

Part 1

Hyperparameters:

```
Hyperparameters
RUNNING VALIDATION 0
FPC value: 0.8894
Cross Entropy: 0.44573515784462253
Best Values So Far:
PC: 0.8894
Neurons: 30
Learning Rate: 0.001
Minibatches: 64
Epochs: 40
Regularization Strength: 0.1
RUNNING VALIDATION 1
FPC value: 0.8706
Cross Entropy: 0.5639125988116117
RUNNING VALIDATION 2
FPC value: 0.9642
Cross Entropy: 0.13186710182535727
Best Values So Far:
PC: 0.9642
Neurons: 50
Learning Rate: 0.05
Minibatches: 32
Epochs: 40
Regularization Strength: 0.01
RUNNING VALIDATION 3
FPC value: 0.921
Cross Entropy: 0.3008006977366516
RUNNING VALIDATION 4
FPC value: 0.9764
Cross Entropy: 0.08960874346656626
Best Values So Far:
PC: 0.9764
Neurons: 50
Learning Rate: 0.1
Minibatches: 16
Epochs: 40
Regularization Strength: 0.001
RUNNING VALIDATION 5
FPC value: 0.2058
Cross Entropy: 1.946630368317069
RUNNING VALIDATION 6
FPC value: 0.9654
Cross Entropy: 0.1249438054992768
RUNNING VALIDATION 7
FPC value: 0.9708
Cross Entropy: 0.13195136369055957
RUNNING VALIDATION 8
FPC value: 0.9234
Cross Entropy: 0.2862626768273551
RUNNING VALIDATION 9
FPC value: 0.9436
Cross Entropy: 0.2206183586402416
```

The best values are:

```
PC: 0.9764
Neurons: 50
Learning Rate: 0.1
Minibatches: 16
Epochs: 40
Regularization Strength: 0.001
```

****NOTE:****

In our zipped file, we have three files of our HW6 (not including the 6 data files). One file is part 1, another is part 2, and the other is our pdf document.

This following is the order of all the functions in our code:

```
loadData
unpack
pack
fPC
softmax
relu
derivRelu
fCE
gradCE
forwardProp
backProp
train <- this is where we run our SGD
testing <- this is where we do the final test
on testing data
findBestHyperparameters
plotSGDPath
```

Running trained network on test set:

