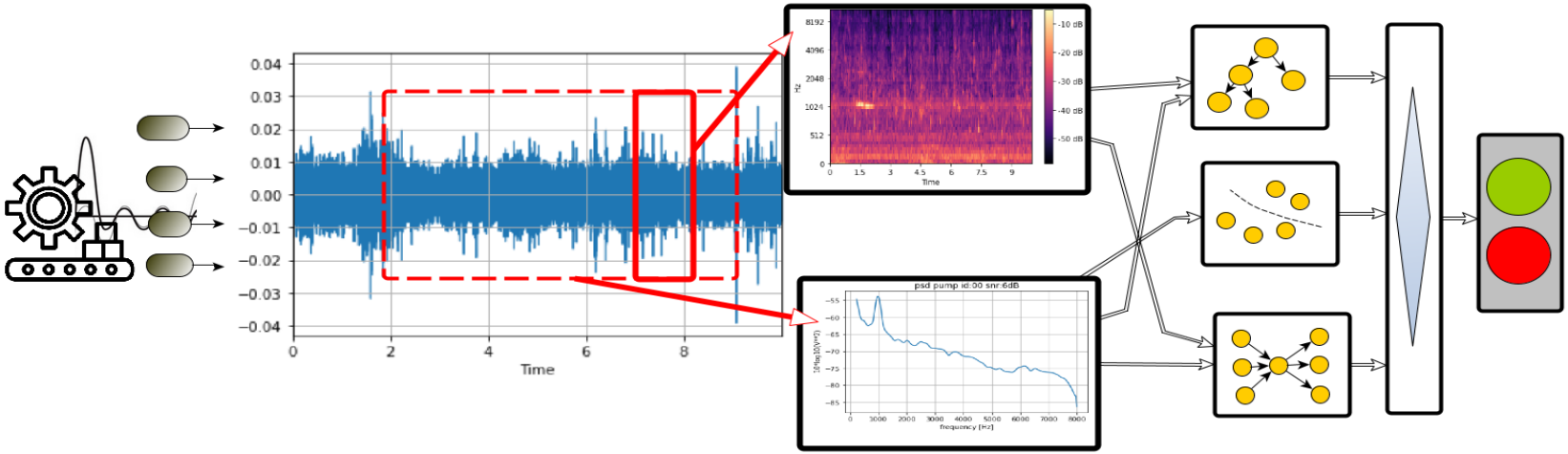
Graded Exercise 2:

Anomaly Detection on Acoustic Data

Teaching Assistant: Chenghao Xu
Dr. Zhan Ma

Date: Oct. 3rd, 2024



Reference: Can a Machine Hear If a Machine Is Broken?

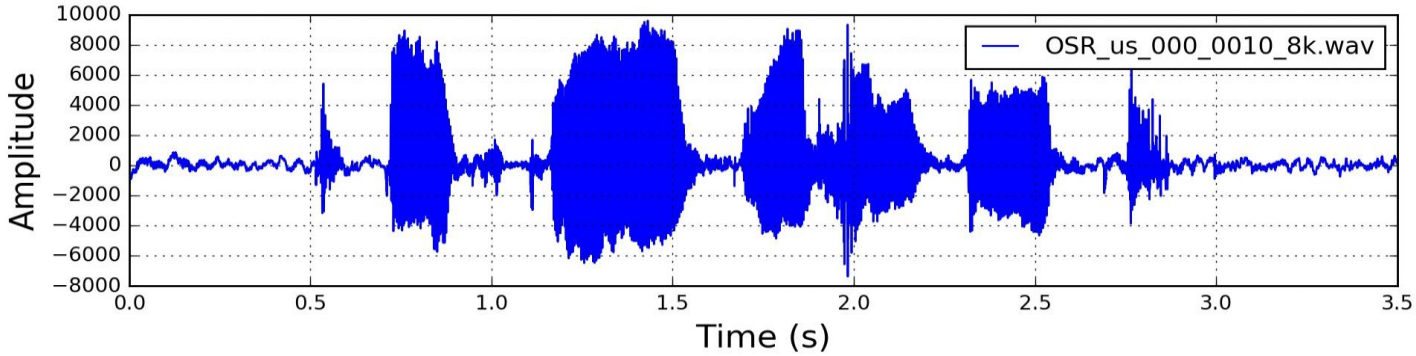
(Step 0: Signal Preprocessing)

