Assignment 1-Jianshu Wang

Note: I am using python for all my assignment and all the library I used is from sklearn

Problem 1:

Part A:

Answer: 63.64%, I write this classifier by myself in python

Part B:

Answer: 63.64%, this does not change a lot even though I did not counter those 0 values in column 3,4,6,8

Part C:

Answer: 74.92%

I used 10 fold cross-validation and get a 10 values in training set which with 80% data, 74.92% is the mean value

Part D:

Answer: 64.29%

Problem 2:

Part A:

Untouched image: 71.25% for Gaussian 83.51% for Bernoulli

Stretched bounding: 83.33% for Gaussian 82.98% for Bernoulli

From the result we can see that Bernoulli is much better than that of using Gaussian distribution for untouched image, and results are pretty close between Gaussian and Bernoulli, so I think they both can be a good choice for stretched bounding box, as they both output good high accuracy

Part B

Untouched pixels:

	Depth = 4	Depth = 8	Depth = 16
#trees = 10	70.36%	87.93%	94.19%
#trees = 20	72.85%	89.16%	95.8%
#trees = 30	73.81%	89.85%	95.84%

Stretched bounding tree

	Depth = 4	Depth = 8	Depth = 16
#trees = 10	73.81%	89.38%	95.07%
#trees = 20	75.04%	90.83%	96.02%
#trees = 30	76.27%	91.46%	96.16%