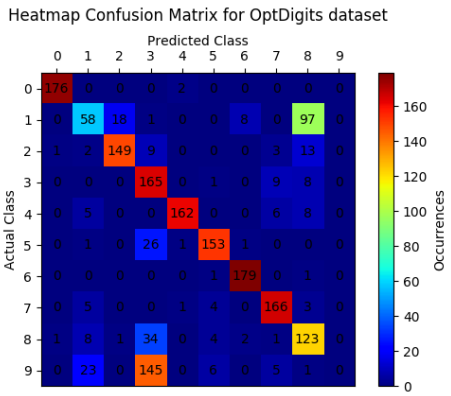
K-means Digit Generation

Joe Arriaga

Description

K-means was used to classify digits from the OptDigits data set. Here are graphical representations of the results.

Confusion Matrix



(a) Heatmap representation of confusion matrix

Confusion Matrix for OptDigits dataset

Actual Class \ Predicted Class	0	1	2	3	4	5	6	7	8	9
0	176	0	0	0	2	0	0	0	0	0
1	0	58	18	1	0	0	8	0	97	0
2	1	2	149	9	0	0	0	3	13	0
3	0	0	0	165	0	1	0	9	8	0
4	0	5	0	0	162	0	0	6	8	0
5	0	1	0	26	1	153	1	0	0	0
6	0	0	0	0	0	1	179	0	1	0
7	0	5	0	0	1	4	0	166	3	0
8	1	8	1	34	0	4	2	1	123	0
9	0	23	0	145	0	6	0	5	1	0

(b) Confusion matrix with values

Figure 1: Confusion matrix on test set after training has been completed.

Notice how some digits get confused ...

Digit Visualizations

After classifying digits using k-means the resulting centroids define the features the model used to classify digits. These could be considered to be "standard" or "ideal" digits. The centroids are visualized here.

The generated images are 8x8px but have been enlarged to 400x400px to be easier to see. However, the magnification sometimes makes it more difficult to see the form of the number. In these cases, smaller sizes are provided.

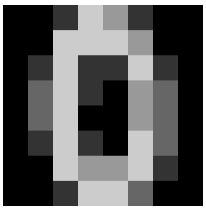
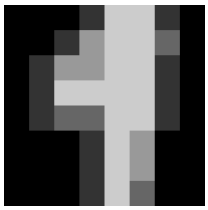


Figure 2: Generated image of a 0



(a) 400x400px

(b) 16x16px

Figure 3: Generated image of a 1 or 9. Examples of 1s and 9s were misclassified.

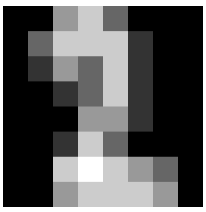
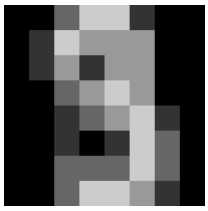


Figure 4: Generated image of a 2



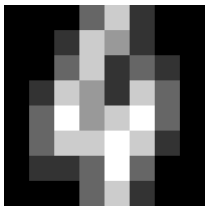
(a) 400x400px

(b) 32x32px

(c) 16x16px

(d) 8x8px

Figure 5: Generated image of a 3



(a) 400x400px

(b) 16x16px

Figure 6: Generated image of a 4

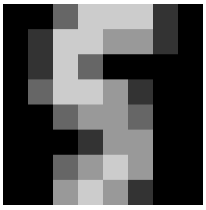


Figure 7: Generated image of a 5

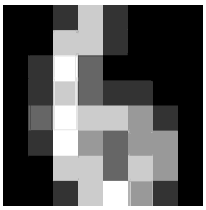
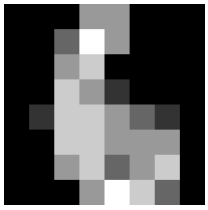


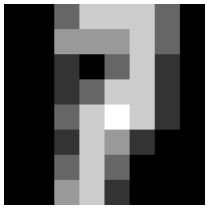
Figure 8: Generated image of a 6



(a) 400x400px

(b) 16x16px

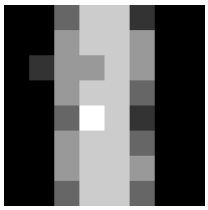
Figure 9: Generated image of a 6



(a) 400x400px

(b) 32x32px

Figure 10: Generated image of a 7



(a) 400x400px

(b) 16x16px

Figure 11: Generated image of an 8

Digit 9

Remember that the 1 and the 9 were combined be the model due to their similar features.