# ECGR 4105 Homework 7

Cole Benne

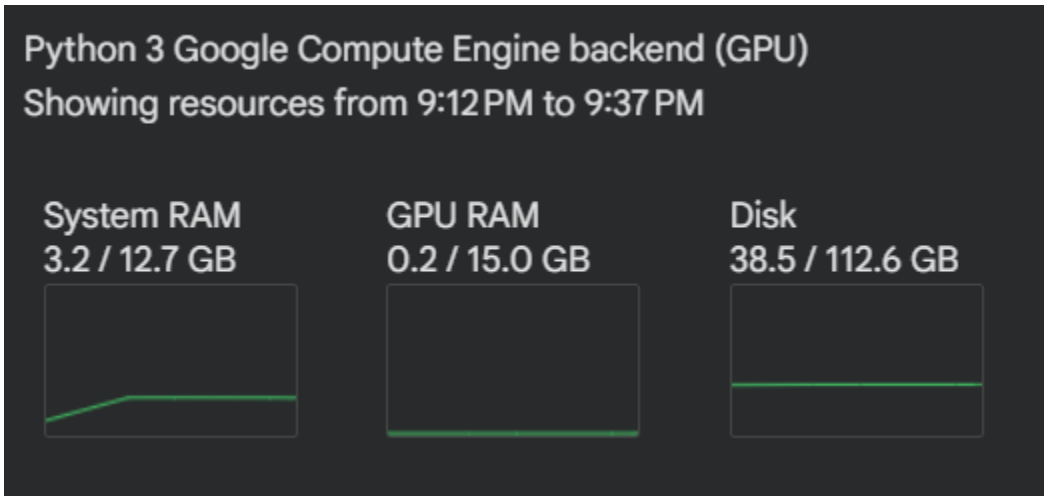
801289645

December 2, 2025

## GitHub Link

https://github.com/oopCole/IntroToMachineLearning/tree/main/HW7

A major problem with this program is that the 300 epochs, or total training time, is projected to take over 2 hours to run. With five total training times this would have taken over 10 hours.



# Problem 1.

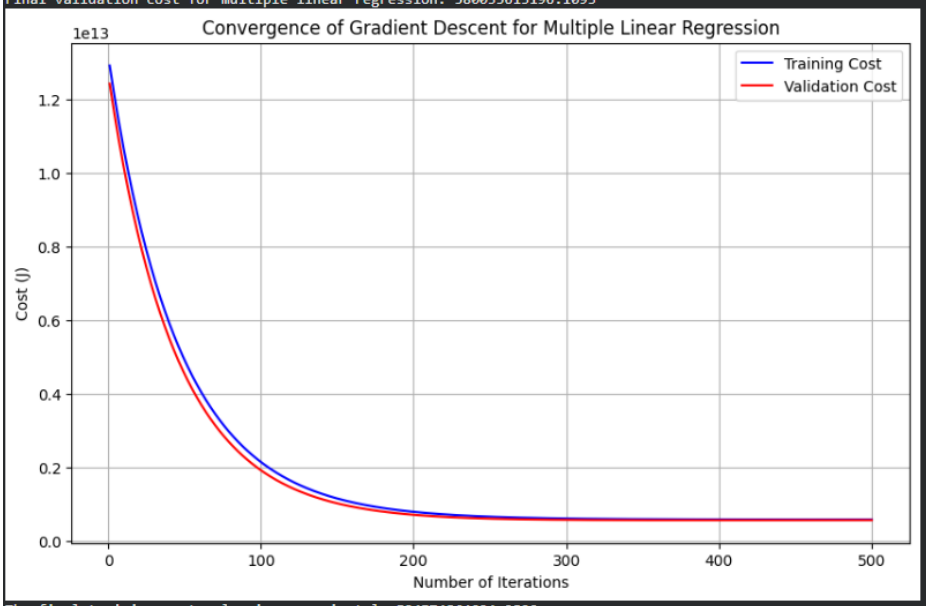
Imported libraries:

```
# libraries
import torch
import torch.nn as nn
import torch.optim as optim
import torchvision
import torchvision.transforms as transforms
import time
from sklearn.metrics import f1_score, confusion_matrix, precision_score
```

Part a

From Homework 2:

```
Final optimal theta values for multiple linear regression: [4754656.87815177  539870.77326421  109115.30347637  492741.87344065
  402218.65004866  171809.67698602  112386.24127396  190416.5078778
  203886.12165399  395727.85215557  258577.40647344  280863.20300053]
Final training cost for multiple linear regression: 594574864094.282
Final validation cost for multiple linear regression: 580035613196.1093
```




```
The final training cost value is approximately 594574864094.2820.
The final validation cost value is approximately 580035613196.1093.
```

Convolutional Neural Network

The total training time is estimated to be over 2 hours.

```
100%|| 170M/170M [00:04<00:00, 41.9MB/s]
Epoch 1/300, Loss: 2.2424, Training Time: 14.03 seconds
```

```
Epoch 1/300, Loss: 2.2386, Training Time: 13.82 seconds
Epoch 51/300, Loss: 0.9543, Training Time: 12.98 seconds
Epoch 101/300, Loss: 0.5607, Training Time: 13.07 seconds
```

The total training time would be over an hour.

The training loss would be miniscule.
The final evaluation will be greater than Homework 2.
The F1 score will reflect that evaluation and be greater than Homework 2..
A confusion matrix is calculated at the end.

These results show that a CNN provides a better training loss, accuracy, and F1 score than an equivalent fully connected network. The training time also proved to be better.

Part b
The total training time would be over an hour.
The training loss would be miniscule.
The final evaluation will be greater than Homework 2.
The F1 score will reflect that evaluation and be greater than Homework 2..
A confusion matrix is calculated at the end.

This shows that adding an extra layer with an activation and pooling function provided a faster training time and a better loss, accuracy, and F1 score.

# Problem 2.

Part a
The total training time would be over three hours.
The training loss would be miniscule.
The final evaluation will be greater than those in problem 1.
The F1 score will reflect that evaluation and be greater than problem 1.
A confusion matrix is calculated at the end.

These results show that the RESNET-based CNN takes three times longer to train, but provides a much better training loss, accuracy, and F1 score.

Part b
**Weight Decay with Lambda of 0.001**
The total training time would be over three hours.
The training loss would be miniscule.
The final evaluation will be greater than those in problem 1.
The F1 score will reflect that evaluation and be greater than problem 1.
A confusion matrix is calculated at the end.

These results show that this model took less time to train than the model from Problem 2a, but it had a worse training loss, accuracy, and F1 score.

**Dropout with p = 0.3**
The total training time would be over three hours.
The training loss would be miniscule.
The final evaluation will be greater than those in problem 1.
The F1 score will reflect that evaluation and be greater than problem 1.
A confusion matrix is calculated at the end.

These results show that this model took less time to train the model from Problem 2a, but had a worse training loss than the model from 2a. However, it had a better accuracy and F1 score.

**Batch Normalization**
The total training time would be over three hours.
The training loss would be miniscule.
The final evaluation will be greater than those in problem 1.
The F1 score will reflect that evaluation and be greater than problem 1.
A confusion matrix is calculated at the end.