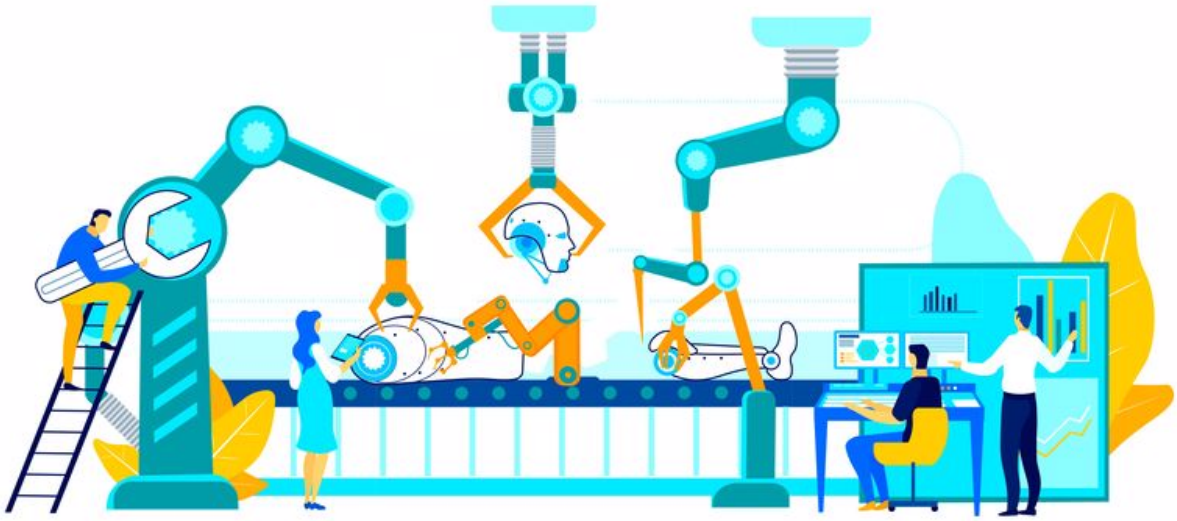


Machine Condition Monitoring

Reena Koshta
Theano 3

Introduction

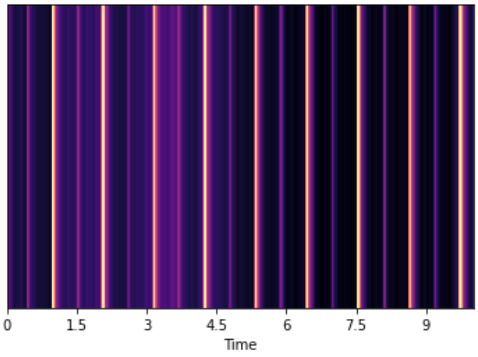
The main objective of this project is to analyse different machine learning models and create the best machine learning model that will be able to monitor the operations and identify anomalies in sound patterns.

We have audio files for 4 different machines. And each machine has normal and abnormal sound files. With the help of these files we have to create a machine learning model which is able to detect abnormal conditions.

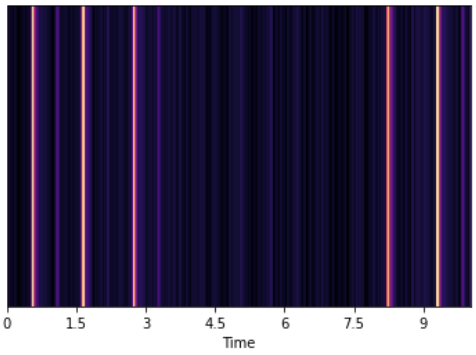
Data Preprocessing

We have analysed information that we can get from the audio file. And we found **Mel Frequency Cepstral Coefficients (MFCCs)** as one of the important sound feature. Here are visuals that we got with MFCC.

Spectrogram For MFCC with respect to Time



Normal



Abnormal

Approaches

1. **Classification**

   In my previous project I worked with my team to create the classification model.

   We have created a universal model that works for all 4 types of machine.

   Where we found that the model is working well with sliders and pumps but for fans and valves it is not giving the correct predictions.

   For better understanding you can refer Machine Condition Monitoring report.

   Also you can find the classification model at git repo .

2. **Clustering**

   In this project we explored some clustering algorithms of machine learning for unsupervised data(data without labels that means we don't know if the data is for normal sound or abnormal sound).

   Here we created separate models for each type of machine and applied different clustering algorithms.

   For code and detail analysis this git repository can be used.

