# Group 12: Audio classification using ensemble methods

Abhinav Dudeja (21111001)
abhinavd21@iitk.ac.in

Arpan Kapoor (21111016)
arpank21@iitk.ac.in

Piyush Gangle (21111046)
piyushg21@iitk.ac.in

February 20, 2022

## 1 Problem Statement

Environmental sound classification is an interesting problem which has different applications ranging from crime detection to environmental context aware processing. Moreover, with the increasing interest in smart cities, IOT devices embedding automatic audio classification can be very useful for urban acoustic monitoring like intelligent audio-based surveillance system in public transportation.

## 2 Solution

We propose to use the following ensemble of DL audio classification techniques in order to try to improve the accuracy:

- 1D CNN that takes the original audio data as input
- 2D CNN or RCNN that takes the spectrogram image of the audio data as input

[1] describes the 1D CNN approach and finds that it performs good on most classes, but there are some classes on which [2] performs better. Hence in order to reach at a state where the model provides a fair justice to all the classes present in dataset, we are trying to come up with a new solution to the problem.

## 3 Dataset

We plan to use the UrbanSound8K dataset available at `https://urbansounddataset.weebly.com/urbansound8k.html`

# References

[1] S. Abdoli, P. Cardinal, and A. Lameiras Koerich, "End-to-end environmental sound classification using a 1d convolutional neural network," *Expert Systems with Applications*, vol. 136, pp. 252–263, 2019.

[2] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 279–283, 2017.