21100118

REPORT – PART 2

Approach

First, the training data is separated based on 'class' value which is essentially splitting the data depending on the patient being normal or abnormal. This is done so that calculating prior probabilities becomes easy and can be treated like two different lists even though they are in the same 'set.' After doing this, the prior probabilities are calculated. These are calculated using the pre-defined formulas. These prior probabilities are then used when predicting if the person is normal or abnormal given the test data. Probabilities are calculated for c0 (Abnormal) and c1 (Normal). Whichever value is greater, that is our 'prediction'. If c0 is greater than c1, the person is abnormal and if c1 is greater than c0 then the person is normal. After determining our prediction, the actual result (the first element of the data for each person) which I had separated already was used to compare with our prediction. After this, the accuracy percentage was calculated.

The accuracy I got was 77.54%

Output

```
harumnaseem@harumnaseem Part2 % python3 classifier.py Spect_train.txt Spect_test.txt
##########

Starting to Train on 80 data points . . .

Training Complete


Testing on 187 data point . . .

Total Accuracy: =  77.54010695187165 %

##########
```