Student ID: 1224239328

Values :

Mean and variance of f1 for digit0 =

Train0

44.19694081632653

114.78696762793733

Test1

43.939138640149935

123.18002137112168

Mean and variance of f2 for digit0 =

Train0

87.43991710811058

101.09764281367967

Test0

87.18577738773497

106.46285124489977

Mean and variance of f1 for digit1 =

Train1

19.56806279780635

34.88200678861763

Test1

19.282945153061224

30.350068252419238

Mean and variance of f2 for digit1 =

Train1

61.229937437452705  
80.11070228540537

Test1

61.671272363141846

90.25124644699872

Predicting new labels for digit0testset and calculating the accuracy =0.24591836734693878

Predicting new labels for digit1testset and calculating the accuracy = 0.6511013215859031

The four sets of data were given. TrainSet0 and TrainSet1 were given to “train” the Test set data. The data Brightness average was calculated, along with the standard deviation. These values were then utilized to calculate the mean, and variance for both data sets with respect to their classes 0 and 1 as shown in their names, e.g “train0, train1”. The average brightness (mean of train0 & train1), and standard deviation of average brightness (standard deviation was utilized of train0 & train1) to predict the class’s of the tests : train0 and train1. This utilized the Naïve Bayers classifier to predict our classes for our test set data of “ymax’.

The process chooses the highest y of test 0 and 1 to assign the correlating indexed values of test0 and test1 data sets respectively to the appropriate class. This is calculated through a product of our prior probability and a total product of our features probability for each trainset. We then divide the number of “successes” for each Classifier 0 and 1 by the total number of values in each test set to get our Accuracy. This concludes that our Test set 1 is more accurate than our Test Set 0.