ECE421

2021

Assignment Four

Chase McDougall

Engineering Science

Part 1

1.

Since the CIFAR10 dataset consists of 60,000 images and we are starting with a training ratio of 0.8 we have the following number of images in our training set:

$$60000 \cdot 0.8 = 48000$$

Training with batch sizes of 32 would require $48000 \div 32 = 1500$ iterations to go through the entire training set. To perform 30 epochs we would require $1500 \cdot 30 = 45000$ iterations.

2.

```python
def my_conv_net(image, weights):
  N, Ci, H, W = image.shape
  k, k2, Ci2, Co = weights.shape
  Ho = H-k+1
  Wo = W-k+1
  out = np.empty(shape=(N,Co,Ho,Wo))
  for n in range(N):
    for c in range(Co):
      for h in range(Ho):
        for w in range(Wo):
          image_colours = image[n,:,h,w]
          weight_slice = weights[0,0,:,c]
          outpoint = jn.multiply(image_colours, weight_slice)
          out[n,c,h,w] = jn.sum(outpoint)
  return out

my_conv_output = my_conv_net(my_image, filter_weights)
```

3.

```python
def my_lin_layer(w,b,X):
  out = 0
  out += b[0]
  out += jn.sum(jn.multiply(jn.transpose(w),X))
  return (out)
```

Note: Includes Bias for completeness though we are assuming a bias of zero

Single line alternative:

```python
jn.sum(jn.multiply(jn.transpose(w), output_rearranged))
```

4.

The training set is used for actually training the model and hence defining the weights and bias of the given model, while the validation set is used in tuning the hyperparameters of the given model. The validation set comes into the tuning of hyperparameters by being used to provide accuracy values during the training of our models.
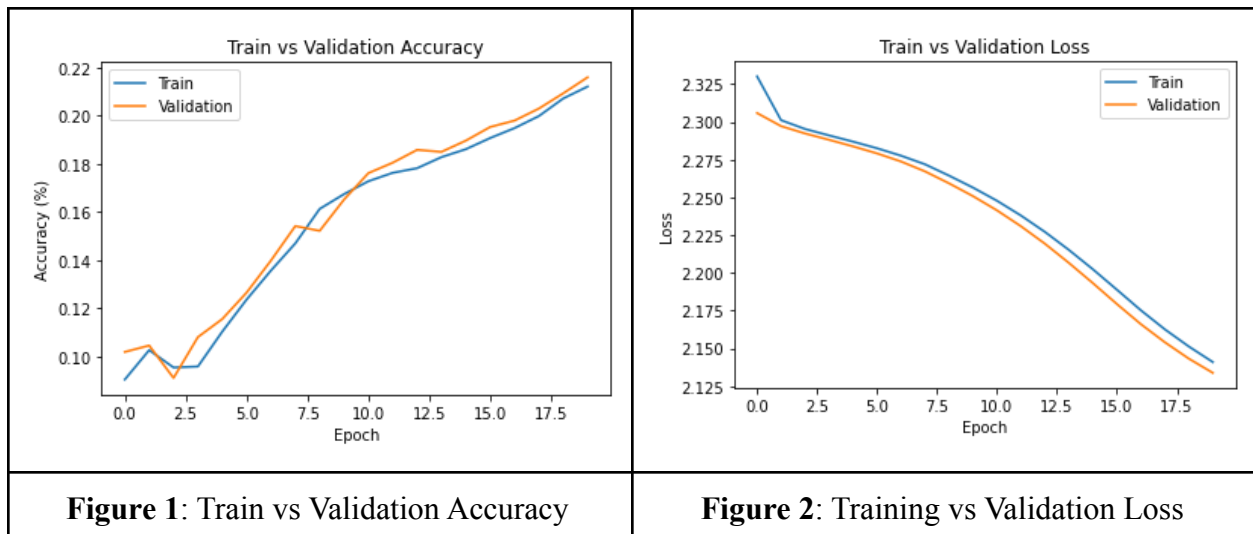
Part 2
1.
```
new_params = params.value - lr*grad
params.assign(new_params)
```

2.
```
batch = <TYPE>_indices[it:it+BATCH]
```

3.
Default Parameters

| **Figure 1**: Train vs Validation Accuracy | **Figure 2**: Training vs Validation Loss |

Full printout can be found in the Appendix.

4.
A hyperparameter is a parameter which is set prior to starting training. These hyperparameter(s) influence the way in which our model will learn from data. In contrast, other training parameters (such as weights or biases) are determined during training.

