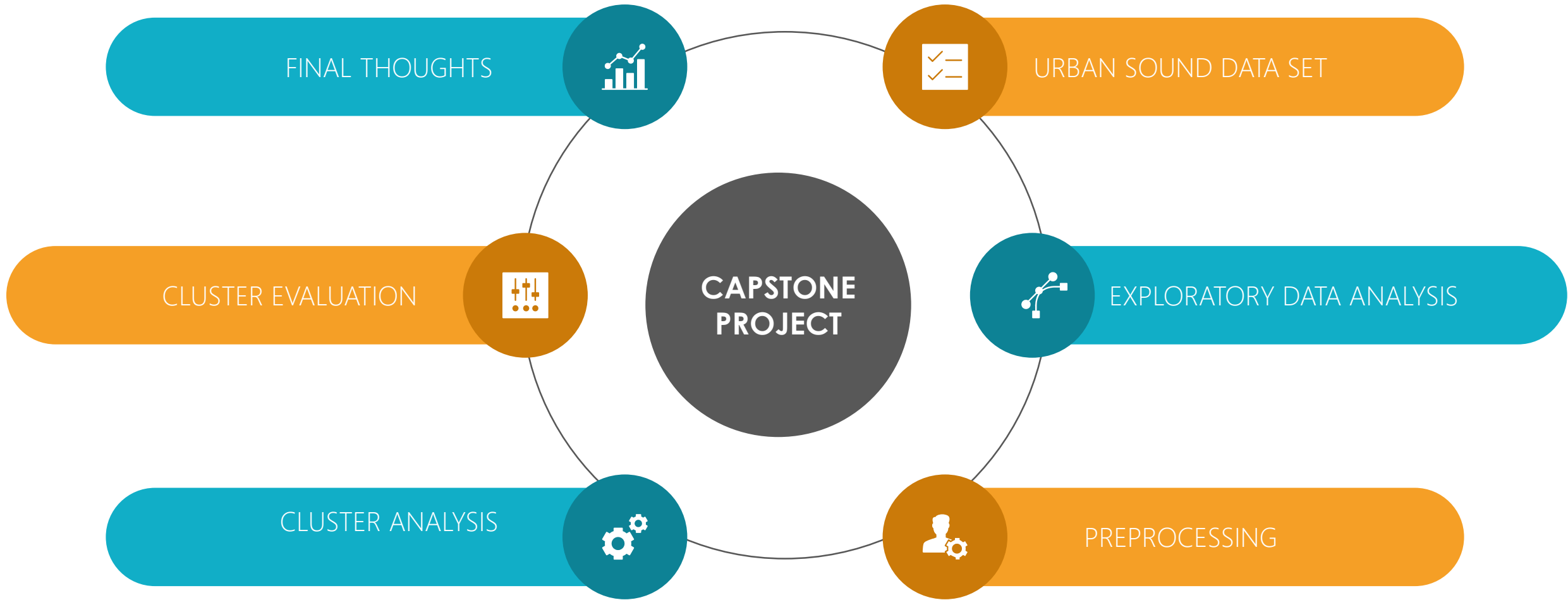




Urban Sound Classification

Unsupervised Learning
Capstone Project

URBAN SOUND CLASSIFICATION



PROJECT DESIGN

INTRODUCTION

- Sonic event classification is a field of growing research.
- Most of these researches focuses on music or speech recognition.
- Scarce works on environment sounds.
- Very few databases for labeled environment audio data.

EDA

- Urban Sound 8k Dataset contains 8723 audio excerpts in .wav formats from 10 classes of different sound sources
- We identified audio files with duration < 1 sec.
- Great variation across audio samples.

PREPROCESS

- Our biggest challenge is feature extraction.
- Audio data cannot be expressed in vector forms like other type of data such as images and texts.
- Applied various feature extraction techniques.

