This repo is about a concept study on the MIMII dataset to detect anomaly of machines or machine parts like fans, slider, pump and valves by means of classic machine learning and deep learning methods.

In condition monitoring of machinery, it is common to use structure-borne sound and order tracking (rpm, etc.) to detect malfunctions, for various reasons like ease of retrofitting or need for a mobile solution and size of the machine part or operational needs like zero downtime, airborne sound anomaly detection could be preferable. The proof of concept study conducted here, shows that by applying machine learning anomaly detection on acoustic sensing, an ML/AI sensor can be constructed that has good performance so that it can compete with a trained technician when detecting anomalous by listen to the machine, with the potential benefit of staying 24/7 at the machine part of interest. The development can be seen as ground work for an embedding solution of smart sensor as a part of IoT plant supervision system (like SCADA).

Anomaly detection with machine learning means mostly unsupervised learning as the base assumption is that abnormal operation is unknown. Abnormal operation could be potentially very diverse of nature, so even if recordings of abnormal operation would exist training on them would lead to over fitting. Furthermore, the application of smart sensor would be lesser useful.

This means a smart senor microphone system needs to be trained by being placed for a reasonable training time in front of a healthy machine part, to become an armed detector that learned what is normal under representative background noise. Here in lays also the limit of the study as in a real-world scenario more machine parts are connected and the dataset specifically focuses on one part. But a general abnormally also of ensemble of machine parts could be possible with the same technique .

To reach a optimum architecture various machine learning techniques are explored and eventually a divers ensemble connecting, the following list summarizes the techniques explored:

* Stochastic model (a multi-dimensional normal distribution is found and outlier defined by significant)
* Random Isolation trees (a decision tree depth is taken assuming outlier need only view decisions to be found as spatial separated in one or more feature)
* Auto encoder (an underrepresented auto encoder reconstruction error is taken)
* Pseudo supervision (where normal observation is augmented/ distorted to train a binary classifier)

In the picture bellow a expletory classification ensemble is sketched, this is a example to show the main parts that are:

* Feature extraction pre-filter like BSS blind source separation or denoising filter
* Feature extraction like welch spectra (PSDs) or MEL spectra
* Classifier like RFC Random Forrest Classifier
* Stochastic models for outlier detection like GMM Gaussian Mixture Model, etc.
* Unsupervised outlier detection like Neural Network Autoencoder
* Or Outlier classification like IFC isolation forest classifier
* The time frame-based ensemble collects different classification over the time processed as some classifier work on longer buffer parts then others this may helps to regulate the training of the algorithms and can improve training speeds.

In order to make any machine learning algorithm able to work with audio it is necessary to use various signal processing steps (feature extraction pre filter) that may be of classic nature or also take use of machine learning methods like clustering for preprocessing on the time buffer like activation detection.

In application the reaction time of such a abnormally detection is around 10 sec at the current construction and training chain build up but some indication could be found to reduce this evetnuly.