**Problem 1**

a. Accuracy: 0.7712418

b. Accuracy: 0.6862745

c. Accuracy: 0.7059

d. Accuracy: 0.7712418

**Problem 2**

a.

|  |  |  |
| --- | --- | --- |
| **Accuracy** | **Gaussian** | **Bernoulli** |
| **Untouched Image** | 0.54 | 0.79 |
| **Stretched Bounding Box** |  |  |

Bernoulli is a better choice because having that threshold helper function to convert numeric data into categorical data allows there to be a more distinct difference for features for different labels. Additionally, if I were able to complete the stretched bounding box it would show that the overall accuracy would be higher for the stretched bounding box version of Gaussian and Bernoulli and ultimately Bernoulli for stretched bounding box would have the best accuracy.

b.

**Untouched Raw Pixels**

|  |  |  |  |
| --- | --- | --- | --- |
| **Accuracy** | **depth = 4** | **depth = 8** | **depth = 16** |
| **#trees = 10** | 0.8776 | 0.8761 | 0.8762 |
| **#trees = 20** | 0.9263 | 0.9283 | 0.9266 |
| **#trees = 30** | 0.9402 | 0.9432 | 0.944 |

**Stretched Bounding Box**

|  |  |  |  |
| --- | --- | --- | --- |
| **Accuracy** | **depth = 4** | **depth = 8** | **depth = 16** |
| **#trees = 10** |  |  |  |
| **#trees = 20** |  |  |  |
| **#trees = 30** |  |  |  |

Although I could not get the stretched bounding box to work because I had issues resizing the image. Conceptually, the accuracy for these tests would be higher for the stretched bounding box because you are cleaning the image more with preprocessing. Additionally, the trend would hold that the more #trees you have the higher accuracy you would get. If I were to fill out the data for the stretched bounding box, the accuracy values would be higher as I increase #trees and in general be a little higher than the untouched raw pixels

**Citations:**

1. Professor Forsyth’s code snippets

2a. <http://api.rpubs.com/JanpuHou/304506>

Used citation in order to help build the Gaussian model and then from there branched out to try building the Bernoulli

2b. <https://www.kaggle.com/iammangod96/mnist-digit-recognition-using-simple-randomforest/code>

Used citation in order to build the random forest classifier