Expectations :

1. The code is expected to be written in a *modular*/*plug and play* manner
2. Each function is expected to be fully parameterised
3. Students are expected to observe impact of parameter value changes on the accuracy and note down the observations/patterns

Assignments

1. An univariate classifier from first principles
    a. Generate distributions (gaussian to start with) for male and female heights (1000 samples each). Fix the mean of female heights to 152 cm and male mean height to 166 cm. label the appropriate gender for samples in each of the distribution (M or F)
    b. Fix the sd of both the distributions to 5
    c. Try classification of gender using following approaches with aim to minimise misclassification
        i. Assign gender based on likelihood calculated from distributions (empirically estimated mean and sd and calculate probability assuming gaussian distributions)
        ii. Derive a threshold hight to separate male female
        iii. quantize the data at scale of 0.5 cm and empirically estimate the likelihood of male female in each segment based on majority
        iv. In each of the above cases output a confusion matrix for classification
    d. Try following values of sd (eg 2.5, 7.5 and 10) repeat 3.a, 3.b, 3.c, 3,d observe impact of change in sd on classification accuracy
    e. Change the quantization interval length (say 0.001, 0.05, 0.1, 0.3, 1, 2, 5,10 cm etc) repeat 3.a, 3.b, 3.c, 3,d observe impact of change in sd on classification accuracy