Xiao Wang wang.xiao1@husky.neu.edu

Problem 1 + 2

Please see code attached in email.

Problem 3

a) My neural network classifier was unable to produce satisfactory results in finite time. But it performs correctly and pretty well on simple purposes (e.g. XOR function). Please try to run test.py to check the correctness of my NN code.

b) With tol=0.0001, C=1.0, max_iter=1000, the 1 vs all SVM classifier gives test accuracy of 0.95.

```
/Users/Breezen/anaconda3/bin/python /Users/Breezen/Study/NEU/Courses/ML/NEU-Machine-Learning/HW4/src/testB.py
0.95

Process finished with exit code 0
```

c) With tol=0.0001, C=1.0, max_iter=100, the 1 vs all Logistic Regression classifier gives test accuracy of 0.957142857143.

```
/Users/Breezen/anaconda3/bin/python /Users/Breezen/Study/NEU/Courses/ML/NEU-Machine-Learning/HW4/src/testC.py
0.957142857143

Process finished with exit code 0
```

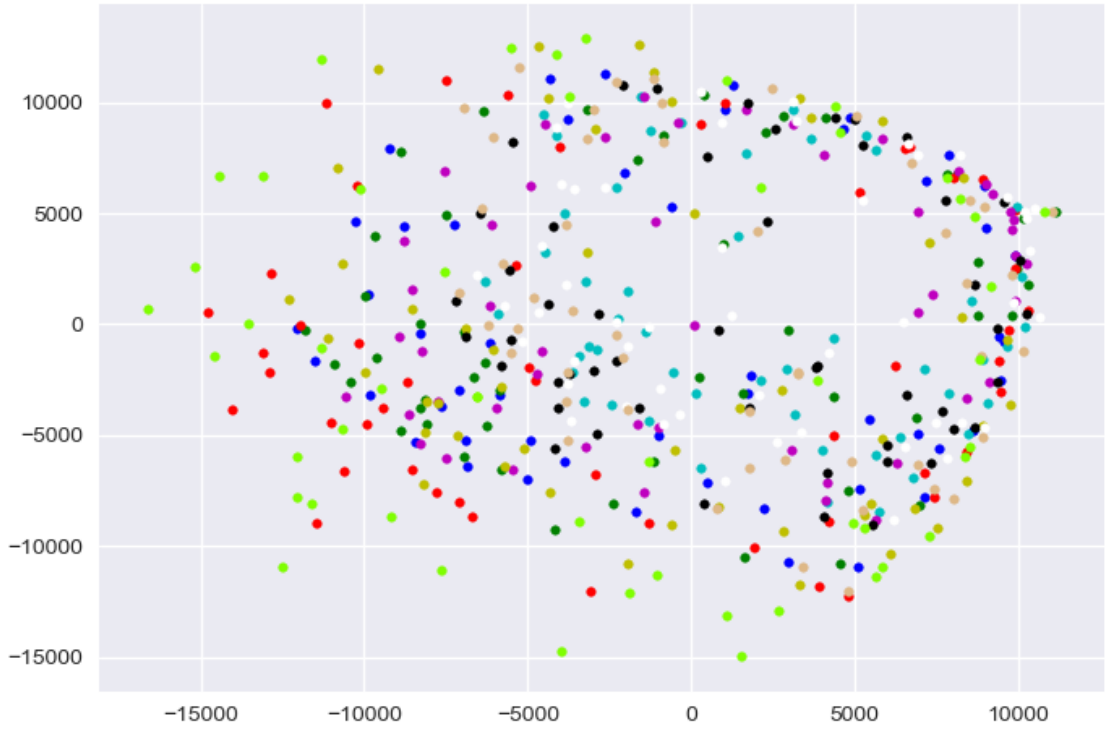
d) After applying PCA to reduce the dimension to d = 100,

```
/Users/Breezen/anaconda3/bin/python /Users/Breezen/Study/NEU/Courses/ML/NEU-Machine-Learning/HW4/src/testD.py
(500, 100)
(140, 100)
SVM accuracy:  0.0428571428571
LR accuracy:  0.0428571428571

Process finished with exit code 0
```

e)

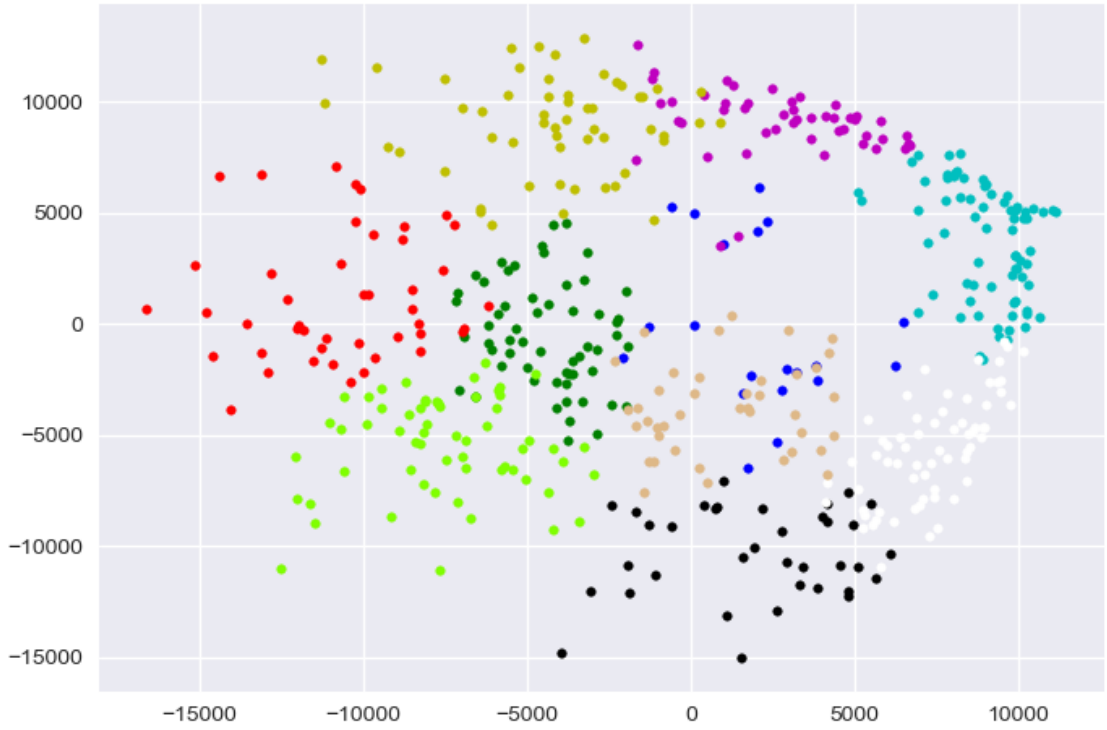
f) After applying PCA to d = 2, the data are not separated according to classes:



g) Kmeans failed in in recovering the true clustering of the data. Because the Euclidean distance between features of images cannot represent the similarity between them.

```
/Users/Breezen/anaconda3/bin/python /Users/Breezen/Study/NEU/Courses/ML/NEU-Machine-Learning/HW4/src/testG.py
Error ratio:  0.896

Process finished with exit code 0
```



h) gamma = 0.010000: Error ratio: 0.904000
gamma = 0.100000: Error ratio: 0.914000
gamma = 1.000000: Error ratio: 0.896000
gamma = 10.000000: Error ratio: 0.894000
gamma = 100.000000: Error ratio: 0.912000
Spectral Clustering performs no better than SVM. Because it's hard to tell the structure of the image data with RBF kernel only.