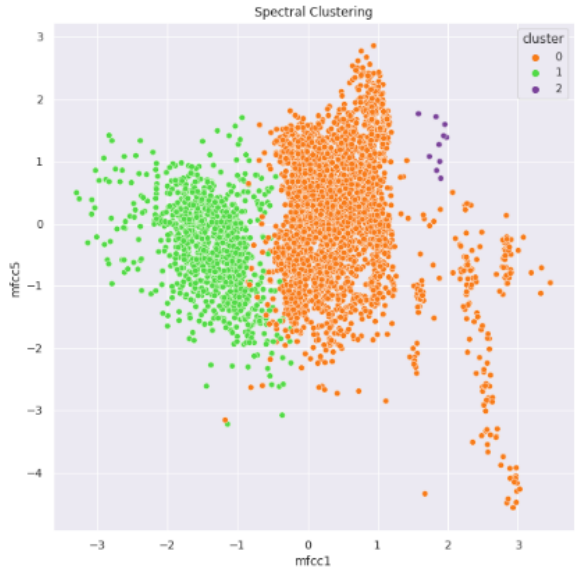
Analysis

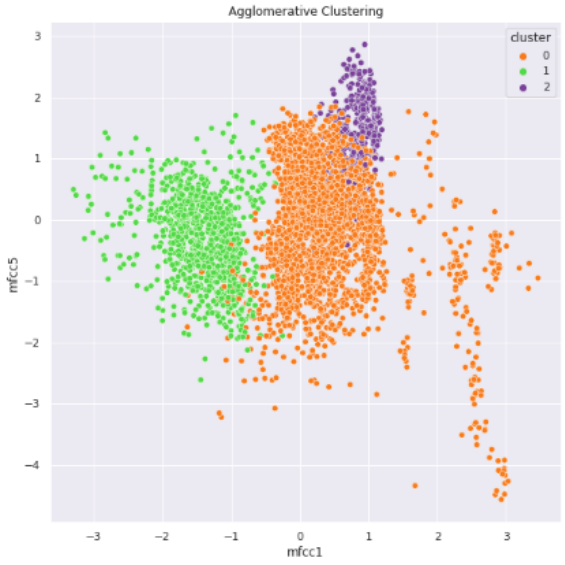
For all the machines we tried the algorithm below.

- Kmean
- Spectral Clustering
- Hierarchical clustering with Agglomerative Algorithm

Machine wise observation are as below:

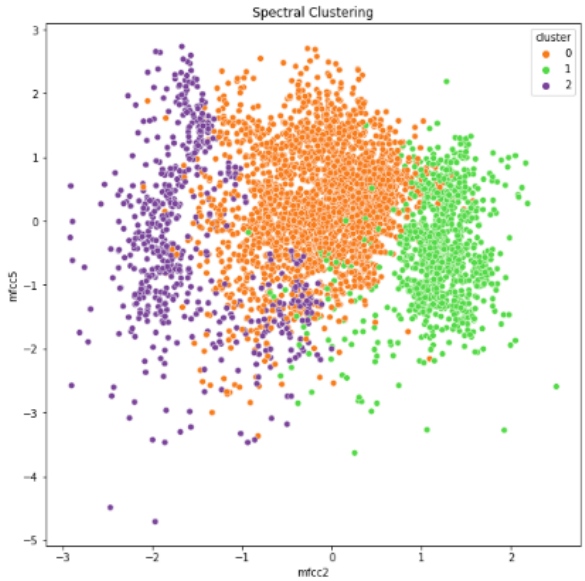
Pumps :

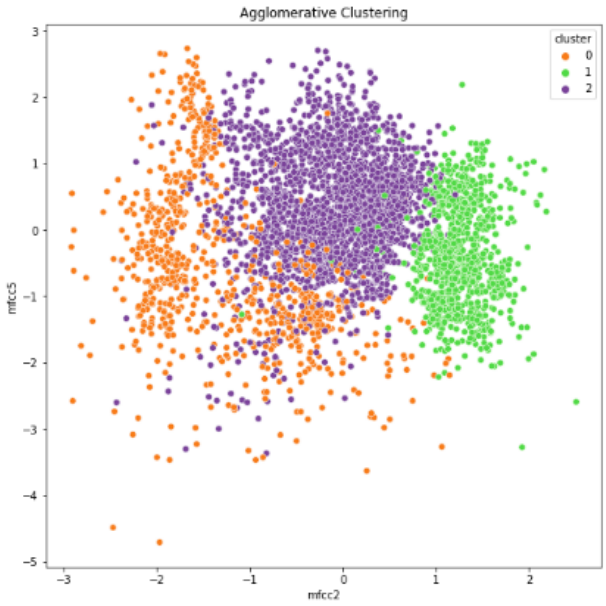




From the scatter plot above I feel Kmean is working well for Pumps and we assume that class 1 is abnormal sounds, class 0 is transition and class 2 is the normal sound.

