Total Marks: 20 **Due Date:** Feb 11, 2024

Speech Understanding

Minor 1

- Any kind of plagiarism is not accepted. We will strictly follow institute policies for plagiarism.
- Recommended programming languages: Python
- Please submit a single zip file containing the report, codes, and readme if required. The zip file should be named Rollno_Minor1.zip.

---

1. Design your custom spectrogram function to accept parameters such as n_fft (length of FFT window), hop_length(number of samples between successive frames), window(a window specification, e.g., Hann, Hamming, gaussian), and win_length(length of the window). Do not use the library functions to create the spectrogram. You can use functions of basic mathematical operations such as Fourier transform. *(5 marks)*
    a. The output should be a spectrogram of your audio speaking – "Hello, everyone, my name is XYZ (your name), and here is the spectrogram for the minor 1 exam." *(1 mark)*
    b. Create the spectrogram on the same sentence spoken by any TTS model of your choice. *(1 mark)*
    c. Present both spectrograms (a and b) side by side, demonstrating the impact of parameter choices on the visualization. Provide a detailed analysis of how variations (atleast four different variations) in the window, n_fft, overlap, and win_length influence the spectrogram's characteristics. Discuss the implications of these parameter choices in the context of speech processing. *(3 marks)*

2. Use the concepts studied in the class so far to build a classification model for the Acoustic Scene Classification . You can only use non-deep learning-based ML models; deep learning models are not allowed.
    a. Explain the choice of the selected ML model and the corresponding hyperparameters. *(2 marks*)

b. Follow the protocol and report the results in terms of the evaluation metrics (class-wise accuracy, precision, and AUC). *(5 marks)*

c. Change the n_fft parameter in the spectrogram and analyze its implication on the evaluation metrics (class-wise accuracy, precision, and AUC) *(3 marks)*