Doruk Özer

70192

Homework 2

I read Section 5.7 from the textbook. After reading, I realized that we will be using multivariate parametric approach that we have used in our first lab class. I estimated the mean and deviation vectors for each class and each column is calculated separately. I used first 30000 images for learning. For example I estimated the mean vectors of the vectors that belongs to first class, and did it for other classes. I repeated it for the deviations. Then I calculated the frequencies of each class. As a result, I got the same results in the pdf.

Later I assumed every column (I assumed pixels of the images as column,784 column)of the image vector has a Gaussian Distribution and has Normal Distribition. Every class had a sample mean, and sample deviation to each corresponding column for those normal distributions. So I wrote an algorithm, gi(sample\_means,class\_priors,sample\_deviations,x) ,that uses the same approach in the lab 1. It finds a score value for each column of vector for each class and stores those scores in a 5x784 sized array.Then I wrote find\_score(a) algorithm that takes the vector returned from gi as an input. It sums every column of rows and stores the summations in a vector sized 5. Then it returns the index of the biggest number in that array.

I created score\_predicted vectors by using those functions. I gave every vector as an input to gi(sample\_means,class\_priors,sample\_deviations,x), and gave every result array to the find\_score(a) and stored the indexes in a vector.

I calculated the confusion matrix from the score\_predicted arrays and used them to generate the confusion matrix

Calendar

Description automatically generatedTable

Description automatically generatedResults for training set Results for test set