Sliders :

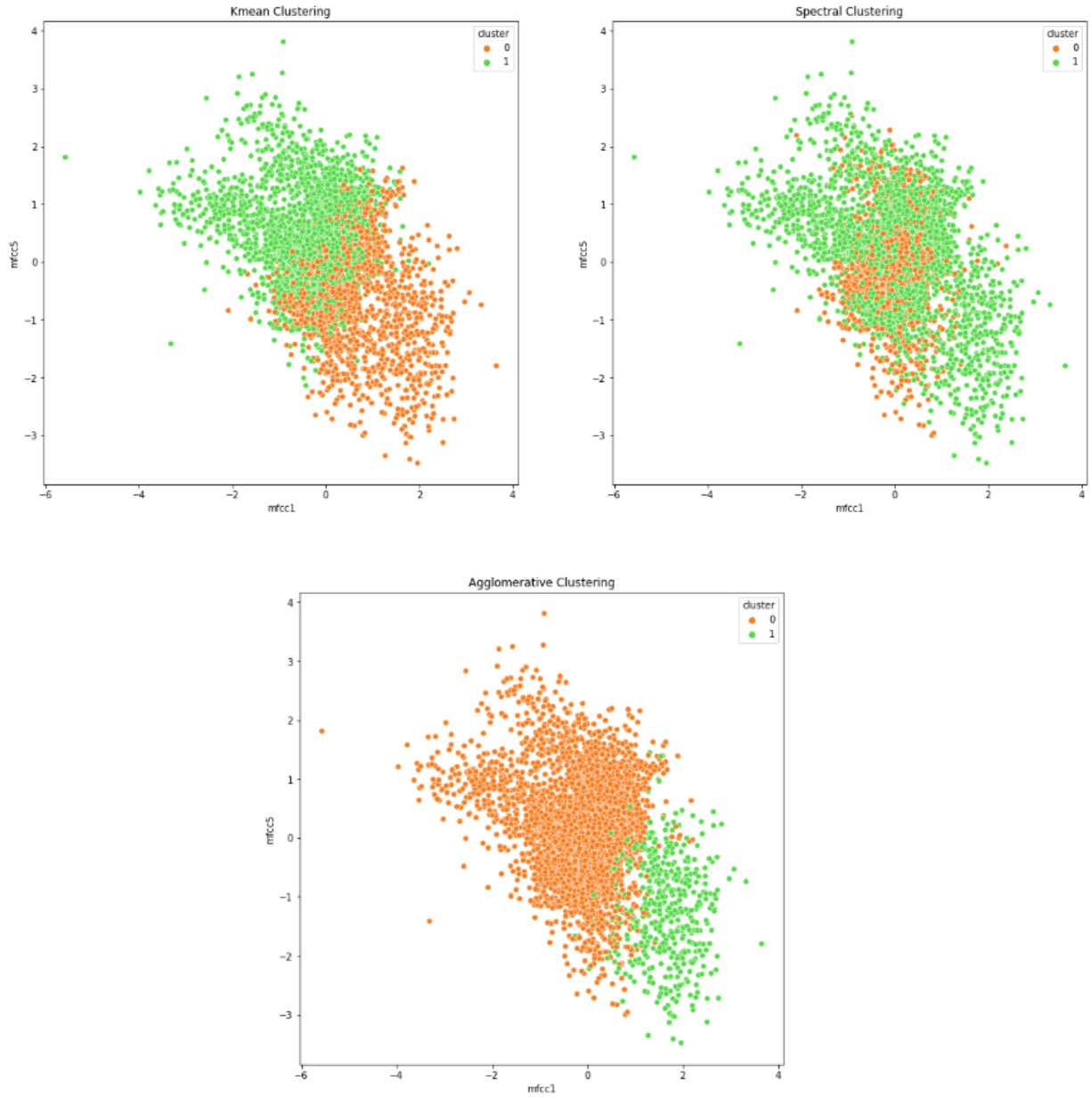




From the scatter plot above I feel Kmean is working well for Sliders also and we assume that class 0 is abnormal sounds, class 1 is transition and class 2 is the normal sound.

We got rand score as 0.5489601737617014 for Kmean which says it is giving almost ~55% predictions correctly.

