GMM:

Work on one of the following datasets.

a) Speaker Identification Dataset

b) Image Dataset (from the previous assignment, use all the classes)

HMM:(DHMM)

a) Work on video dataset (optional)
b) Work on digit data (one HMM for one digit)(from assignment 2)

HMM: (HTK)

Work on **both** of the following datasets.

a) TIDIGITS Dataset (Perform Embedded Reestimation)

b) Handwritten Dataset (Take from the previous assignment)

DTW:

Work on one of the following datasets.

a) Music Dataset

b) Mandi Dataset

Non Parametric Methods:

Build the following classifiers for the Image dataset alone (Use all the classes):

1)Bayes classifier using Parzen window method for class-conditional density estimation
(h to be emprically chosen)

```
        * Hypersphere as region


        * Gaussian kernel


   2)Bayes classifier using K-nearest neighbour method for
class-conditional density estimation
       (K to be empirically chosen)




   3)Fisher Discriminant Analysis (FDA) based classifier.


   4)Perceptron based classifier.


   5)Support Vector machine (SVM) with linear kernel.
(Please go through this link
       to know more about libsvm library download the
library according to your system
       configurations. There are other tools available for
SVM,
       its not mandatory that you use libsvm, but it is
convenient. )
```

Dataset Formats:

1. **TIDIGITS:**
   - The dataset contains the utterances of multiple digits. The set of digits uttered are given below: symbol - uttered word 1 - one 2 - two 3 - three 4 - four 5 - five 6 - six 7 - seven 8 - eight 9 - nine **z - zero o - o**
   - The MFCC features, extracted from the wav files, are what is given.
   - The MFCC filename specifies the sequence of symbols that are spoken in the corresponding wav file.
   - Ignore the characters 'a' and 'b' in the filenames.
   - Eg. In file 534z29a.mfcc, the digits spoken are five three four zero two nine **Structure of MFCC file:** The first line of MFCC file contains two space separated integers. First integer N_C - The dimension of the feature vector (The number of MFC coefficients) Second integer N_F - The number of frames , the wav file is divided into. Starting from the second line, there will be N_F lines. Each line contains N_C numbers corresponding to the MFCC feature vector for a particular frame.
2. **ONLINE HANDWRITTEN DATASET:**
   - The data given are the sampled coordinates of telugu letters, as they are written.
   - A stroke is a coordinate sequence traced without lifting the pen.
   - A letter can be written with different number of strokes.

- o   The ldf file of a particular letter consists of multiple examples for that letter.
- o   First line of each example is N_S - number of strokes.
- o   Followed by 2 * N_S lines, the stroke name, and the stroke coordinate sequence for each stroke.
- o   The stroke coordinate sequence is given in a single line, in the following format $<N_C> <x_1> <y_1> <x_2> <y_2> ... <x_{NC}> <y_{NC}>$ where N_C - Num of coordinates

3. **MANDI SPEECH DATASET:**
   - o   The data given are the MFCC features of speech data.
   - o   The structure of MFCC file is explained above.
   - o   The name of the mfcc file is what is spoken in the corresponding wav file

4. **Video Dataset**

   You have to process the features from the pixel values of video. Details of video can be obtained from this website. The three classes of video (running, walking and clapping) can be downloaded from this link.