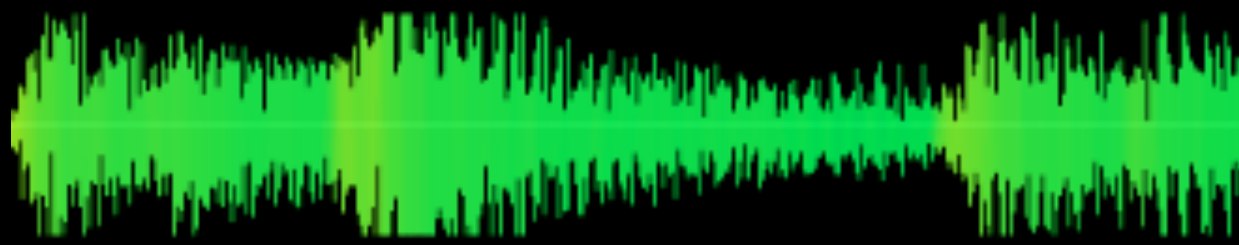


# Freesound General-Purpose Audio Tagging Challenge

<https://www.kaggle.com/c/freesound-audio-tagging>



## Data Set :

The train data set includes ~9.5k samples unequally distributed among 41 categories. Each file is an audio file(.wav) and there are labels so that each file matches a label.

The test data set includes ~9.4k samples and they have no label. We need to upload our answer to Kaggle to know our model's accuracy.

## Problem :

Labeling the test data using the model learned by train data.

## Sound Classification With Deep Learning

### Data Processing 1 : MFCC

MFCC is the most generally used method for the sound processing. So we converted the data into arrays using MFCC (librosa module of python).

### Data Processing 2 : STFT

While MFCC is most commonly used in sound processing, passing a Mel filter filters out the frequency range that corresponds to the human voice, resulting in the loss of necessary data. So we also did it with STFT.

### Deep Learning Model : CNN

We trained the model using three convolutional layers and two fully connected layers.

41 Categories including  
"Acoustic\_guitar", "Applause",  
"Bark", "Bass\_drum"

### Result (MFCC) : 75.4%

The Accuracy when we converted the data using MFCC.

### Result (STFT) : 63.7%

The Accuracy when we converted the data using STFT.

### Result (MFCC X STFT) : 82.3%

The result of multiplying the predicted probability using MFCC and the predicted probability using STFT. (The accuracy is much better than in both cases)

We only submitted the most likely answer when submitting the answer. Submitting the three answers as required by the original problem will be more accurate and is currently under revision.

**Kim SeungHyuck**

GitHub : <https://github.com/smallerhand>

Project : <https://github.com/kylusmethod/KYLius-method/tree/master/PROJECT2>