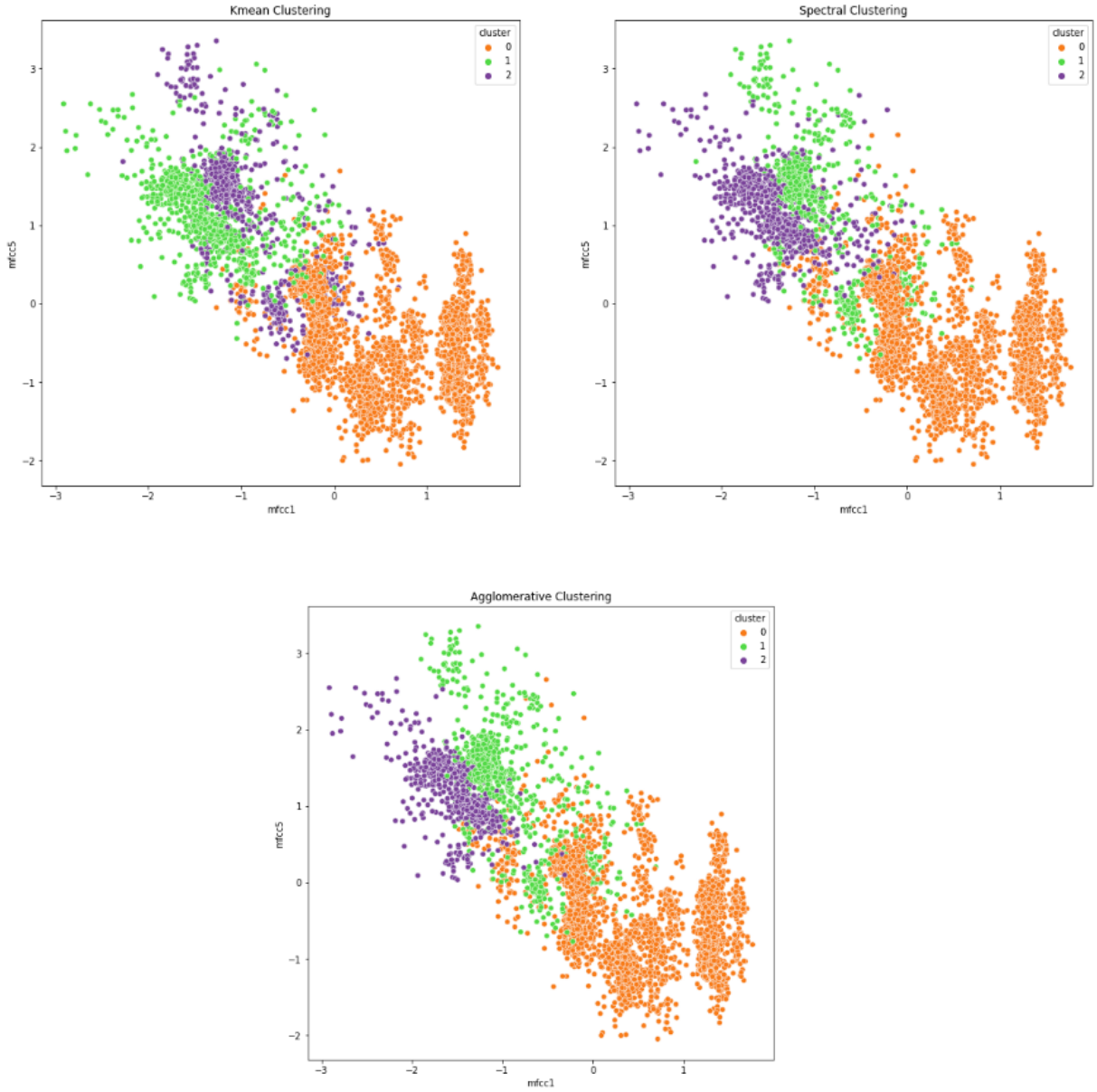
Valves :



Here we have only checked for cluster count 2 . From the scatter plot above I feel Agglomerative clustering is working well for Valves and we assume that class 0 is normal sounds and class 1 is the abnormal sound.

We got rand score as 0.6753472731529336 for Agglomerative clustering which says it is giving almost ~67% predictions correctly.

Fans :



From the scatter plot above I feel Spectral clustering is working well for Fans and we assume that class 0 is normal sounds, class 1 is transition and class 2 is the abnormal sound.

Conclusion

- It is hard to create the one universal model which work for all type of machine.
- Based on analysis with different algorithms for different machine I can say that clustering is also not an effective way of finding the anomaly in this sound dataset. Because if you compare the result from the clustering model with the actual data we found that there are so much difference.
- There might be some other machine learning approaches which work well with this type of data.