If you try to evaluate the accuracy before tuning hyperparameters you will end up with a misleading view of the model, since any training done before tuning would not be the desired representation.
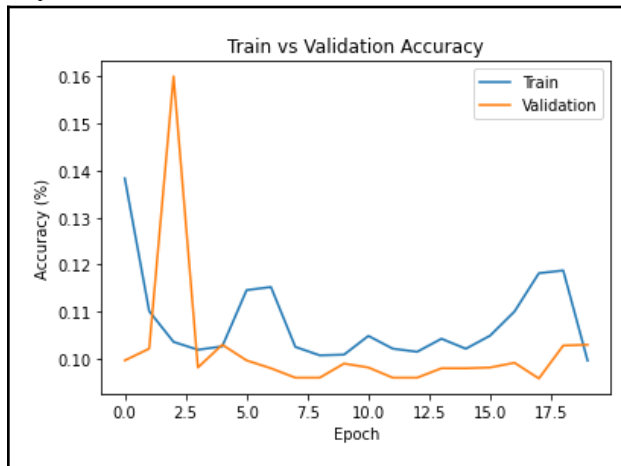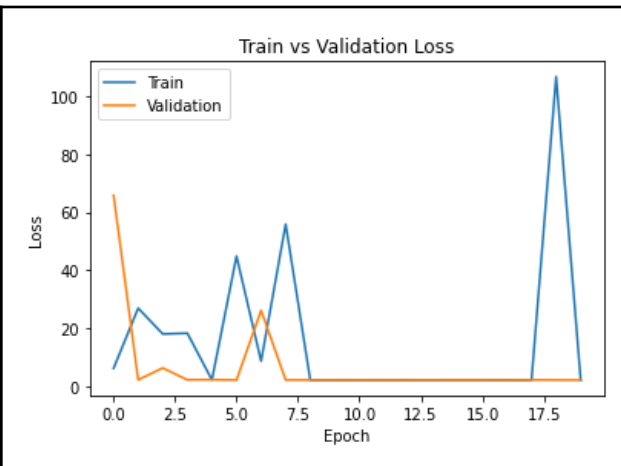
5.
H = {batch_size, learning_rate, # of classes from linear transformation, # of convolutional layers}
$$H_{default} = \{32, 9*10^{-4}, 10, 2\}$$
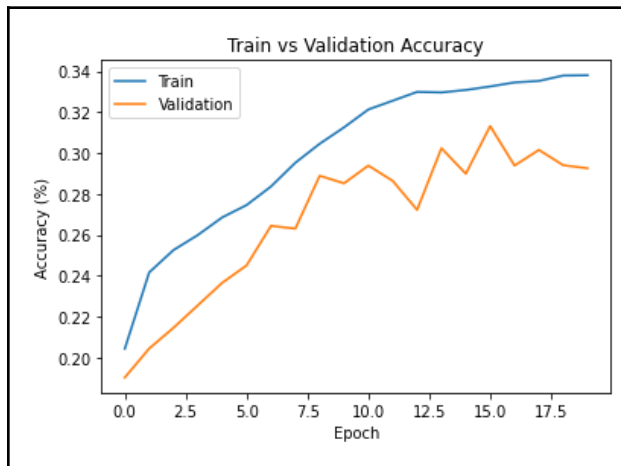$$H_1 = \{35, 0.01, 9, 3\}$$
$$H_2 = \{40, 0.1, 15, 1\}$$

6.

M_1

| **Figure 3**: Train vs Validation Accuracy | **Figure 4**: Training vs Validation Loss |

$H_1 = \{35, 0.01, 9, 3\}$

Full printout of data can be found in the Appendix.

M_2

| **Figure 5**: Train vs Validation Accuracy | **Figure 6**: Training vs Validation Loss |
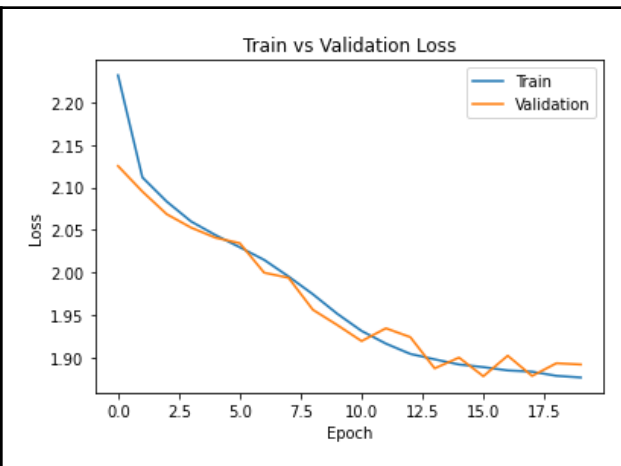
$H_2 = \{40, 0.1, 15, 1\}$

Full printout of data can be found in the Appendix.

Best Validation Validation Accuracy and Time:

Base Model	Model 1	Model 2
21.60% at Epoch 20	16.00% at Epoch 3	31.32% at Epoch 16
13m 14s	15m 18s	7m 8s

*Times are meant to be used as relative measures of speed

7.
Based on the validation accuracy indicated in the previous question I would pick Model 2, since we see a considerable increase in validation accuracy. Further analysis of the time requirement shows that not only did Model 2 obtain higher validation accuracy values, but also did so in a considerably reduced time. We saw a 47% reduction in compute time when going from the Base Model to Model 2. This can lead to considerable cost and/or time savings when running larger datasets.

The final test accuracy of Model 2 is 30.62%
Performing the same test evaluation on the Base Model returns a Base Test Accuracy of 20.47%.

Therefore we find that Model 2 has a considerably higher Test Accuracy than the provided Base Model.
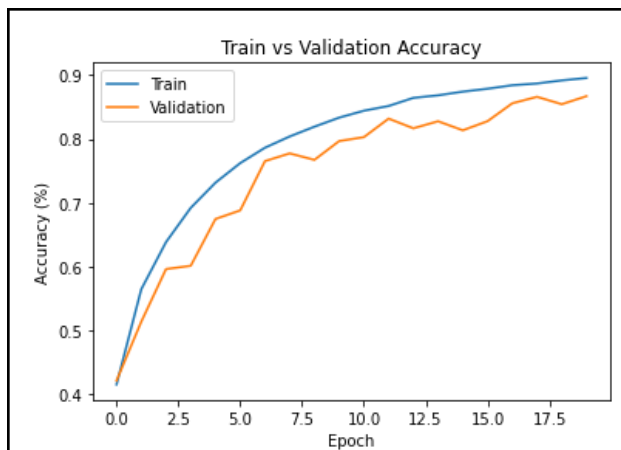
Part 3

1.

Using MNIST dataset.

2.

Using the same base model used in part 1 of the question we get the following outputs:

| **Figure 7**: Train vs Validation Accuracy | **Figure 8**: Training vs Validation Loss |
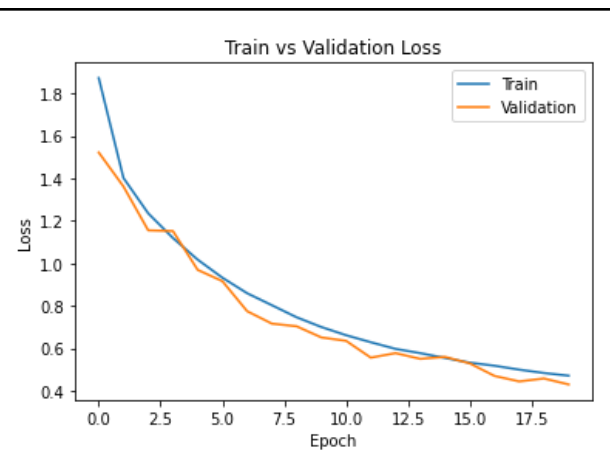
Over 20 Epochs

	Peak Validation Accuracy	Test Accuracy
Base Model	89.69% at Epoch 20	86.53%

3.

To determine which parameters to tune for optimal predictions we will start by tweaking a couple. We saw in Part 2 that increasing the learning rate led to an increase in accuracy over the same number of epochs . We will first start by adjusting the learning rate as follows:

Learning Rate Testing:

Learning Rate	Peak Validation Accuracy	Test Accuracy
Base Model (0.0009)	89.69% at Epoch 20	86.53%
0.001	85.81% at Epoch 20	85.34%
0.01	92.47% at Epoch 19	91.34%
0.1	10.67% at Epoch 19	11.17%

From trying different learning rates while holding other factors constant, we see that a learning rate of 0.01 appears to provide the best Validation and Test Accuracy. Exceeding the Base Model Test Accuracy by nearly 5%.

Activation Function Testing:

Activation Function	Peak Validation Accuracy	Test Accuracy
Base Model (Relu)	89.69% at Epoch 20	86.53%
Elu	85.59% at Epoch 20	85.40%
Selu	87.26% at Epoch 19	86.59%
Log Sigmoid	70.53% at Epoch 19	71.26%

Looking at the different accuracies from varying the activation function, we see marginal differences in accuracy though the Base Model (Relu) is slightly better in terms of Validation accuracy, so we will choose Relu as the best tested Activation Function.

Batch Size Testing:

Batch Size	Peak Validation Accuracy	Test Accuracy
Base Model (32)	89.69% at Epoch 20	86.53%
20	88.34% at Epoch 20	88.29%
25	88.43% at Epoch 20	87.97%
40	82.23% at Epoch 20	82.65%

While we don't see any increases in the Validation accuracy of our model, we see a nearly 2% jump when decreasing the batch size to 20. Therefore we'll select a smaller size of 20 to provide Test Accuracy improvements, though not too small as to significantly increase run times.

Number of Convolutional Layers Testing:

# of Convolutional Layers	Peak Validation Accuracy	Test Accuracy
Base Model (2)	89.69% at Epoch 20	86.53%
1	47.67% at Epoch 20	46.73%
3	95.56% at Epoch 20	95.04%

We see that adding an additional convolution layer gives us a considerate bump in Accuracy. Therefore we will choose 3 convolutional layers to be the best of the tested options.

Using the the following found parameters in our optimized model:

Learning Rate	Activation Function	Batch Size	# of Convolutional Layers
0.01	Relu	20	3

We get the following results:

	Peak Validation Accuracy	Test Accuracy
Base Model	89.69% at Epoch 20	86.53%
Optimized Model	96.81% at Epoch 18	96.64%

From the above results we can see that the tuning of these four parameters resulted in approximately a 10% increase in the model's Test Accuracy. This is a considerable improvement especially when it comes to increasing accuracy in the upper percentages.

In addition to looking at the peak and fully trained accuracies it is interesting to look at earlier accuracies as well. In the following table we have the Validation accuracies after the first few Epochs:

Validation Accuracies:

Epoch #	Base Model	Optimized Model
1	42.14%	86.17%
2	51.37%	93.57%
3	59.61%	94.40%

Looking at these results we see that not only did the final optimized model have better performance than the base model, but it also achieved high accuracies much earlier on in training. This means that by using a better optimized model (such as the one presented here) we can end up saving compute time by doing fewer Epochs to achieve similar if not better accuracies, since the optimized model shows comparable validation accuracies after a single epoch as the base model did for a fully trained model.

Part 4

Hyperparameter exploration

Here we will further explore 3 additional hyperparameters for our Convolutional Neural Network. These hyperparameters are:

Convolution Kernel Size

# of Output Channels from the First Convolutional Layer

Convolution Dilation

For each hyperparameter we tune we will run two trials: one for a larger value and one for a smaller value (as compared to the base model). If the base model is already the smallest possible value we will instead run two trials with larger values. The final optimized model is chosen based on the values of each hyperparameter that led to the highest Test Accuracy.

This means that we ran a total of 6 trials for the various hyperparameters plus 2 other trials for the Base Model and Final Optimized Model for a total of 8 trained models.

Continuing with the same base model we have the following base accuracies:

	Peak Validation Accuracy	Test Accuracy
Base Model	89.69% at Epoch 20	86.53%

Size of the Convolution Kernel	Peak Validation Accuracy	Test Accuracy
Base Model (2x2)	89.69% at Epoch 20	86.53%
1x1	24.11% at Epoch 17	24.36%
3x3	95.19% at Epoch 20	95.04%

We see that by increasing the Kernel Size from 2x2 to 3x3 we get a roughly 5% jump in peak validation accuracy and nearly 9% jump in Test Accuracy.

# of Output Channels on the First Convolution Layer	Peak Validation Accuracy	Test Accuracy
Base Model (16)	89.69% at Epoch 20	86.53%
10	87.70% at Epoch 20	87.46%
20	84.34% at Epoch 19	82.83%

These trials show that the base model is better in terms of optimal Validation Accuracy.

Convolution Dilation	Peak Validation Accuracy	Test Accuracy
Base Model (1)	89.69% at Epoch 20	86.53%
2	93.23% at Epoch 17	93.69%
3	93.97% at Epoch 16	93.33%

The above trials show that by increasing the Dilation of each convolution yields a higher Validation and Test Accuracy.

Combining the optimal hyperparameters from these trials we get the following:
We will define the "optimal hyperparameters" to be those that provided the highest test accuracy.

	Size of Convolution Kernel	# of Output Channels of First Convolution Layer	Convolution Dilation
Base Model	2x2	16	1
Optimized Model	3x3	10	2

Now training a model to access these hyperparameter values provides:

	Peak Validation Accuracy	Test Accuracy
Base Model	89.69% at Epoch 20	86.53%
Optimized Model	95.99% at Epoch 18	95.76%

Here we see approximately a 6% increase in peak validation accuracy, while we see a roughly 9% increase in Test Accuracy.

Furthermore, we can look at the accuracies of earlier Epochs:
Validation Accuracies:

Epoch #	Base Model	Optimized Model
1	42.14%	83.89%
2	51.37%	89.19%
3	59.61%	91.67%

We again see that the optimized model can achieve significantly higher degrees of accuracy at earlier epochs. This shows that there is potential to save significant amounts of time and money in training if one does not necessarily require a fully trained model.

However, when looking at the fully trained models we notice that the ~9% increase in Test Accuracy is significant enough to make the Optimized Model an excellent alternative to the base model, as we can expect to see improved classification performance.

Appendix

Base Model Printout:

Epoch 0001 Training Loss 2.33 Validation Loss 2.31 Training Accuracy 9.04 Validation Accuracy 10.18

Epoch 0002 Training Loss 2.30 Validation Loss 2.30 Training Accuracy 10.26 Validation Accuracy 10.45

Epoch 0003 Training Loss 2.30 Validation Loss 2.29 Training Accuracy 9.54 Validation Accuracy 9.10

Epoch 0004 Training Loss 2.29 Validation Loss 2.29 Training Accuracy 9.58 Validation Accuracy 10.80

Epoch 0005 Training Loss 2.29 Validation Loss 2.28 Training Accuracy 11.03 Validation Accuracy 11.55

Epoch 0006 Training Loss 2.28 Validation Loss 2.28 Training Accuracy 12.36 Validation Accuracy 12.65

Epoch 0007 Training Loss 2.28 Validation Loss 2.27 Training Accuracy 13.57 Validation Accuracy 13.98

Epoch 0008 Training Loss 2.27 Validation Loss 2.27 Training Accuracy 14.70 Validation Accuracy 15.42

Epoch 0009 Training Loss 2.26 Validation Loss 2.26 Training Accuracy 16.13 Validation Accuracy 15.22

Epoch 0010 Training Loss 2.26 Validation Loss 2.25 Training Accuracy 16.74 Validation Accuracy 16.52

Epoch 0011 Training Loss 2.25 Validation Loss 2.24 Training Accuracy 17.28 Validation Accuracy 17.62

Epoch 0012 Training Loss 2.24 Validation Loss 2.23 Training Accuracy 17.63 Validation Accuracy 18.05

Epoch 0013 Training Loss 2.23 Validation Loss 2.22 Training Accuracy 17.82 Validation Accuracy 18.58

Epoch 0014 Training Loss 2.22 Validation Loss 2.21 Training Accuracy 18.28 Validation Accuracy 18.50

Epoch 0015 Training Loss 2.20 Validation Loss 2.19 Training Accuracy 18.61 Validation Accuracy 18.97

Epoch 0016 Training Loss 2.19 Validation Loss 2.18 Training Accuracy 19.07 Validation Accuracy 19.53

Epoch 0017 Training Loss 2.18 Validation Loss 2.17 Training Accuracy 19.49 Validation Accuracy 19.80

Epoch 0018 Training Loss 2.16 Validation Loss 2.15 Training Accuracy 19.99 Validation Accuracy 20.30

Epoch 0019 Training Loss 2.15 Validation Loss 2.14 Training Accuracy 20.72 Validation Accuracy 20.93

Epoch 0020 Training Loss 2.14 Validation Loss 2.13 Training Accuracy 21.21 Validation Accuracy 21.60

Base Model with lr=0.01 Printout:

Epoch 0001 Training Loss 2.26 Validation Loss 2.20 Training Accuracy 16.81 Validation Accuracy 18.62

Epoch 0002 Training Loss 2.15 Validation Loss 2.11 Training Accuracy 20.67 Validation Accuracy 21.33

Epoch 0003 Training Loss 2.10 Validation Loss 2.09 Training Accuracy 22.28 Validation Accuracy 21.85

Epoch 0004 Training Loss 2.09 Validation Loss 2.08 Training Accuracy 23.35 Validation Accuracy 23.12

Epoch 0005 Training Loss 2.07 Validation Loss 2.07 Training Accuracy 23.85 Validation Accuracy 22.62

Epoch 0006 Training Loss 2.06 Validation Loss 2.05 Training Accuracy 24.73 Validation Accuracy 24.28

Epoch 0007 Training Loss 2.05 Validation Loss 2.05 Training Accuracy 25.52 Validation Accuracy 24.05

Epoch 0008 Training Loss 2.03 Validation Loss 2.02 Training Accuracy 26.64 Validation Accuracy 26.28

Epoch 0009 Training Loss 2.02 Validation Loss 2.01 Training Accuracy 27.15 Validation Accuracy 26.88

Epoch 0010 Training Loss 2.00 Validation Loss 2.00 Training Accuracy 27.66 Validation Accuracy 26.42

Epoch 0011 Training Loss 1.99 Validation Loss 1.99 Training Accuracy 28.18 Validation Accuracy 28.00

Epoch 0012 Training Loss 1.98 Validation Loss 1.97 Training Accuracy 28.89 Validation Accuracy 27.50

Epoch 0013 Training Loss 1.96 Validation Loss 1.96 Training Accuracy 29.35 Validation Accuracy 27.87

Epoch 0014 Training Loss 1.95 Validation Loss 1.95 Training Accuracy 30.20 Validation Accuracy 28.82

Epoch 0015 Training Loss 1.94 Validation Loss 1.94 Training Accuracy 30.55 Validation Accuracy 28.77

Epoch 0016 Training Loss 1.92 Validation Loss 1.92 Training Accuracy 31.28 Validation Accuracy 29.57

Epoch 0017 Training Loss 1.91 Validation Loss 1.90 Training Accuracy 31.75 Validation Accuracy 30.07

Epoch 0018 Training Loss 1.90 Validation Loss 1.89 Training Accuracy 32.15 Validation Accuracy 29.77

Epoch 0019 Training Loss 1.88 Validation Loss 1.88 Training Accuracy 32.76 Validation Accuracy 30.92

Epoch 0020 Training Loss 1.87 Validation Loss 1.88 Training Accuracy 33.27 Validation Accuracy 29.52

Model 1 Printout:

Epoch 0001 Training Loss 6.27 Validation Loss 65.80 Training Accuracy 13.83 Validation Accuracy 9.97

Epoch 0002 Training Loss 26.95 Validation Loss 2.29 Training Accuracy 11.01 Validation Accuracy 10.22

Epoch 0003 Training Loss 18.07 Validation Loss 6.36 Training Accuracy 10.36 Validation Accuracy 16.00

Epoch 0004 Training Loss 18.34 Validation Loss 2.29 Training Accuracy 10.19 Validation Accuracy 9.82

Epoch 0005 Training Loss 2.26 Validation Loss 2.34 Training Accuracy 10.26 Validation Accuracy 10.30

Epoch 0006 Training Loss 44.86 Validation Loss 2.19 Training Accuracy 11.46 Validation Accuracy 9.97

Epoch 0007 Training Loss 8.71 Validation Loss 26.19 Training Accuracy 11.53 Validation Accuracy 9.80

Epoch 0008 Training Loss 55.86 Validation Loss 2.22 Training Accuracy 10.25 Validation Accuracy 9.60

Epoch 0009 Training Loss 2.21 Validation Loss 2.21 Training Accuracy 10.07 Validation Accuracy 9.60

Epoch 0010 Training Loss 2.22 Validation Loss 2.22 Training Accuracy 10.09 Validation Accuracy 9.90

Epoch 0011 Training Loss 2.22 Validation Loss 2.21 Training Accuracy 10.49 Validation Accuracy 9.82

Epoch 0012 Training Loss 2.22 Validation Loss 2.22 Training Accuracy 10.22 Validation Accuracy 9.60

Epoch 0013 Training Loss 2.22 Validation Loss 2.22 Training Accuracy 10.15 Validation Accuracy 9.60

Epoch 0014 Training Loss 2.23 Validation Loss 2.23 Training Accuracy 10.43 Validation Accuracy 9.80

Epoch 0015 Training Loss 2.23 Validation Loss 2.24 Training Accuracy 10.21 Validation Accuracy 9.80

Epoch 0016 Training Loss 2.23 Validation Loss 2.24 Training Accuracy 10.49 Validation Accuracy 9.82

Epoch 0017 Training Loss 2.22 Validation Loss 2.21 Training Accuracy 11.01 Validation Accuracy 9.92

Epoch 0018 Training Loss 2.22 Validation Loss 2.26 Training Accuracy 11.82 Validation Accuracy 9.58

Epoch 0019 Training Loss 106.68 Validation Loss 2.22 Training Accuracy 11.88 Validation Accuracy 10.28

Epoch 0020 Training Loss 2.18 Validation Loss 2.18 Training Accuracy 9.96 Validation Accuracy 10.30

Model 2 Printout:

Epoch 0001 Training Loss 2.23 Validation Loss 2.13 Training Accuracy 20.43 Validation Accuracy 19.02

Epoch 0002 Training Loss 2.11 Validation Loss 2.10 Training Accuracy 24.16 Validation Accuracy 20.43

Epoch 0003 Training Loss 2.08 Validation Loss 2.07 Training Accuracy 25.25 Validation Accuracy 21.45

Epoch 0004 Training Loss 2.06 Validation Loss 2.05 Training Accuracy 25.99 Validation Accuracy 22.55

Epoch 0005 Training Loss 2.04 Validation Loss 2.04 Training Accuracy 26.85 Validation Accuracy 23.65

Epoch 0006 Training Loss 2.03 Validation Loss 2.03 Training Accuracy 27.45 Validation Accuracy 24.50

Epoch 0007 Training Loss 2.01 Validation Loss 2.00 Training Accuracy 28.36 Validation Accuracy 26.43

Epoch 0008 Training Loss 2.00 Validation Loss 1.99 Training Accuracy 29.52 Validation Accuracy 26.30

Epoch 0009 Training Loss 1.97 Validation Loss 1.96 Training Accuracy 30.45 Validation Accuracy 28.88

Epoch 0010 Training Loss 1.95 Validation Loss 1.94 Training Accuracy 31.25 Validation Accuracy 28.52

Epoch 0011 Training Loss 1.93 Validation Loss 1.92 Training Accuracy 32.13 Validation Accuracy 29.38

Epoch 0012 Training Loss 1.92 Validation Loss 1.93 Training Accuracy 32.56 Validation Accuracy 28.63

Epoch 0013 Training Loss 1.90 Validation Loss 1.92 Training Accuracy 32.99 Validation Accuracy 27.22

Epoch 0014 Training Loss 1.90 Validation Loss 1.89 Training Accuracy 32.96 Validation Accuracy 30.23

Epoch 0015 Training Loss 1.89 Validation Loss 1.90 Training Accuracy 33.08 Validation Accuracy 28.98

Epoch 0016 Training Loss 1.89 Validation Loss 1.88 Training Accuracy 33.25 Validation Accuracy 31.32

Epoch 0017 Training Loss 1.89 Validation Loss 1.90 Training Accuracy 33.45 Validation Accuracy 29.38
Epoch 0018 Training Loss 1.88 Validation Loss 1.88 Training Accuracy 33.52 Validation Accuracy 30.15
Epoch 0019 Training Loss 1.88 Validation Loss 1.89 Training Accuracy 33.79 Validation Accuracy 29.40
Epoch 0020 Training Loss 1.88 Validation Loss 1.89 Training Accuracy 33.80 Validation Accuracy 29.25

Part 3 Appendix

Base Model Printout:

```
Epoch 0001   Training Loss 1.87 Validation Loss 1.52 Training Accuracy 41.55 Validation Accuracy 42.14
Epoch 0002   Training Loss 1.40 Validation Loss 1.36 Training Accuracy 56.46 Validation Accuracy 51.37
Epoch 0003   Training Loss 1.24 Validation Loss 1.15 Training Accuracy 63.84 Validation Accuracy 59.61
Epoch 0004   Training Loss 1.12 Validation Loss 1.15 Training Accuracy 69.20 Validation Accuracy 60.13
Epoch 0005   Training Loss 1.02 Validation Loss 0.97 Training Accuracy 73.15 Validation Accuracy 67.46
Epoch 0006   Training Loss 0.93 Validation Loss 0.92 Training Accuracy 76.21 Validation Accuracy 68.80
Epoch 0007   Training Loss 0.86 Validation Loss 0.77 Training Accuracy 78.61 Validation Accuracy 76.51
Epoch 0008   Training Loss 0.80 Validation Loss 0.72 Training Accuracy 80.38 Validation Accuracy 77.73
Epoch 0009   Training Loss 0.75 Validation Loss 0.70 Training Accuracy 81.91 Validation Accuracy 76.71
Epoch 0010   Training Loss 0.70 Validation Loss 0.65 Training Accuracy 83.33 Validation Accuracy 79.66
Epoch 0011   Training Loss 0.66 Validation Loss 0.64 Training Accuracy 84.41 Validation Accuracy 80.27
Epoch 0012   Training Loss 0.63 Validation Loss 0.56 Training Accuracy 85.15 Validation Accuracy 83.17
Epoch 0013   Training Loss 0.60 Validation Loss 0.58 Training Accuracy 86.42 Validation Accuracy 81.66
Epoch 0014   Training Loss 0.58 Validation Loss 0.55 Training Accuracy 86.82 Validation Accuracy 82.76
Epoch 0015   Training Loss 0.55 Validation Loss 0.56 Training Accuracy 87.40 Validation Accuracy 81.34
Epoch 0016   Training Loss 0.53 Validation Loss 0.53 Training Accuracy 87.85 Validation Accuracy 82.77
Epoch 0017   Training Loss 0.52 Validation Loss 0.47 Training Accuracy 88.40 Validation Accuracy 85.57
Epoch 0018   Training Loss 0.50 Validation Loss 0.44 Training Accuracy 88.67 Validation Accuracy 86.59
Epoch 0019   Training Loss 0.48 Validation Loss 0.46 Training Accuracy 89.16 Validation Accuracy 85.43
Epoch 0020   Training Loss 0.47 Validation Loss 0.43 Training Accuracy 89.54 Validation Accuracy 86.69
```