*** Epoch 21 Statistics ***
FPC = 0.9867636363636364
Cross Entropy = 0.043496344355615425

*** Epoch 22 Statistics ***
FPC = 0.9868
Cross Entropy = 0.04362344114328186

*** Epoch 23 Statistics ***
FPC = 0.9868727272727272
Cross Entropy = 0.0431135444799004

*** Epoch 24 Statistics ***
FPC = 0.9872
Cross Entropy = 0.042027408469194026

*** Epoch 25 Statistics ***
FPC = 0.9869636363636364
Cross Entropy = 0.04220663670909303

*** Epoch 26 Statistics ***
FPC = 0.9875636363636363
Cross Entropy = 0.041632061430159906

*** Epoch 27 Statistics ***
FPC = 0.9872181818181818
Cross Entropy = 0.042843411307491065

*** Epoch 28 Statistics ***
FPC = 0.9873272727272727
Cross Entropy = 0.041602181429454155

*** Epoch 29 Statistics ***
FPC = 0.9864727272727273
Cross Entropy = 0.04370441976343429

*** Epoch 30 Statistics ***
FPC = 0.9870363636363636
Cross Entropy = 0.04148656038231037

*** Epoch 31 Statistics ***
FPC = 0.9870727272727273
Cross Entropy = 0.0424283376240996

*** Epoch 32 Statistics ***
FPC = 0.9868727272727272
Cross Entropy = 0.04151134000554301

*** Epoch 33 Statistics ***
FPC = 0.9872181818181818
Cross Entropy = 0.04093880616608359

*** Epoch 34 Statistics ***
FPC = 0.9874181818181819
Cross Entropy = 0.04089377560898408

*** Epoch 35 Statistics ***
FPC = 0.9865454545454545
Cross Entropy = 0.041824056272197783

*** Epoch 36 Statistics ***
FPC = 0.9861818181818182
Cross Entropy = 0.04403824039177476

*** Epoch 37 Statistics ***
FPC = 0.9873636363636363
Cross Entropy = 0.03991584116725973

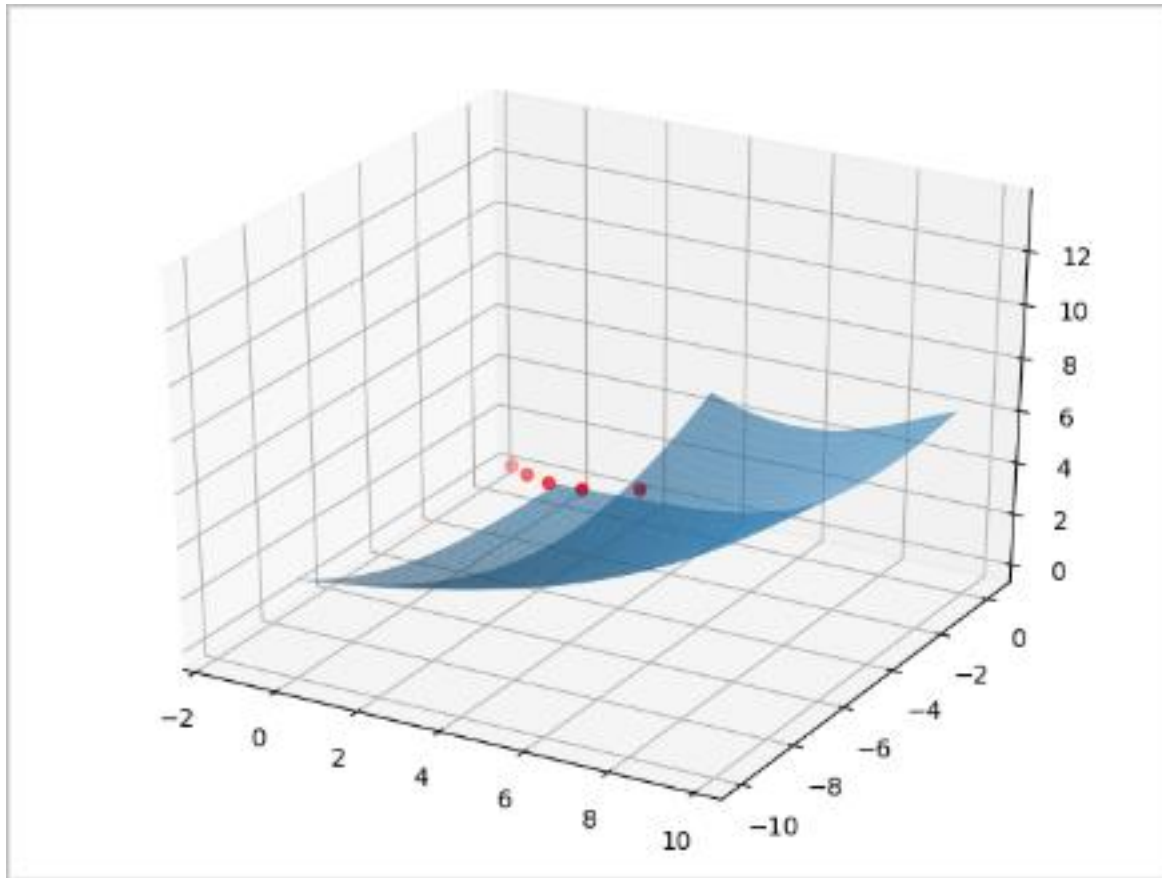
*** Epoch 38 Statistics ***
FPC = 0.9869818181818182
Cross Entropy = 0.04094409667005886

*** Epoch 39 Statistics ***
FPC = 0.9869818181818182
Cross Entropy = 0.04112114839693129

*** Epoch 40 Statistics ***
FPC = 0.9865272727272727
Cross Entropy = 0.04200076776303092

Final Test
FPC value: 0.9749
Cross Entropy: 0.08819793411754362

Part 2:



This is our final graph for part 2.

For part 2, first we have a WS list (a global variable) to store all of our w values. We append the W s for each epoch in our train function. Then, we found the 2 components using the built-in PCA function. Next, we adjusted our axis1 and axis2 of equally spaced points. We then iterated through the axis1 and axis2 points in the meshgrid to find all our costs and made this our Zaxis. To superimpose the plot, we set the components of PCA equal to the Xaxis and Yaxis. We used a similar approach as described above to find our Zaxis values. After all these steps, we were able to obtain the graph above.