EE 624 Speech Technology
Report for Project 2
Isolated Digit Recognition using GMM

By: Kota Yamini Suguna
Roll number: 194102309
M.Tech Signal Processing and Machine Learning

Objective:

Recognition of isolated digits using GMM. Model the GMMs using 25 utterances of each digit from 12 different people and test it using 25 utterances of each digit in own voice.

Parameters used:

- Sampling Frequency: 16 kHz
- Top dB for trimming training audios: 20dB
- Top dB for trimming training audios: 36dB
- Features used: MFCC, Delta and Delta-Delta components
- Window length used for MFCC: 25ms
- Window Shift: 10ms
- Densities used to model GMMs for individual digits: 16

Confusion Matrix obtained:

conf - NumPy array

	0	1	2	3	4	5	6	7	8	9
0	19	0	2	0	2	2	0	0	0	0
1	0	22	0	0	0	2	0	0	0	1
2	19	2	1	0	2	1	0	0	0	0
3	0	0	1	0	0	0	2	0	22	0
4	3	2	0	0	17	3	0	0	0	0
5	0	0	0	0	0	19	0	3	3	0
6	0	0	0	4	0	3	3	14	1	0
7	0	0	2	0	0	1	0	16	6	0
8	0	0	1	0	0	0	2	0	21	1
9	0	1	0	7	1	12	0	0	1	3

Conclusion:

Using GMM, the accuracy for isolated digit recognition is reasonably good for 6 digits.
i.e., ‘0’, ‘1’, ‘4’, ‘5’, ‘7’ and ‘8’.