizer = Adam

Results:
At the end of the 1st epoch, models on both computing nodes achieve an accuracy of around 15% on the test set. The accuracies increase with number of epochs trained for the next several epochs. The accuracy for the first computing node reaches 60% in the 17th epoch. After exceeding 60%, the test accuracies climb a little bit and stays quite stable at around 63%, which fulfills the requirement of this task. For more details, please see the training log and the learning curves (for Node 1) in following 4 pages.

Training Log for Task 4:

Epoch 1 / 30 is complete in 182.28497505187988 seconds with a training accuracy of 8.78
Epoch 1 / 30 is complete in 182.26946544647217 seconds with a training accuracy of 8.956
Accuracy of predictions on the test set is 13.28
Accuracy of predictions on the test set is 13.81
==========
Epoch 2 / 30 is complete in 370.14060854911804 seconds with a training accuracy of 19.332
Epoch 2 / 30 is complete in 370.1250047683716 seconds with a training accuracy of 19.724
Accuracy of predictions on the test set is 22.53
Accuracy of predictions on the test set is 23.21
==========
Epoch 3 / 30 is complete in 557.9323196411133 seconds with a training accuracy of 28.55
Epoch 3 / 30 is complete in 557.9479334354401 seconds with a training accuracy of 28.372
Accuracy of predictions on the test set is 30.46
Accuracy of predictions on the test set is 26.19
==========
Epoch 4 / 30 is complete in 745.7491908073425 seconds with a training accuracy of 36.278
Epoch 4 / 30 is complete in 745.7648255825043 seconds with a training accuracy of 35.92
Accuracy of predictions on the test set is 37.71
Accuracy of predictions on the test set is 33.22
==========
Epoch 5 / 30 is complete in 933.5421853065491 seconds with a training accuracy of 42.176
Epoch 5 / 30 is complete in 933.5266463756561 seconds with a training accuracy of 42.196
Accuracy of predictions on the test set is 43.87
Accuracy of predictions on the test set is 42.48
==========
Epoch 6 / 30 is complete in 1121.3671624660492 seconds with a training accuracy of 47.264
Epoch 6 / 30 is complete in 1121.3515496253967 seconds with a training accuracy of 47.374
Accuracy of predictions on the test set is 46.85
Accuracy of predictions on the test set is 44.98
==========
Epoch 7 / 30 is complete in 1309.217069387436 seconds with a training accuracy of 52.006
Epoch 7 / 30 is complete in 1309.2014436721802 seconds with a training accuracy of 52.058
Accuracy of predictions on the test set is 48.05
Accuracy of predictions on the test set is 48.72
==========
Epoch 8 / 30 is complete in 1497.039943933487 seconds with a training accuracy of 55.192
Epoch 8 / 30 is complete in 1497.0243623256683 seconds with a training accuracy of 55.346
Accuracy of predictions on the test set is 49.38
Accuracy of predictions on the test set is 48.28
==========
Epoch 9 / 30 is complete in 1684.7652678489685 seconds with a training accuracy of 57.852
Epoch 9 / 30 is complete in 1684.7496654987335 seconds with a training accuracy of 57.998
Accuracy of predictions on the test set is 52.27

Accuracy of predictions on the test set is 51.7

==========

Epoch 10 / 30 is complete in 1872.621220111847 seconds with a training accuracy of 60.898

Epoch 10 / 30 is complete in 1872.6056716442108 seconds with a training accuracy of 60.594

Accuracy of predictions on the test set is 53.57

Accuracy of predictions on the test set is 51.67

==========

Epoch 11 / 30 is complete in 2060.4769852161407 seconds with a training accuracy of 62.922

Epoch 11 / 30 is complete in 2060.461402654648 seconds with a training accuracy of 63.228

Accuracy of predictions on the test set is 55.22

Accuracy of predictions on the test set is 53.1

==========

Epoch 12 / 30 is complete in 2248.360791683197 seconds with a training accuracy of 65.03

Epoch 12 / 30 is complete in 2248.34517621994 seconds with a training accuracy of 65.064

Accuracy of predictions on the test set is 55.64

Accuracy of predictions on the test set is 54.61

==========

Epoch 13 / 30 is complete in 2436.2143874168396 seconds with a training accuracy of 66.674

Epoch 13 / 30 is complete in 2436.1987862586975 seconds with a training accuracy of 66.964

Accuracy of predictions on the test set is 57.32

Accuracy of predictions on the test set is 56.5

==========

Epoch 14 / 30 is complete in 2624.1052153110504 seconds with a training accuracy of 68.852

Epoch 14 / 30 is complete in 2624.1209115982056 seconds with a training accuracy of 68.88

Accuracy of predictions on the test set is 58.31

Accuracy of predictions on the test set is 57.05

==========

Epoch 15 / 30 is complete in 2811.9650597572327 seconds with a training accuracy of 70.298

