# ECGR 4105 Homework 5

Cole Bennett

801289645

October 25, 2025

## GitHub Link

https://github.com/oopCole/IntroToMachineLearning/tree/main/HW5

## Problem 1.

Learning rate at 0.1.

```
Training Learning Rate: 0.1
epoch   500 | loss 2.131903 | lr 0.1
epoch  1000 | loss 2.125364 | lr 0.1
epoch  1500 | loss 2.119942 | lr 0.1
epoch  2000 | loss 2.115369 | lr 0.1
epoch  2500 | loss 2.111510 | lr 0.1
epoch  3000 | loss 2.108257 | lr 0.1
epoch  3500 | loss 2.105512 | lr 0.1
epoch  4000 | loss 2.103198 | lr 0.1
epoch  4500 | loss 2.101246 | lr 0.1
epoch  5000 | loss 2.099599 | lr 0.1
```

Learning rate at 0.01.

```
Training Learning Rate: 0.01
epoch   500 | loss 17.333752 | lr 0.01
epoch  1000 | loss 6.317653 | lr 0.01
epoch  1500 | loss 3.279615 | lr 0.01
epoch  2000 | loss 2.441494 | lr 0.01
epoch  2500 | loss 2.210000 | lr 0.01
epoch  3000 | loss 2.145788 | lr 0.01
epoch  3500 | loss 2.127707 | lr 0.01
epoch  4000 | loss 2.122352 | lr 0.01
epoch  4500 | loss 2.120515 | lr 0.01
epoch  5000 | loss 2.119651 | lr 0.01
```

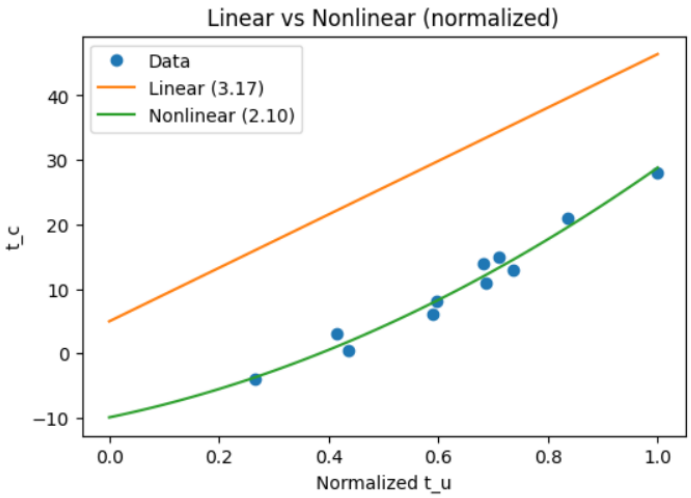
Learning rate at 0.001.

```
Training Learning Rate: 0.001
epoch   500 | loss 48.301525 | lr 0.001
epoch  1000 | loss 37.399002 | lr 0.001
epoch  1500 | loss 32.932919 | lr 0.001
epoch  2000 | loss 29.202843 | lr 0.001
epoch  2500 | loss 25.930910 | lr 0.001
epoch  3000 | loss 23.054548 | lr 0.001
epoch  3500 | loss 20.525681 | lr 0.001
epoch  4000 | loss 18.302317 | lr 0.001
epoch  4500 | loss 16.347551 | lr 0.001
epoch  5000 | loss 14.628925 | lr 0.001
```

Learning rate at 0.0001.

```
Training Learning Rate: 0.0001
epoch   500 | loss 169.676285 | lr 0.0001
epoch  1000 | loss 136.590195 | lr 0.0001
epoch  1500 | loss 112.525436 | lr 0.0001
epoch  2000 | loss 94.977486 | lr 0.0001
epoch  2500 | loss 82.137650 | lr 0.0001
epoch  3000 | loss 72.699776 | lr 0.0001
epoch  3500 | loss 65.720772 | lr 0.0001
epoch  4000 | loss 60.519421 | lr 0.0001
epoch  4500 | loss 56.603916 | lr 0.0001
epoch  5000 | loss 53.618935 | lr 0.0001
```

Linear vs Nonlinear pot.