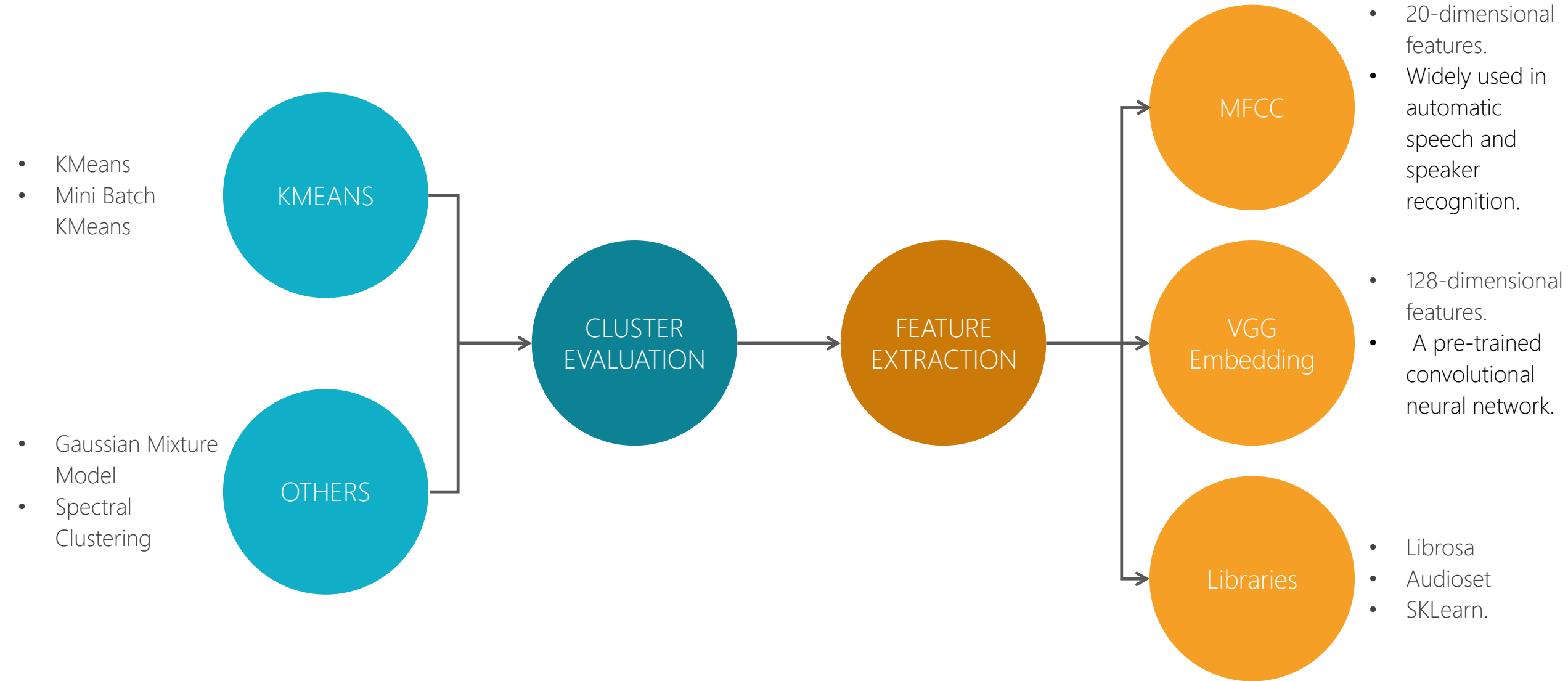
CLUSTER ANALYSIS

- Most common method in Unsupervised Learning to group unsorted information.
- Various methods to find the optimal number of clusters:
 - Elbow
 - Silhouette
 - AIC/BIC

