**Audio Dataset**

**Description**

The audio dataset contains 8732 labelled sound files which are less that or equal to 4 seconds in length. The audio wav files contain urban sounds from 10 classes: air conditioner, car horn, children playing, dog barking, drilling, idle engine, gun shot, jack hammer, siren and street music.

All recordings are taken from field recordings and the files are pre-sorted into ten folders. In addition to the sound excerpts, a csv file containing meta data about each excerpt is also provided.

**Meta Data Files**

A csv file is provided that contains the meta-data information about every audio file in the dataset. The headers in this csv file are:

* Slice\_file\_name: The name of the audio file – the filename is formatted as: [fsID]-[ClassID]-[OccurenceID]-[slideID].wav, where:
  + [fsID] = the sound ID of the recording from which this excerpt (slice) is taken.
  + [classID] = a numeric identifier of the sound class.
  + [occuranceID] = a numeric identifier to distinguish different occurrences of the sound within the original recording.
  + [sliceID] = a numeric identifier to distinguish different slices taken from the same occurrence.
* fsID: The sound ID of the recording.
* start: The start time of the slice.
* end: The end time of the slice.
* salience: a subjective salience rating of the sound. 1=foreground, 2=background.
* fold: The fold number (1-10) to which this file has been allocated.
* classID: a numeric identifier of the sound class:
  + 0=air\_conditioner
  + 1=car\_horn
  + 2=children\_playing
  + 3=dog\_bark
  + 4=drilling
  + 5=engine\_idling
  + 6=gun\_shot
  + 7=jackhammer
  + 8=siren
  + 9=street\_music
* Class: The class name: air\_conditioner, car\_horn, children\_playing, dog\_bark, drilling, engine\_idling, gun\_shot, jackhammer, siren, street\_music.

**Class Distributions in Each Fold**

Many audio classification experiments are evaluated using 10-fold cross validation – hence why you have 10 folds. It is strongly recommended that you use this method. Do not just evaluate one split – use 10-fold (not 5-fold) cross validation and average the scores (statistically you should use 30 but it is not necessary for this coursework). The figure below shows you the class distribution for each fold. As you can see there is class imbalance so you may need to consider this in your modelling.

A picture containing chart

Description automatically generated