Epoch 15 / 30 is complete in 2811.980707168579 seconds with a training accuracy of 70.734

Accuracy of predictions on the test set is 57.73

Accuracy of predictions on the test set is 57.52

==========

Epoch 16 / 30 is complete in 2999.81981921196 seconds with a training accuracy of 72.392

Epoch 16 / 30 is complete in 2999.8354454040527 seconds with a training accuracy of 72.246

Accuracy of predictions on the test set is 57.52

Accuracy of predictions on the test set is 56.27

==========

Epoch 17 / 30 is complete in 3187.67307472229 seconds with a training accuracy of 74.016

Epoch 17 / 30 is complete in 3187.657535791397 seconds with a training accuracy of 73.826

Accuracy of predictions on the test set is 58.82

Accuracy of predictions on the test set is 59.17

==========

Epoch 18 / 30 is complete in 3375.5220081806183 seconds with a training accuracy of 75.56

Epoch 18 / 30 is complete in 3375.5064191818237 seconds with a training accuracy of 75.06

Accuracy of predictions on the test set is 60.26
Accuracy of predictions on the test set is 59.64
==========
Epoch 19 / 30 is complete in 3563.342886686325 seconds with a training accuracy of 76.754
Epoch 19 / 30 is complete in 3563.3273210525513 seconds with a training accuracy of 76.348
Accuracy of predictions on the test set is 59.75
Accuracy of predictions on the test set is 58.47
==========
Epoch 20 / 30 is complete in 3751.146908044815 seconds with a training accuracy of 77.778
Epoch 20 / 30 is complete in 3751.1625938415527 seconds with a training accuracy of 77.826
Accuracy of predictions on the test set is 60.15
Accuracy of predictions on the test set is 59.32
==========
Epoch 21 / 30 is complete in 3938.981353998184 seconds with a training accuracy of 79.188
Epoch 21 / 30 is complete in 3938.9656932353973 seconds with a training accuracy of 79.056
Accuracy of predictions on the test set is 59.67
Accuracy of predictions on the test set is 59.53
==========
Epoch 22 / 30 is complete in 4126.78848862648 seconds with a training accuracy of 85.172
Epoch 22 / 30 is complete in 4126.7728724479675 seconds with a training accuracy of 84.736
Accuracy of predictions on the test set is 63.23
Accuracy of predictions on the test set is 62.62
==========
Epoch 23 / 30 is complete in 4334.613925933838 seconds with a training accuracy of 86.648
Epoch 23 / 30 is complete in 4334.629611492157 seconds with a training accuracy of 86.862
Accuracy of predictions on the test set is 63.08
Accuracy of predictions on the test set is 62.69
==========
Epoch 24 / 30 is complete in 4502.381586074829 seconds with a training accuracy of 87.454
Epoch 24 / 30 is complete in 4502.397217750549 seconds with a training accuracy of 87.816
Accuracy of predictions on the test set is 63.58
Accuracy of predictions on the test set is 62.94
==========
Epoch 25 / 30 is complete in 4690.178122758865 seconds with a training accuracy of 88.072
Epoch 25 / 30 is complete in 4690.162544250488 seconds with a training accuracy of 88.162
Accuracy of predictions on the test set is 63.51
Accuracy of predictions on the test set is 62.45
==========
Epoch 26 / 30 is complete in 4877.974770307541 seconds with a training accuracy of 88.436
Epoch 26 / 30 is complete in 4877.990465641022 seconds with a training accuracy of 88.674
Accuracy of predictions on the test set is 63.79
Accuracy of predictions on the test set is 62.46
==========
Epoch 27 / 30 is complete in 5065.719406366348 seconds with a training accuracy of 89.25

Epoch 27 / 30 is complete in 5065.7037987709045 seconds with a training accuracy of 88.986
Accuracy of predictions on the test set is 63.69
Accuracy of predictions on the test set is 62.31
==========
Epoch 28 / 30 is complete in 5253.542608499527 seconds with a training accuracy of 89.21
Epoch 28 / 30 is complete in 5253.55822968483 seconds with a training accuracy of 89.458
Accuracy of predictions on the test set is 63.75
Accuracy of predictions on the test set is 62.89
==========
Epoch 29 / 30 is complete in 5441.231100797653 seconds with a training accuracy of 89.808
Epoch 29 / 30 is complete in 5441.215489864349 seconds with a training accuracy of 89.686
Accuracy of predictions on the test set is 63.72
Accuracy of predictions on the test set is 62.45
==========
Epoch 30 / 30 is complete in 5629.042144775391 seconds with a training accuracy of 90.088
Epoch 30 / 30 is complete in 5629.026530265808 seconds with a training accuracy of 89.858
Accuracy of predictions on the test set is 63.4
Accuracy of predictions on the test set is 62.75
==========
Model is saved. Total time consumed is 5640.066656827927 seconds
Model is saved. Total time consumed is 5640.401250362396 seconds

Learning Curve for Task 4: