# ECGR 5105 Homework 6: Neural Networks

Lauren Bourque (Student ID: 801264851)

November 26, 2023

## GitHub Link

Click here to view the code

## Problem 1a.

**Training and Validation Loss Results:**

The loss values for the regression from Homework 5 with a learning rate of 0.001 can be seen below.

Figure 1: Homework 5 Loss Values



The loss values for the neural network can be seen below. As can be seen from the results from Figure 2 compared to the results from Figure 1, the neural network achieves an eventual better training loss but slightly worse validation loss than the linear regression model from Homework 5. The neural network has $(12{\times}32{+}32){+}(32{\times}1{+}1) = 417$ trainable parameters compared to the 12 parameters from the linear regression model.

Figure 2: Neural Network Loss Values

# Problem 1b.

**Training and Validation Loss Results:**

The loss values for the neural network with 3 hidden layers can be seen below.

Figure 3: Neural Network with 3 Layers Loss Values

```
Epoch 0/5000, Training Loss: 1.186692237854004, Validation Loss: 1.1106796264648438
Epoch 500/5000, Training Loss: 0.26614847779273987, Validation Loss: 0.41631677746772766
Epoch 1000/5000, Training Loss: 0.571448028087616, Validation Loss: 0.46390071511268616
Epoch 1500/5000, Training Loss: 0.3825329542160034, Validation Loss: 0.468228816986084
Epoch 2000/5000, Training Loss: 0.07898759841918945, Validation Loss: 0.4665388762950897
Epoch 2500/5000, Training Loss: 0.10419765114784241, Validation Loss: 0.4629688858985901
Epoch 3000/5000, Training Loss: 0.08431097865104675, Validation Loss: 0.4802996814250946
Epoch 3500/5000, Training Loss: 0.19947503507137299, Validation Loss: 0.5038023591041565
Epoch 4000/5000, Training Loss: 0.04506227746605873, Validation Loss: 0.530514121055603
Epoch 4500/5000, Training Loss: 0.08443551510572433, Validation Loss: 0.5555590391159058
Epoch 5000/5000, Training Loss: 0.07330196350812912, Validation Loss: 0.5717246532440186
```

The neural network with 3 hidden layers has a much better training loss than both the neural network with 1 hidden layer and the linear regression model. However, it has a worse validation loss than both the neural network with 1 hidden layer and the linear regression model. There are $(12 \times 32 + 32) + (32 \times 64 + 64) + (64 \times 16 + 16) + (16 \times 1 + 1) = 9409$ parameters versus the 417 trainable parameters from problem 1a and the 12 trainable parameters from the linear regression model.

# Problem 2a.

**Training Time, Training Loss, and Evaluation Accuracy Results:**

The total training time was 8241 seconds or 2.289 hours.
The training loss was 0.5180834466234192 at epoch 291.
The final evaluation accuracy was 49.72%, but the best recorded accuracy was 50.49%
The F1 score was found to be 0.4902.
The confusion matrix can be seen below:

Figure 4: Problem 2a Confusion Matrix

```
Final Confusion Matrix:
[[584  35  64  29  37  21  22  35 122  51]
 [ 42 566  38  42  15  24  33  36  67 137]
 [ 89  25 386 100 130  72  89  74  20  15]
 [ 36  23  97 322  68 195 124  53  35  47]
 [ 50   9 158  71 417  64 103  79  30  19]
 [ 36  18  97 192  87 374  65  74  34  23]
 [ 18  17 112 105 101  73 530  11  15  18]
 [ 62  23  77  70  83  77  27 522  19  40]
 [ 83  51  26  34  27  37   9  14 672  47]
 [ 71 164  15  25  21  22  35  44  72 531]]
```

# Problem 2b.

**Training Time, Training Loss, and Evaluation Accuracy Results:**

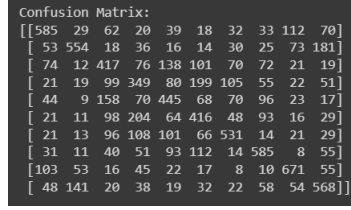The total training time was 8654 seconds or 2.404 hours.
The training loss was 0.009872898049152378 at epoch 291.
The final evaluation accuracy was 51.40%, but the best recorded accuracy was 54.01%
The F1 score was found to be 0.5128.
The confusion matrix can be seen below:

Figure 5: Problem 2b Confusion Matrix

```
Confusion Matrix:
[[585  29  62  20  39  18  32  33 112  70]
 [ 53 554  18  36  16  14  30  25  73 181]
 [ 74  12 417  76 138 101  70  72  21  19]
 [ 21  19  99 349  80 199 105  55  22  51]
 [ 44   9 158  70 445  68  70  96  23  17]
 [ 21  11  98 204  64 416  48  93  16  29]
 [ 21  13  96 108 101  66 531  14  21  29]
 [ 31  11  40  51  93 112  14 585   8  55]
 [103  53  16  45  22  17   8  10 671  55]
 [ 48 141  20  38  19  32  22  58  54 568]]
```

The model had a better accuracy than the model from problem 2a, showing that adding more hidden layers improved the accuracy.

The model in 2a had $(3072 \times 512 + 512) + (512 \times 10 + 10) = 1578506$ parameters, while the model in 2b had $(3072 \times 512 + 512) + (512 \times 256 + 256) + (256 \times 128 + 128) + (128 \times 10 + 10) = 1748890$ parameters.