## Problem 2.

Learning rate at 0.1.

```
Training Learning Rate: 0.1
epoch   500 | train 0.437832 | val 0.743571 | lr 0.1
epoch  1000 | train 0.437832 | val 0.743571 | lr 0.1
epoch  1500 | train 0.437832 | val 0.743571 | lr 0.1
epoch  2000 | train 0.437832 | val 0.743571 | lr 0.1
epoch  2500 | train 0.437832 | val 0.743571 | lr 0.1
epoch  3000 | train 0.437832 | val 0.743571 | lr 0.1
epoch  3500 | train 0.437832 | val 0.743571 | lr 0.1
epoch  4000 | train 0.437832 | val 0.743571 | lr 0.1
epoch  4500 | train 0.437832 | val 0.743571 | lr 0.1
epoch  5000 | train 0.437832 | val 0.743571 | lr 0.1
```

Learning rate at 0.01.

```
Training Learning Rate: 0.01
epoch   500 | train 0.437841 | val 0.742912 | lr 0.01
epoch  1000 | train 0.437832 | val 0.743570 | lr 0.01
epoch  1500 | train 0.437832 | val 0.743571 | lr 0.01
epoch  2000 | train 0.437832 | val 0.743571 | lr 0.01
epoch  2500 | train 0.437832 | val 0.743571 | lr 0.01
epoch  3000 | train 0.437832 | val 0.743571 | lr 0.01
epoch  3500 | train 0.437832 | val 0.743571 | lr 0.01
epoch  4000 | train 0.437832 | val 0.743571 | lr 0.01
epoch  4500 | train 0.437832 | val 0.743571 | lr 0.01
epoch  5000 | train 0.437832 | val 0.743571 | lr 0.01
```

Learning rate at 0.001.

```
Training Learning Rate: 0.001
epoch   500 | train 2.104640 | val 2.834990 | lr 0.001
epoch  1000 | train 0.828557 | val 1.132878 | lr 0.001
epoch  1500 | train 0.537014 | val 0.804240 | lr 0.001
epoch  2000 | train 0.463697 | val 0.741267 | lr 0.001
epoch  2500 | train 0.444669 | val 0.733784 | lr 0.001
epoch  3000 | train 0.439654 | val 0.736245 | lr 0.001
epoch  3500 | train 0.438320 | val 0.739176 | lr 0.001
epoch  4000 | train 0.437963 | val 0.741134 | lr 0.001
epoch  4500 | train 0.437867 | val 0.742267 | lr 0.001
epoch  5000 | train 0.437842 | val 0.742886 | lr 0.001
```

Learning rate at 0.0001.

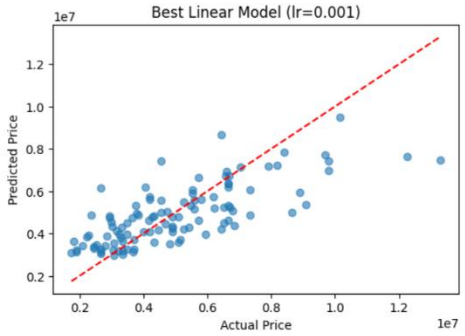
```
Training Learning Rate: 0.0001
epoch   500 | train 8.842999 | val 12.765259 | lr 0.0001
epoch  1000 | train 7.228313 | val 10.383260 | lr 0.0001
epoch  1500 | train 5.984525 | val 8.539974 | lr 0.0001
epoch  2000 | train 5.013280 | val 7.098298 | lr 0.0001
epoch  2500 | train 4.244779 | val 5.959111 | lr 0.0001
epoch  3000 | train 3.629064 | val 5.050177 | lr 0.0001
epoch  3500 | train 3.130044 | val 4.318377 | lr 0.0001
epoch  4000 | train 2.721360 | val 3.724301 | lr 0.0001
epoch  4500 | train 2.383546 | val 3.238401 | lr 0.0001
epoch  5000 | train 2.102040 | val 2.838310 | lr 0.0001
```

Baseline linear model comparison.

```
Final val losses
lr=0.1 | val 0.743571
lr=0.01 | val 0.743571
lr=0.001 | val 0.742886
lr=0.0001 | val 2.838310

Best lr=0.001 | val 0.742886
```

Best linear model plot.



## Problem 3.

Fully connected neural network with one hidden layer.

Fully connected neural network with three hidden layers.

```
Model 1: One Hidden Layer
epoch   50 | train loss: 0.426109 | val loss: 0.743059 | r²: 0.5467
epoch  100 | train loss: 0.400184 | val loss: 0.726423 | r²: 0.5569
epoch  150 | train loss: 0.393532 | val loss: 0.727658 | r²: 0.5561
epoch  200 | train loss: 0.382810 | val loss: 0.721488 | r²: 0.5599

Train time: 0.21s | Final Train Loss: 0.382810 | Final val loss: 0.721488 | r²: 0.5599

Model 2: Three Hidden Layers
epoch   50 | train loss: 0.432784 | val loss: 0.750481 | r²: 0.5422
epoch  100 | train loss: 0.388295 | val loss: 0.737615 | r²: 0.5500
epoch  150 | train loss: 0.340441 | val loss: 0.757707 | r²: 0.5378
epoch  200 | train loss: 0.311748 | val loss: 0.776771 | r²: 0.5262

Train Time: 0.31s | Final Train Loss: 0.311748 | Final val loss: 0.776771 | r²: 0.5262
```

The train time took 0.1 seconds longer with less train loss. The final val loss was barely higher.