LR = 0.01

```
Epoch 0001   Training Loss 1.61 Validation Loss 0.99 Training Accuracy 55.60 Validation Accuracy 66.40
Epoch 0002   Training Loss 0.88 Validation Loss 0.76 Training Accuracy 78.04 Validation Accuracy 74.37
Epoch 0003   Training Loss 0.68 Validation Loss 0.53 Training Accuracy 84.49 Validation Accuracy 82.14
Epoch 0004   Training Loss 0.56 Validation Loss 0.44 Training Accuracy 87.81 Validation Accuracy 85.69
Epoch 0005   Training Loss 0.50 Validation Loss 0.45 Training Accuracy 89.55 Validation Accuracy 85.54
Epoch 0006   Training Loss 0.45 Validation Loss 0.41 Training Accuracy 90.83 Validation Accuracy 86.41
Epoch 0007   Training Loss 0.43 Validation Loss 0.37 Training Accuracy 91.87 Validation Accuracy 88.33
Epoch 0008   Training Loss 0.40 Validation Loss 0.35 Training Accuracy 92.41 Validation Accuracy 88.70
Epoch 0009   Training Loss 0.38 Validation Loss 0.57 Training Accuracy 92.97 Validation Accuracy 80.96
Epoch 0010   Training Loss 0.36 Validation Loss 0.33 Training Accuracy 93.48 Validation Accuracy 89.09
Epoch 0011   Training Loss 0.34 Validation Loss 0.32 Training Accuracy 93.81 Validation Accuracy 89.59
Epoch 0012   Training Loss 0.33 Validation Loss 0.29 Training Accuracy 94.05 Validation Accuracy 90.37
Epoch 0013   Training Loss 0.32 Validation Loss 0.29 Training Accuracy 94.25 Validation Accuracy 90.43
Epoch 0014   Training Loss 0.32 Validation Loss 0.30 Training Accuracy 94.54 Validation Accuracy 90.54
Epoch 0015   Training Loss 0.30 Validation Loss 0.39 Training Accuracy 94.93 Validation Accuracy 87.26
Epoch 0016   Training Loss 0.29 Validation Loss 0.28 Training Accuracy 94.97 Validation Accuracy 90.74
Epoch 0017   Training Loss 0.29 Validation Loss 0.47 Training Accuracy 95.09 Validation Accuracy 84.79
Epoch 0018   Training Loss 0.28 Validation Loss 0.29 Training Accuracy 95.28 Validation Accuracy 90.24
Epoch 0019   Training Loss 0.28 Validation Loss 0.23 Training Accuracy 95.41 Validation Accuracy 92.47
Epoch 0020   Training Loss 0.27 Validation Loss 0.25 Training Accuracy 95.46 Validation Accuracy 91.91
```

Selu

```
Epoch 0001   Training Loss 2.06 Validation Loss 1.52 Training Accuracy 39.43 Validation Accuracy 44.87
Epoch 0002   Training Loss 1.38 Validation Loss 1.22 Training Accuracy 59.58 Validation Accuracy 58.56
Epoch 0003   Training Loss 1.15 Validation Loss 1.04 Training Accuracy 68.60 Validation Accuracy 66.13
Epoch 0004   Training Loss 1.01 Validation Loss 0.91 Training Accuracy 73.87 Validation Accuracy 71.07
Epoch 0005   Training Loss 0.90 Validation Loss 0.85 Training Accuracy 77.32 Validation Accuracy 72.33
Epoch 0006   Training Loss 0.82 Validation Loss 0.79 Training Accuracy 79.92 Validation Accuracy 74.23
Epoch 0007   Training Loss 0.76 Validation Loss 0.72 Training Accuracy 81.75 Validation Accuracy 76.76
Epoch 0008   Training Loss 0.70 Validation Loss 0.73 Training Accuracy 83.23 Validation Accuracy 74.71
Epoch 0009   Training Loss 0.66 Validation Loss 0.68 Training Accuracy 84.33 Validation Accuracy 77.53
Epoch 0010   Training Loss 0.63 Validation Loss 0.63 Training Accuracy 85.28 Validation Accuracy 79.56
Epoch 0011   Training Loss 0.60 Validation Loss 0.60 Training Accuracy 86.30 Validation Accuracy 80.71
Epoch 0012   Training Loss 0.57 Validation Loss 0.54 Training Accuracy 87.09 Validation Accuracy 83.66
Epoch 0013   Training Loss 0.55 Validation Loss 0.50 Training Accuracy 87.79 Validation Accuracy 85.09
Epoch 0014   Training Loss 0.53 Validation Loss 0.51 Training Accuracy 88.09 Validation Accuracy 84.01
Epoch 0015   Training Loss 0.51 Validation Loss 0.48 Training Accuracy 88.67 Validation Accuracy 85.23
Epoch 0016   Training Loss 0.49 Validation Loss 0.51 Training Accuracy 89.02 Validation Accuracy 83.54
Epoch 0017   Training Loss 0.48 Validation Loss 0.44 Training Accuracy 89.41 Validation Accuracy 86.90
Epoch 0018   Training Loss 0.46 Validation Loss 0.43 Training Accuracy 89.72 Validation Accuracy 86.59
Epoch 0019   Training Loss 0.45 Validation Loss 0.42 Training Accuracy 90.10 Validation Accuracy 87.26
Epoch 0020   Training Loss 0.44 Validation Loss 0.42 Training Accuracy 90.34 Validation Accuracy 86.96
```