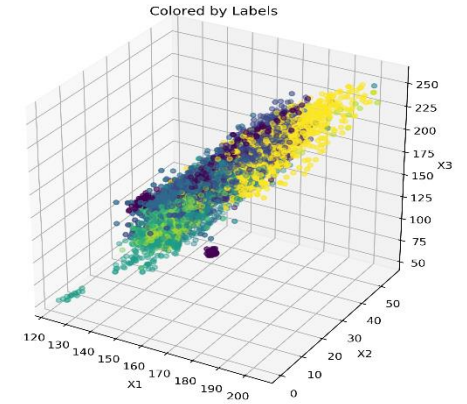
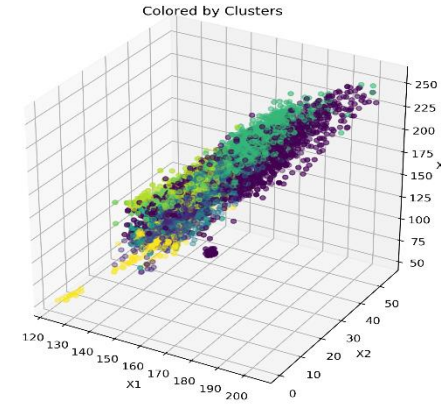
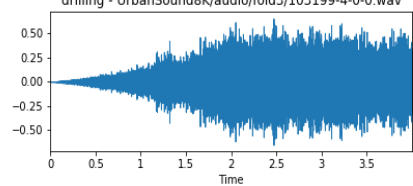
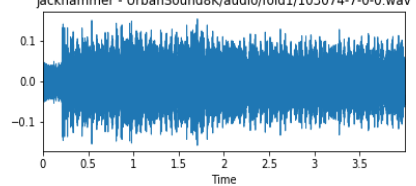
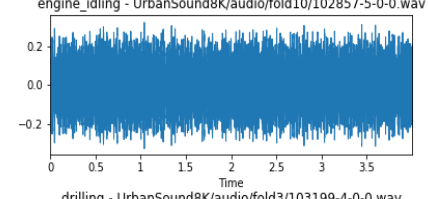
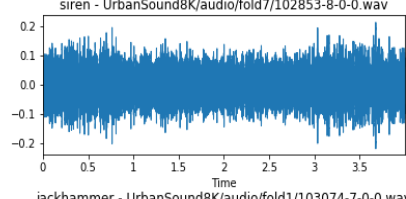
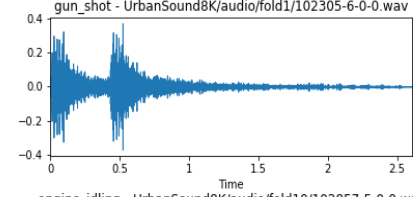
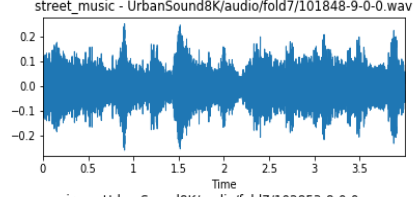
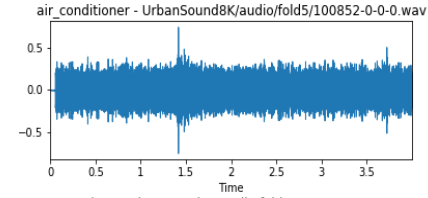
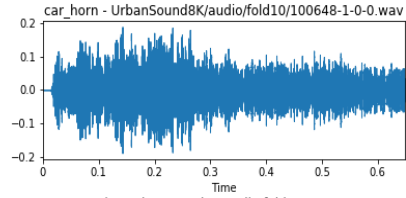
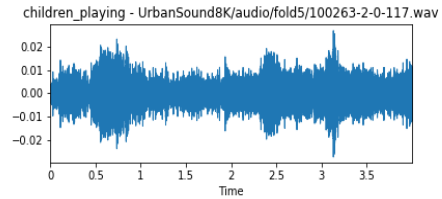
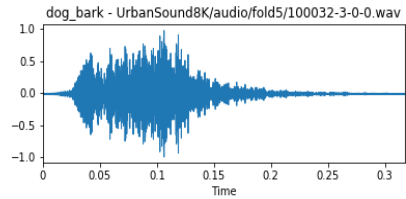
CLUSTER EVALUATION

