Nitheese Thirumoorthy

ASU ID:1225417336

**PROJECT PART 1**

**Feature Extraction, Density Estimation and Bayesian Classification.**

**1)Steps followed in the Project:**

1. First, we import the data from the .mat files.
2. We then extract the data and store them in Xtrain,ytrain,Xtest,ytest respectively.
3. Then we create the following functions.

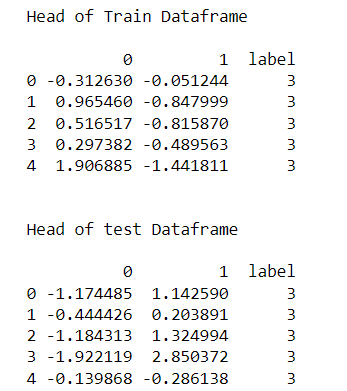
“skewness”- to calculate the skewness of every image in the given data.

”brighttodark”- to calculate the ratio of bright pixels to dark pixels of every image in the given data using a threshold T.

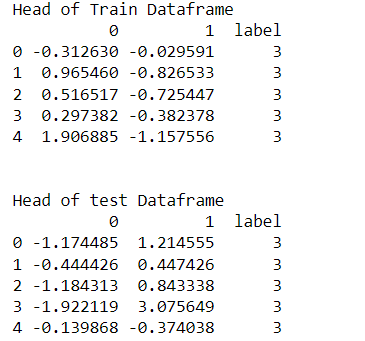
“normalize”- which uses these functions to normalize our data.

1. We then, convert these normalized data into a dataframe, which makes it easier for us to do the remaining tasks.
2. We then use Maximum Likelihood estimation, to find the estimates of mean and co-variance of both the classes.
3. We then calculate the probability density function using these estimates.
4. Then we define a function which does minimum error rate classification given the prior probabilities and X.
5. Then we calculate the error rate for every configuration.

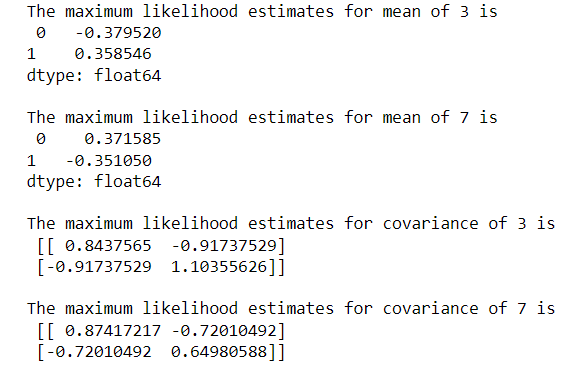
**2)Head of the Normalized Features when the Value of T is 150:**



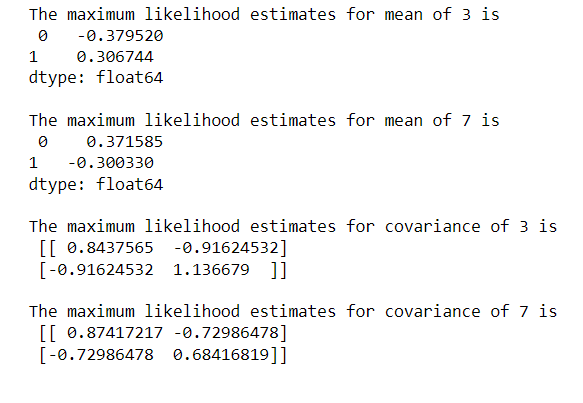
**3)Head of the Normalized Features when the Value of T is 200:**



**4)Mean and Co-variance when T is 150:**



**5)Mean and Co-variance when T is 200:**



**6)Error Rates for all possible values of T and all possible prior probabilities:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Threshold (T)** | **P(3) and p(7)** | **Train Error Rate** | **Test Error Rate** |
| 150 | 0.5 and 0.5 | 0.3428299272601316 | 0.33853083853083854 |
| 150 | 0.3 and 0.7 | 0.40656390717007274 | 0.4078309078309078 |
| 200 | 0.5 and 0.5 | 0.28827502597852445 | 0.2955647955647956 |
| 200 | 0.3 and 0.7 | 0.40318669899549703 | 0.41302841302841303 |