Experiment-1: iteration 3

Average mean square error: 648.987
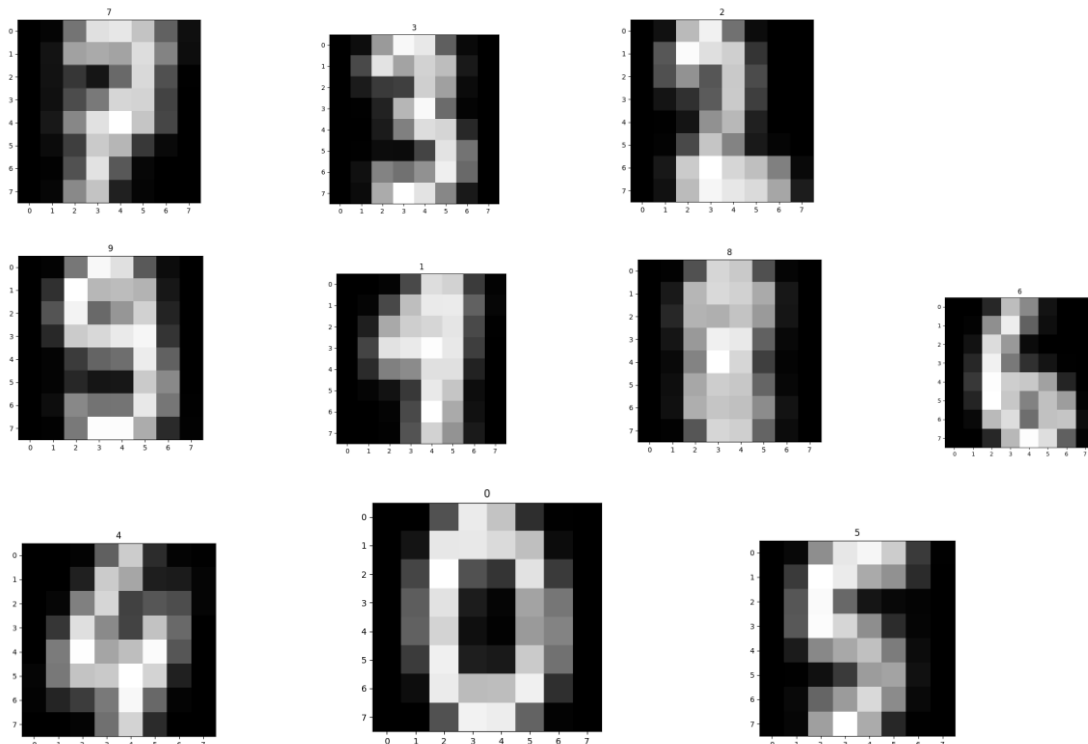
Mean square separation: 102075.979445

Mean entropy: 1.1120842487

Accuracy: 80.634390%

Confusion matrix:

```
[[ 176.    0.    0.    0.    2.    0.    0.    0.    0.    0.]
 [   0.   58.   21.    2.    0.    1.    4.    0.   96.    0.]
 [   1.    2.  149.    5.    0.    0.    0.    3.   14.    3.]
 [   0.    0.    1.  150.    0.    1.    0.    8.    7.   16.]
 [   0.    5.    0.    0.  162.    0.    0.    6.    8.    0.]
 [   0.    0.    0.    0.    1.  144.    1.    0.    0.   36.]
 [   1.    1.    0.    0.    1.    0.  176.    0.    2.    0.]
 [   0.    5.    0.    0.    1.    3.    0.  167.    3.    0.]
 [   0.    8.    1.    5.    0.    4.    1.    2.  125.   28.]
 [   0.   23.    0.    3.    0.    5.    0.    5.    2.  142.]]
```


Summary:

After 5 iterations I found that this result has minimum MSE. You can see all possible 10 classes of digits. Yes, I can see the digit to which cluster is associated. Because, ultimately the centers represent the

cluster. And clusters are formed using the input data. Thus, we can see features of centers as a proper digit. The iteration in which some clusters had zero elements, accuracy is not quite high because the clustering was not done right. I ignored such iteration for this experiment.

EXPERIMENT 2:

Average mean square error: 351.610000125

Mean square separation: 947917.783284

Mean entropy: 0.790478976868

Accuracy: 89.7050639955

Confusion matrix:

```
[[ 177.    0.    0.    0.    1.    0.    0.    0.    0.    0.]
 [   0.  143.   19.    1.    0.    1.    5.    0.   10.    3.]
 [   1.    4.  151.    3.    0.    0.    0.    4.   10.    4.]
 [   0.    0.    1.  147.    0.    3.    0.    8.    7.   17.]
 [   0.    6.    0.    0.  172.    0.    0.    2.    1.    0.]
 [   0.    0.    0.    0.    1.  169.    1.    0.    0.   11.]
 [   1.    1.    0.    0.    1.    1.  175.    0.    2.    0.]
 [   0.    0.    0.    0.    1.    0.    0.  168.    2.    8.]
 [   0.   16.    1.    1.    0.    1.    1.    0.  146.    8.]
 [   0.    3.    0.    3.    1.    3.    0.    3.    3.  164.]]
```

Summary:

I can clearly see all the digits. It is observed that accuracy increases when you increase the number of clusters. This is because, clusters get aligned closely to all the data points. Thus, mean square error is decreased, and as a result accuracy is increased. Entropy is decreased as desired because clusters are now aligned more closely to data points. Mean square separation is increased because clusters are now more separated than each other than the previous experiment.

Thus, we can say that, accuracy increases as the number of clusters increases. However, if multiple clusters classify one class, then it is not quite right. Thus we should find a trade-off between more accuracy and more generalization.

I have included all the clusters except those who had zero elements in their clusters.