Batch Size of 40

```
Epoch 0001   Training Loss 2.03 Validation Loss 1.66 Training Accuracy 35.87 Validation Accuracy 36.89
Epoch 0002   Training Loss 1.53 Validation Loss 1.41 Training Accuracy 50.96 Validation Accuracy 49.77
Epoch 0003   Training Loss 1.35 Validation Loss 1.23 Training Accuracy 58.89 Validation Accuracy 58.46
Epoch 0004   Training Loss 1.21 Validation Loss 1.12 Training Accuracy 64.67 Validation Accuracy 62.80
Epoch 0005   Training Loss 1.11 Validation Loss 1.06 Training Accuracy 68.44 Validation Accuracy 62.99
Epoch 0006   Training Loss 1.03 Validation Loss 0.95 Training Accuracy 71.17 Validation Accuracy 68.61
Epoch 0007   Training Loss 0.96 Validation Loss 0.96 Training Accuracy 73.42 Validation Accuracy 65.66
Epoch 0008   Training Loss 0.91 Validation Loss 0.99 Training Accuracy 75.02 Validation Accuracy 66.21
Epoch 0009   Training Loss 0.86 Validation Loss 0.85 Training Accuracy 76.64 Validation Accuracy 72.13
Epoch 0010   Training Loss 0.82 Validation Loss 0.78 Training Accuracy 77.72 Validation Accuracy 73.94
Epoch 0011   Training Loss 0.79 Validation Loss 0.77 Training Accuracy 78.92 Validation Accuracy 73.63
Epoch 0012   Training Loss 0.76 Validation Loss 0.70 Training Accuracy 79.91 Validation Accuracy 77.16
Epoch 0013   Training Loss 0.73 Validation Loss 0.81 Training Accuracy 80.64 Validation Accuracy 71.89
Epoch 0014   Training Loss 0.71 Validation Loss 0.70 Training Accuracy 81.42 Validation Accuracy 76.33
Epoch 0015   Training Loss 0.69 Validation Loss 0.67 Training Accuracy 82.13 Validation Accuracy 78.16
Epoch 0016   Training Loss 0.67 Validation Loss 0.67 Training Accuracy 82.76 Validation Accuracy 78.41
Epoch 0017   Training Loss 0.65 Validation Loss 0.61 Training Accuracy 83.34 Validation Accuracy 80.71
Epoch 0018   Training Loss 0.63 Validation Loss 0.61 Training Accuracy 83.74 Validation Accuracy 80.99
Epoch 0019   Training Loss 0.62 Validation Loss 0.63 Training Accuracy 84.22 Validation Accuracy 79.64
Epoch 0020   Training Loss 0.60 Validation Loss 0.57 Training Accuracy 84.67 Validation Accuracy 82.23
```

Part 3 Optimized Model

```
Epoch 0001   Training Loss 0.82 Validation Loss 0.43 Training Accuracy 83.14 Validation Accuracy 86.17
Epoch 0002   Training Loss 0.29 Validation Loss 0.21 Training Accuracy 97.10 Validation Accuracy 93.57
Epoch 0003   Training Loss 0.22 Validation Loss 0.18 Training Accuracy 98.19 Validation Accuracy 94.40
Epoch 0004   Training Loss 0.19 Validation Loss 0.23 Training Accuracy 98.51 Validation Accuracy 92.86
Epoch 0005   Training Loss 0.17 Validation Loss 0.21 Training Accuracy 98.75 Validation Accuracy 93.86
Epoch 0006   Training Loss 0.16 Validation Loss 0.21 Training Accuracy 98.93 Validation Accuracy 93.39
Epoch 0007   Training Loss 0.15 Validation Loss 0.13 Training Accuracy 98.99 Validation Accuracy 96.04
Epoch 0008   Training Loss 0.14 Validation Loss 0.15 Training Accuracy 99.11 Validation Accuracy 95.69
Epoch 0009   Training Loss 0.13 Validation Loss 0.13 Training Accuracy 99.15 Validation Accuracy 96.00
Epoch 0010   Training Loss 0.12 Validation Loss 0.12 Training Accuracy 99.25 Validation Accuracy 96.37
Epoch 0011   Training Loss 0.12 Validation Loss 0.13 Training Accuracy 99.32 Validation Accuracy 96.34
Epoch 0012   Training Loss 0.12 Validation Loss 0.16 Training Accuracy 99.33 Validation Accuracy 95.46
Epoch 0013   Training Loss 0.11 Validation Loss 0.12 Training Accuracy 99.47 Validation Accuracy 96.31
Epoch 0014   Training Loss 0.11 Validation Loss 0.12 Training Accuracy 99.39 Validation Accuracy 96.49
Epoch 0015   Training Loss 0.11 Validation Loss 0.13 Training Accuracy 99.49 Validation Accuracy 96.14
Epoch 0016   Training Loss 0.10 Validation Loss 0.15 Training Accuracy 99.49 Validation Accuracy 95.59
Epoch 0017   Training Loss 0.10 Validation Loss 0.13 Training Accuracy 99.52 Validation Accuracy 96.11
Epoch 0018   Training Loss 0.10 Validation Loss 0.10 Training Accuracy 99.58 Validation Accuracy 96.81
Epoch 0019   Training Loss 0.09 Validation Loss 0.11 Training Accuracy 99.54 Validation Accuracy 96.63
Epoch 0020   Training Loss 0.09 Validation Loss 0.11 Training Accuracy 99.60 Validation Accuracy 96.81
```