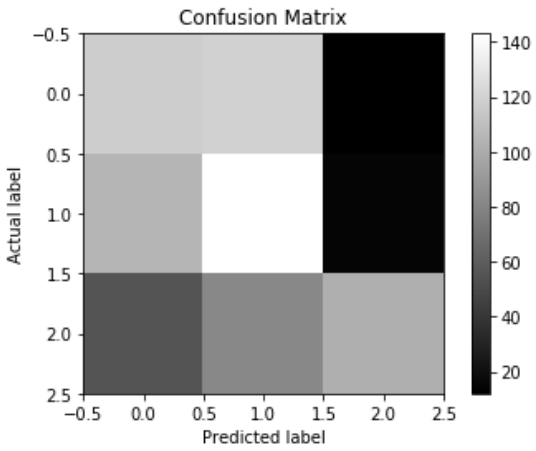
Peter Victoratos
CSC 441 Computer Vision
April 3, 2018

Lab 14 KNN

Question A: KNN is faster to train, which is not preferable. Training follows the same steps every time. Testing has to evaluate each image and compare it to all of the other ones which can take a longer time. Faster testing would be better to have since the results of the test will be what you will adjust based on the results.

```
In [11]: runfile('/Users/pvictoratos/Desktop/
lab14_starter.py', wdir='/Users/pvictoratos/Desktop')
Reloaded modules: datasetloader
Loading data...
[INFO] processed 500/3000
[INFO] processed 1000/3000
[INFO] processed 1500/3000
[INFO] processed 2000/3000
[INFO] processed 2500/3000
[INFO] processed 3000/3000
Seconds elapsed: 0.95
Creating KNN classifier...
Seconds elapsed: 0.19
Evaluating KNN classifier...
Seconds elapsed: 7.66
accuracy = 48.27%
[[117 119  12]
 [105 143  15]
 [ 55  82 102]]
```

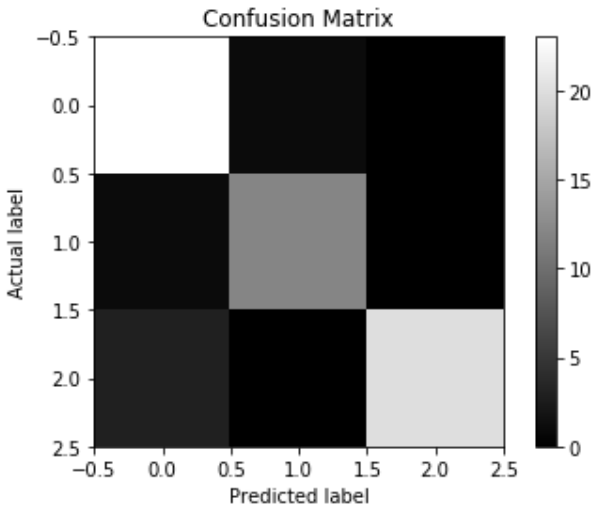


Question B: The overall classification accuracy for the test shown above was 48.27%. It's pretty good at recognizing the first object in question, okay with the second, and not so great with the third.

Question C: With the confusion matrix above, we can infer some mistakes that the classifier is making. Sometimes pandas and dogs can look similar if the dog has a similar color/pattern of spots on it. There can also be issues regarding the backgrounds of the images, their quality, orientation (and a whole other list of object characteristics). The model most likely doesn't mistaken a cat for a panda since generally cats have skinners bodies and legs, while pandas have more *robust* features.

6. Flower Dataset Results

```
In [12]: runfile('/Users/pvictoratos/Desktop/
lab14_starter.py', wdir='/Users/pvictoratos/Desktop')
Reloaded modules: datasetloader
Loading data...
Seconds elapsed: 1.32
Creating KNN classifier...
Seconds elapsed: 0.01
Evaluating KNN classifier...
Seconds elapsed: 0.04
accuracy = 91.67%
[[23  1  0]
 [ 1 12  0]
 [ 3  0 20]]
```



The flowers dataset took a much faster time to train, and an even faster time to test. This is because there was less data to look at. The results were a lot better (an accuracy rate of 91.67%), the distinctions between the images seem to have been more definite, which may have also contributed to the speed of the program.

7. Different k-values on the flowers data set

```
In [14]: runfile('/Users/pvictoratos/Desktop/
lab14_starter.py', wdir='/Users/pvictoratos/Desktop')
Reloaded modules: datasetloader
Loading data...
Seconds elapsed: 1.29
Creating KNN classifier...
k value: 5
Seconds elapsed: 0.01
Evaluating KNN classifier...
Seconds elapsed: 0.04
accuracy = 83.33%
[[22  2  0]
 [ 6 13  0]
 [ 2  0 15]]
```

```
In [15]: runfile('/Users/pvictoratos/Desktop/
lab14_starter.py', wdir='/Users/pvictoratos/Desktop')
Reloaded modules: datasetloader
Loading data...
Seconds elapsed: 1.29
Creating KNN classifier...
k value: 13
Seconds elapsed: 0.01
Evaluating KNN classifier...
Seconds elapsed: 0.04
accuracy = 85.00%
[[20  0  0]
 [ 7 11  0]
 [ 2  0 20]]
```

```
In [16]: runfile('/Users/pvictoratos/Desktop/
lab14_starter.py', wdir='/Users/pvictoratos/Desktop')
Reloaded modules: datasetloader
Loading data...
Seconds elapsed: 1.29
Creating KNN classifier...
k value: 33
Seconds elapsed: 0.01
Evaluating KNN classifier...
Seconds elapsed: 0.05
accuracy = 85.00%
[[17  0  0]
 [ 5 14  0]
 [ 4  0 20]]
```

It seems that adding more neighbors brings the accuracy down but levels out around 85%. There seems to be no difference in running time