- Review various clustering algorithms.
 - KMeans
 - GMM
 - Spectral Clustering
- Evaluate cluster performance metrics:
 - Adjusted Rand Index Score
 - Silhouette Score

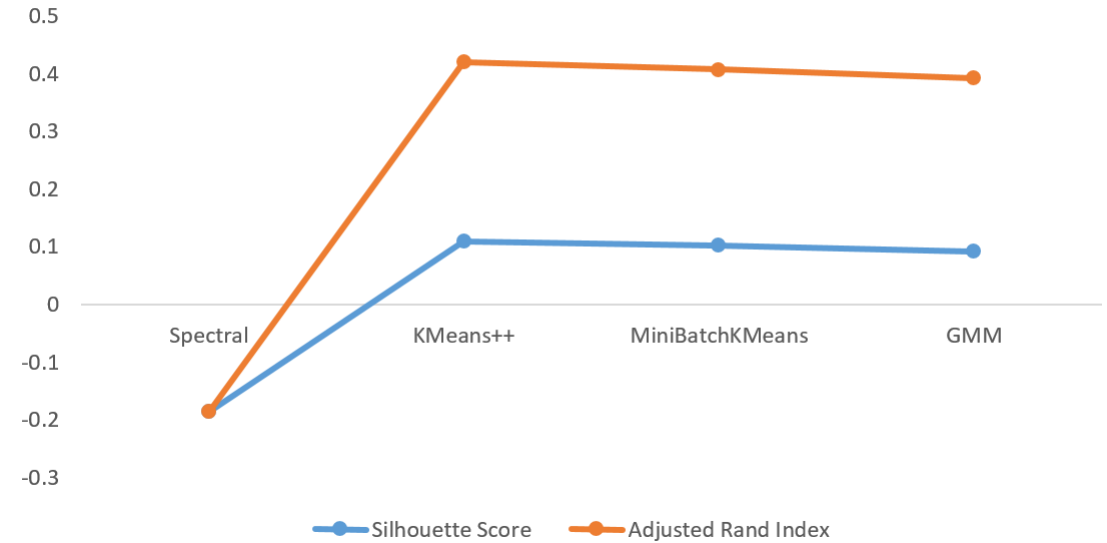
IMPLEMENTATION



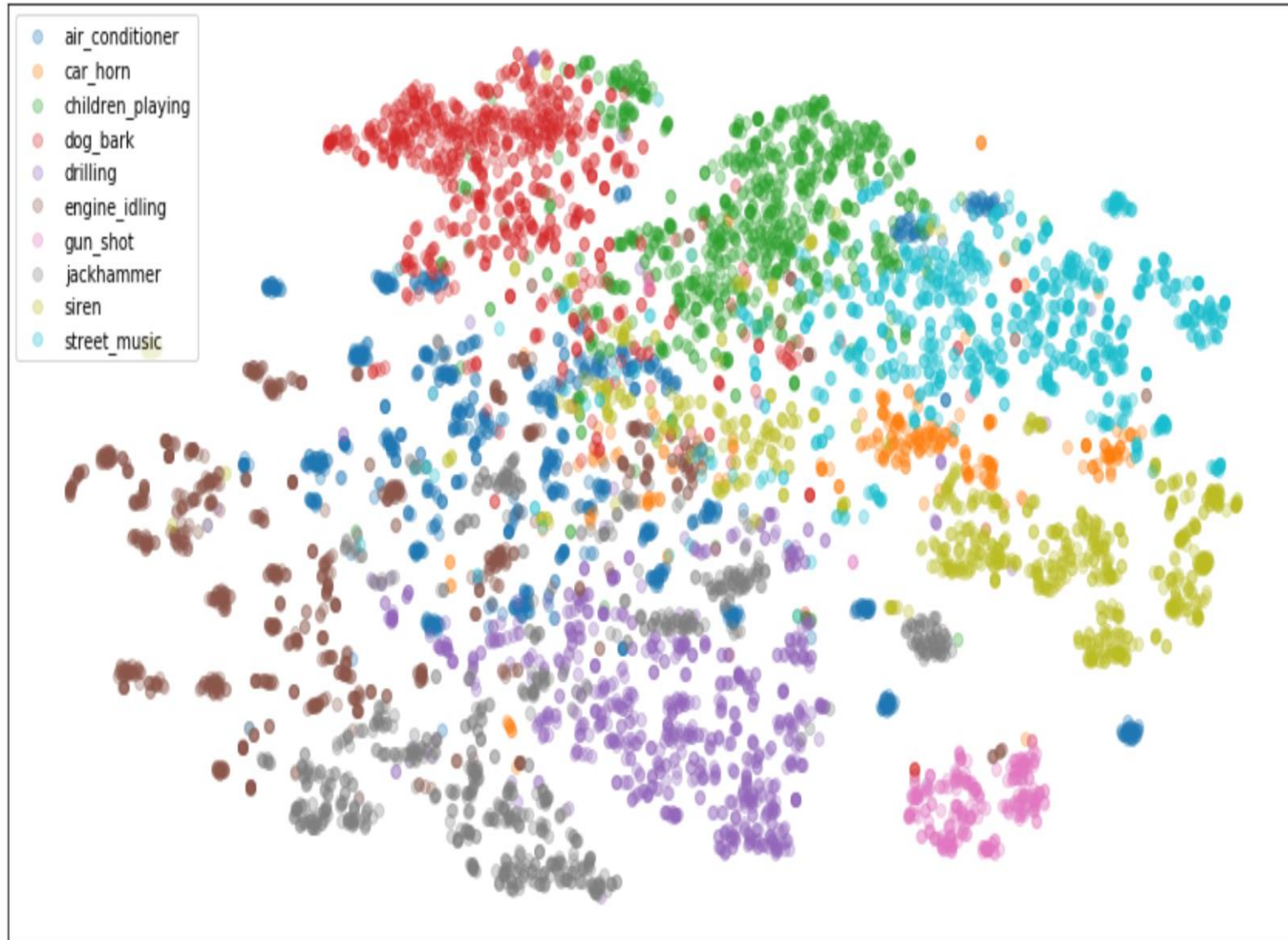
VISUALIZATIONS



Cluster Metrics



HIGHLIGHTS



- We can see that our models is making the same types of mistakes that a human might: air conditioners for drilling and drilling for jack hammering.
- Car horn and gun shot have less than 300 samples compare to other classes, which have around 1000 samples each.



- Having the lowest number of samples, gunshot is still managed to have the highest proportion for true positive value. However, the car horn is often misclassified as the street music.
- It is difficult to differentiate between jackhammer and drilling, however it is easy to discern between dog bark and drilling.

FINAL THOUGHTS

PRACTICAL USES

The automatic classification of audio events in an urban setting has a variety of applications. Some of them are listed below.

- Audio Event Detection
- Home security or Audio Surveillance
- Assisted living, elder or infant care
- Accident and crime surveillance

FUTURE CONSIDERATIONS

This capstone project focused on the unsupervised machine learning methods such as “**Clustering**” to draw inferences from datasets consisting of Unstructured Data without labeled responses.

- Using Supervised Learning to model the data to give us predictive power to classify the sonic events accurately. Improving this model to encompass several different methods to optimize prediction of the audio classification. In addition to the modeling, one can also implement tensorflow, keras and neural networks.
- We can potentially improve the quality of life of city dwellers by providing a data-driven understanding of urban sound and noise patterns, partly enabled by the move towards “smart cities” equipped with multimedia sensor networks.



Thank You