

---

# Urban Sound Classification

---

**Jay Jagtap**      **Rohan Pillai**      **Sri Harsha Varma**      **Sai Shashank N**  
jjjagtap@ncsu.edu    rspillai@ncsu.edu    suppala@ncsu.edu    snayabu@ncsu.edu

## 1 Dataset

The dataset for this project is publicly available on <https://urbansounddataset.weebly.com/urbansound8k.html>. It contains 8732 labeled sound excerpts ( $\leq 4$ s) of urban sounds from 10 different classes.

## 2 Project Idea

The aim of this project is to implement a Machine/ Deep learning model that is capable of classifying sounds in an urban environment. The dataset has an unstructured data of sounds in .wav format and is divided into 10 parts. We will perform 10 fold cross validation to train our classification system and report the average accuracy.

Our models will be based on extracted MFCC data and/or the spectrogram image of the audio. In [1], the average accuracy for SVM and Random-Forest classifier is around 68 percent and is much lower for the other algorithms, we aim at achieving an accuracy greater than 68%. Our data set size is 6 GB and it might not be possible to train it on Google Colab/personal computers, in such a scenario, we will trim our dataset to use around 1500 samples and will benchmark against 65% accuracy.

## 3 Software

We will develop our models in Python 3.6 and following libraries will be required for our implementation: Scikit-learn (for classification, and other Machine Learning algorithms) , Librosa (for music and audio analysis), Numpy (for mathematical operations), Pandas (for Data Analysis), Tensorflow (for deploying ML models), Keras (wrapper for Tensorflow).

## 4 Relevant Papers

[1] Justin Salamon, Christopher Jacoby, and Juan Pablo Bello. (2014). *A Dataset and Taxonomy for Urban Sound Research*.

[2] Cai, Rui Lu, Lie Hanjalic, Alan Cai, Lian-Hong. (2006). flexible framework for key audio effects detection and auditory context inference.

## 5 Teammates

Sri Harsha Varma (suppala) will be responsible for preprocessing, Sai Shashank Nayabu (snayabu) will work on extracting features, Rohan Pillai (rspillai) will work on training Machine Learning models, and Jay Jagtap (jjjagtap) will train Deep Learning architectures. The work distribution is also collaborative in nature and teammates will help each other in their tasks.

## 6 Midterm Milestone

We are planning to finish exploratory data analysis and the data preprocessing steps in 2-3 weeks. By midterm, we have planned to develop a baseline model for the classification with atleast 50% accuracy.