

MCA Assignment 2

Vaibhav Goel

2016111

The code is commented to run evaluation on the validation set.

The background noise was added randomly to some of the samples and the model was then separately tested

Spectrogram: Short time fourier transform is used to create the spectrogram. The given sample is divided into windows of size 256 frames at an interval of 128 frames. Then FFT is calculated for each window.

These features are then stored in a pickle file and then testing is performed using the validation set.

The size of each window is taken to calculate stft is 256.

The overlap is taken as 128.

Without Noise:

1. Precision:
 - a. Macro average: 0.35
 - b. Weighted average: 0.36
2. Recall:
 - a. Macro average: 0.35
 - b. Weighted average: 0.35
3. F1:
 - a. Macro average: 0.34
 - b. Weighted average: 0.35

With Noise:

1. Precision:
 - a. Macro average: 0.36
 - b. Weighted average: 0.36
2. Recall:
 - a. Macro average: 0.35
 - b. Weighted average: 0.35
3. F1:
 - a. Macro average: 0.35
 - b. Weighted average: 0.35

MFCC: we use the same spectrogram code to get the power signal for the given audio sample. Next we create filter banks, by taking equally spaced intervals between two frequencies and using the filter bank scales, we create the mel filters. After applying the filter on our power signal, we calculate DCT which yields us the MFCC.

Without Noise:

1. Precision:
 - a. Macro average: 0.34
 - b. Weighted average: 0.34
2. Recall:
 - a. Macro average: 0.27
 - b. Weighted average: 0.28
3. F1:
 - a. Macro average: 0.26
 - b. Weighted average: 0.26

With Noise:

1. Precision:
 - a. Macro average: 0.31
 - b. Weighted average: 0.31
2. Recall:
 - a. Macro average: 0.27
 - b. Weighted average: 0.28
3. F1:
 - a. Macro average: 0.24
 - b. Weighted average: 0.25