

Urban Sound Classification Project Proposal

1. Problem to be solved:

The question I aim to investigate here is given a brief sound excerpt from an urban environment, what is the source of the sound? The goal is to develop a simple classifier system that can output the sound label after learning an appropriate data classification model from a set of training samples that have been manually annotated. System performance will be measured using a classification accuracy score using an independently labeled test set of sound samples. Deliverables for this project will be a text report detailing the methods and data analysis techniques used, the code and tools used to analyze, model and test the data. This will be in addition to a slide deck to be used for brief project presentations in the future.

2. Clients interested in the solution:

The developed solution is of interest to audio processing system software devices. Classifying ubiquitous sounds according to their sources can be a preprocessing step for many audio processing systems before more sophisticated data analyses. For example, speech recognition devices need to remove non-speech sounds before language translation. As a result, identifying these non-speech patterns can help improve the recognition capabilities of many current speech recognition software.

3. Dataset:

The dataset consists of 8732 sound excerpts that are shorter than 4 seconds each. Sound excerpts can be categorized into ten classes that represent common sounds often experienced in an urban environment. The labeled classes available in the dataset are grouped into the following sounds: Air Conditioner; Car Horn; Children Playing; Dog Bark; Drilling; Engine Idling; Gun Shot; Jackhammer; Siren; Street Music.

The dataset is available for download at:

<https://drive.google.com/drive/folders/0By0bAi7hOBAFUHVXd1JCN3MwTEU>

The training sample dataset is 3.4GB and the test sample dataset is 2.15GB. Data files consist of .wav audio file formats where each file has a unique numeric ID. The class labels are in a separate excel file in the same folder that has two columns containing the file ID and its class annotation. The training data is completely labeled while the test data is not. This could be problematic in the accuracy score calculations part of the project since there is no ground truth. I can overcome this in two ways either by adding an extra labeling step to fill in the missing labels or using part of the training data files in the testing stage. It is to be noted, the same file cannot be used for training and testing simultaneously.

4. Methods:

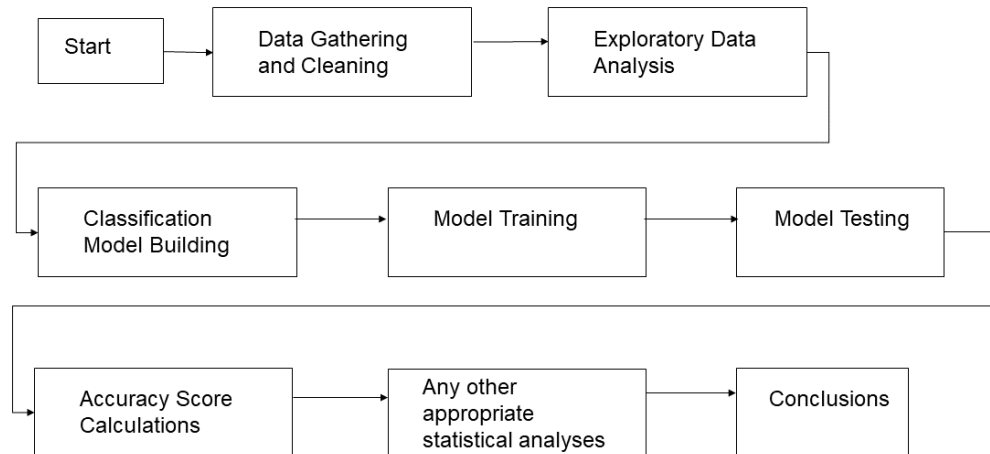


Figure 1: Project Steps and Milestones

A rough outline of the project stages is shown in figure 1. I will use Python 3 packages to do all the data analysis, modeling and visualization needed to draw the appropriate conclusions.

5. References:

Salamon, Justin & Jacoby, Christopher & Bello, Juan. (2014). *A Dataset and Taxonomy for Urban Sound Research*. Proceedings of the 22nd ACM International Conference on Multimedia. 10.1145/2647868.2655045