

Hyperparameters:

Regularization strength: 1e-3

NUM_HIDDEN: 20

batchSize: 256

epochCount: 25

Annealing learning rate: 1e-4, /10 every 100 epochs

Results:

```
CORRECT GRAD CONFIRMATION =====
0.0
0.0
3.642919299551295e-16
0.0

batch 1: TrainingLoss = 61.875272940766195    TestingLoss = 84.53519694539742
batch 2: TrainingLoss = 62.9983004278822      TestingLoss = 84.506202681058
batch 3: TrainingLoss = 66.10439681496739    TestingLoss = 83.03301050832528
batch 4: TrainingLoss = 64.68813321433598    TestingLoss = 85.79178641509858
batch 5: TrainingLoss = 68.16247356434616    TestingLoss = 89.35851258308415
batch 6: TrainingLoss = 70.93676895923306    TestingLoss = 81.98345203045392
batch 7: TrainingLoss = 60.13110504460267    TestingLoss = 85.12334552269564
batch 8: TrainingLoss = 60.588704678379386    TestingLoss = 87.40181599877042
batch 9: TrainingLoss = 59.13346792161422    TestingLoss = 92.50151191945261
batch 10: TrainingLoss = 67.85079092458186    TestingLoss = 94.09500059195695
batch 11: TrainingLoss = 82.81753385878457    TestingLoss = 91.26568317565871
batch 12: TrainingLoss = 75.1848416086813    TestingLoss = 89.11471787405095
batch 13: TrainingLoss = 69.47130362209072    TestingLoss = 91.0001849278244
batch 14: TrainingLoss = 67.25890466947016    TestingLoss = 81.8774420105119
batch 15: TrainingLoss = 65.66382059224965    TestingLoss = 82.6946222454682
batch 16: TrainingLoss = 51.24208791200338    TestingLoss = 83.04175122688748
batch 17: TrainingLoss = 61.31898451049639    TestingLoss = 81.9081993296788
batch 18: TrainingLoss = 58.977242317683555    TestingLoss = 84.80595698222548
batch 19: TrainingLoss = 60.787172533776356    TestingLoss = 95.05648794166301
batch 20: TrainingLoss = 73.42378518617186    TestingLoss = 81.41920021886722

final testing accuracy: 81.41920021886722
final training accuracy: 58.83724701330778
```

Weight Vectors:







