CS5691: Pattern Recognition and Machine Learning
Programming Assignment 3

Deadline: April 07, 11:59 pm

A) K-means, GMM

The task is to perform classification by modelling the class-conditional density as a Gaussian Mixture Model for the following datasets:
a) Image dataset
b) Synthetic dataset

Image Dataset: Image data corresponding to five different classes — Coast, Forest, Mountain, Highway, and Open country.

You are required to work on the features extracted (download from here). Each image is divided into 6x6 blocks and a 23-dimensional feature vector is extracted from each block. The 23-dimensional features include colour histogram, edge-directed histograms, and Entropy of wavelet coefficients extracted from local blocks of an image for a particular scene. Original images can be accessed from here.

Synthetic Dataset: The synthetic dataset can be found here. The folder name corresponds to the team number. Please use only the data allotted to your team.

Format: Each row contains
Dim1, Dim2, Class

Guidelines:
0. Sample plot. Add the contours of the mixture components on the plot as shown in the sample plot.
1. Try with different values of 'K' (the number of mixtures in each class' model). You can use the same or a different number of mixtures for each class. Your 'K' can be as low as 2 and as high as 30 though you should get an idea of how many mixtures to choose on observing the data.
2. Try with both Diagonal and Non-diagonal covariances and report results for different values of K.
3. Plot ROC and DET for both datasets. Each of your ROC and DET plots should compare the models with diagonal and non-diagonal covariances.
5. Try normalising your dataset. Experiment with the number of K-means iterations and EM iterations.

6. Please write generalised codes for GMM, K-means, ROC and other parts as we will be asked to revisit them in future assignments.

Libraries not allowed: GMM, K-means, mean, covariance estimation, Gaussian likelihood estimation, ROC.

Bonus question (Optional): Use the UBM-GMM concept covered in the class and perform the aforementioned experiments.

Page limit: 4 pages

B) DTW and Discrete HMM

You need to do the following task:
1. DTW
2. Discrete HMM (using this code)

Isolated Spoken-Digit dataset: This dataset consists of spoken utterances. Sounds in speech are generated by the resonances of the vocal tract. MFCC (Mel Frequency Cepstral Coefficients) represent the resonances of the vocal tract. These features have already been extracted. Any sound in speech is represented by a sequence of spectral feature vectors. Along with the MFCC files, the original .wav files are also given. The dataset and the team mapping can be found here.

Structure of MFCC file: The first line of the MFCC file contains two space-separated integers. First integer NC — the dimension of the feature vector (the number of MFCC coefficients), second integer NF — the number of frames the .wav file is divided into. The next NF rows contain the MFCC features of dimension NC. Each row corresponds to a feature vector in the sequence. Please note that NF varies with each example.

Online Handwritten-Character dataset: Handwritten Telugu characters. The dataset and the team mapping can be found here.

Structure of the data: The data given are the sampled coordinates of Telugu letters, as they are written. Each file consists of one handwritten character.

The handwritten character sequence is given in a single line, in the following format <NC> <x1> <y1> <x2> <y2> ... <xNC> <yNC> where NC — number of coordinates.

Apart from using the (x,y) positions as given in the data, you can extract other features such as slope and curvature to perform the handwritten letter recognition task.

Guidelines:
1. You need to plot ROC, DET and confusion matrices (your own).
2. You can include graphs and tables for your results.

Bonus question (Optional): Use the discrete HMMs trained to recognise connected digits and handwritten characters. The datasets can be found here (link to be added). Write code to concatenate the HMMs trained in the previous part. Using the concatenated HMMs, recognise the continuous digits/characters.

Connected digits:

Download the development and blind test data from here. The data contains directories with team numbers. Each directory contains MFCC features from utterances of multiple digits (corresponding to the isolated digits assigned to your batch). The data format is the same as isolated digits in the previous part. The set of digits uttered is as follows:

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

In development data, the file name represents spoken digits. For example, in file 534.mfcc, the digits spoken are five three four. The blind test dataset consists of 5 unlabeled sequences. Provide the possible sequence of digits obtained in the report.

Connected Handwritten Characters:

Download the development and blind test data from here. The data contains directories with group numbers. Each directory contains sample x and y locations from the continuous characters. The data format is the same as that for the handwritten characters' data in the previous part. In development data, the file name represents characters. For example, in file bA_chA_dA_1.txt, the characters bA, chA, and dA are written in sequence. The blind test dataset consists of 5 unlabeled sequences. Provide the possible sequence of digits obtained in the report.

Page limit: 6 pages

Please note that no deadline extension will be given for this assignment.