Step 1: Feature Extraction

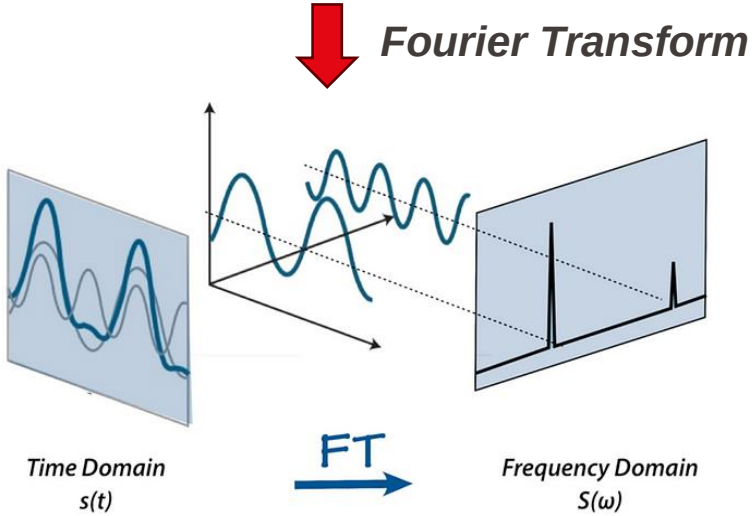
Step 2: Model Training

Step 3: Decision-making / Evaluation

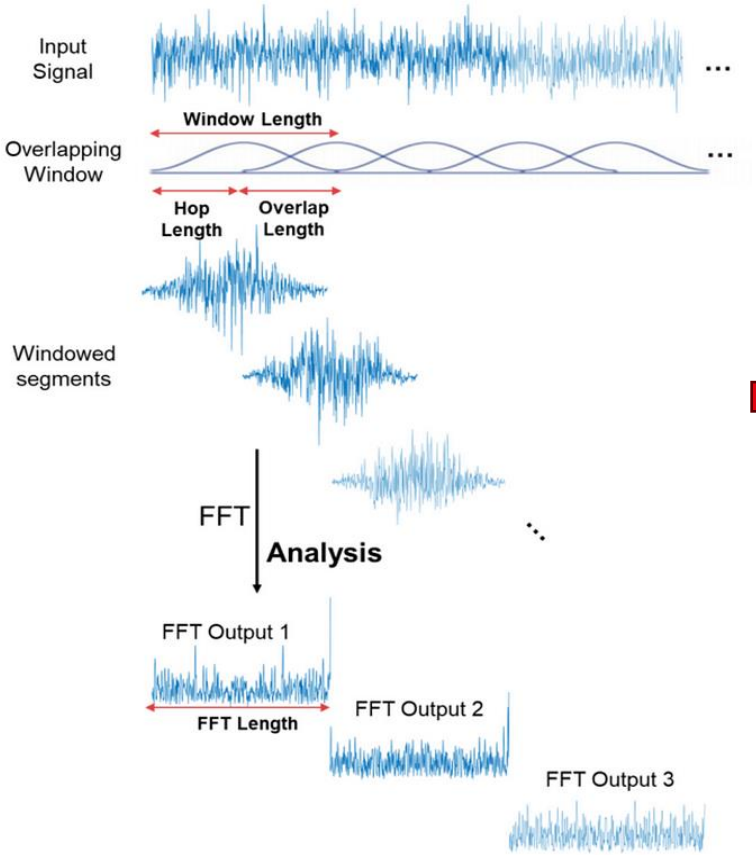
I – Mel Spectrogram

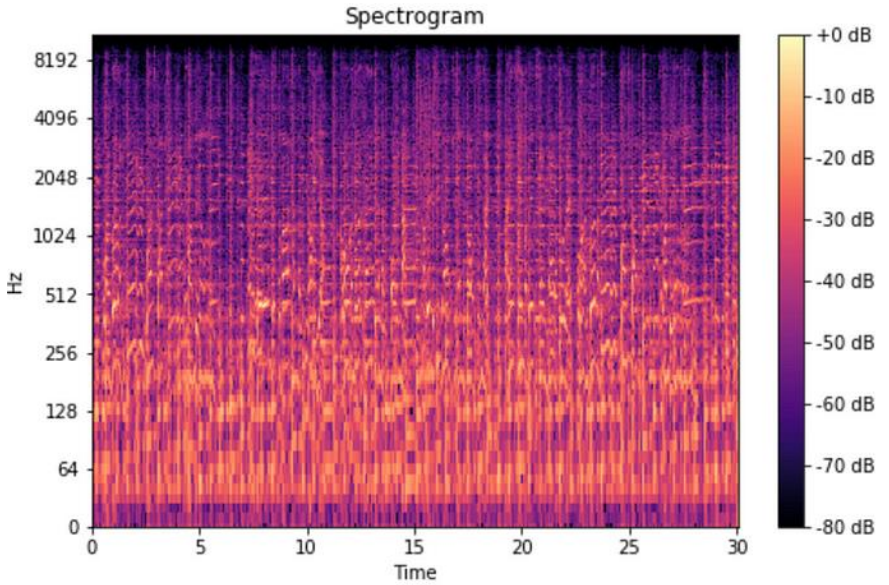


Digital representation of audio signal



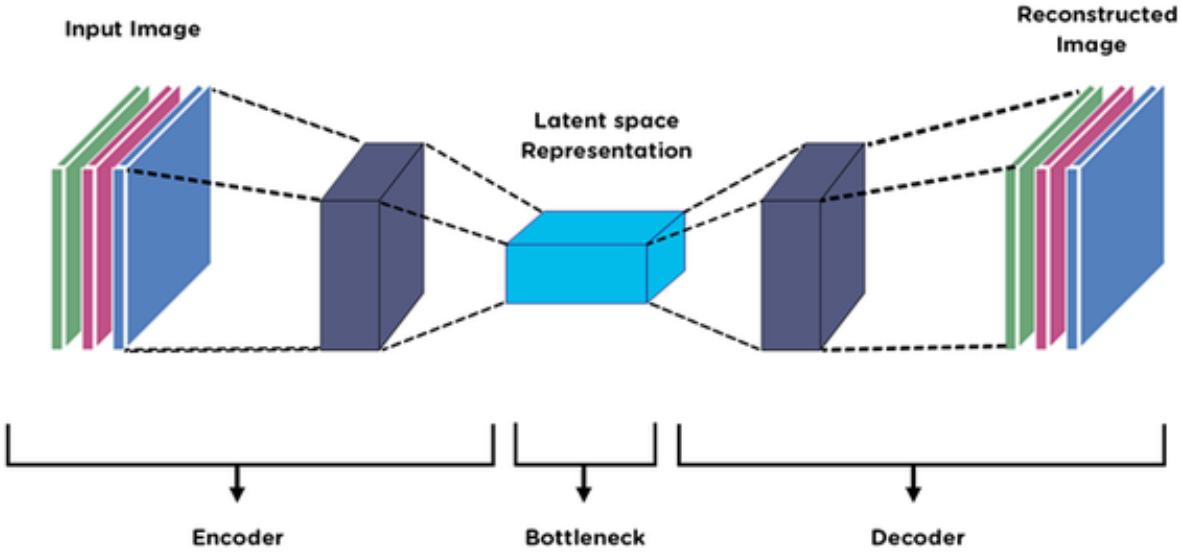
- **How to obtain spectrum on non-periodic signals?**





Mel Scale: Human better at detecting differences in lower frequencies than higher frequencies.

II – AutoEncoder

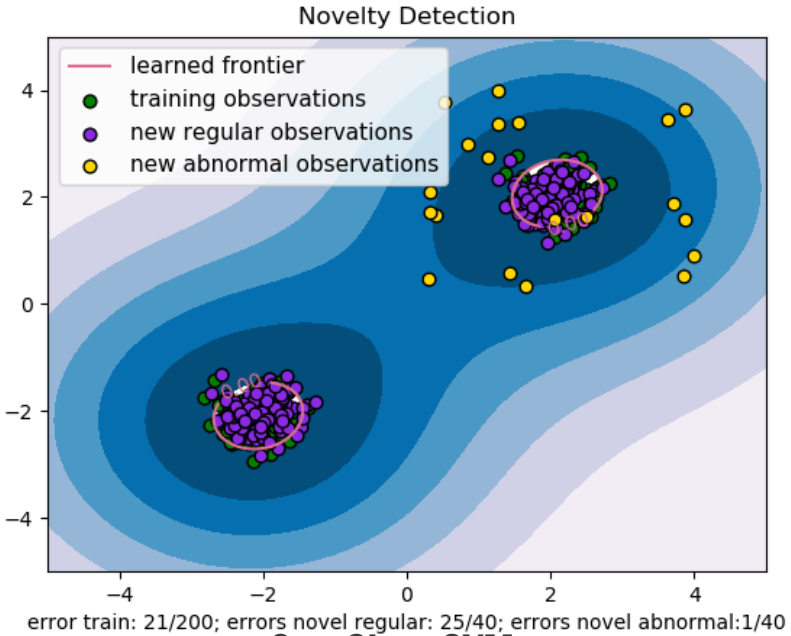


Autoencoder Architecture

An autoencoder is a type of neural network architecture designed to efficiently compress (encode) input data down to its essential features, then reconstruct (decode) the original input from this compressed representation.

III – One-Class SVM & Isolation Forest

- Anomaly Detection is widely applied in industrial application or imbalanced data distribution
- One-class SVM: find a hypersphere to separate normal samples and minimize its volume.
- Isolation Forest: find the sample point most likely to be separated (low density, and far from the cluster with high density)



One-Class SVM

Algorithm 1 : $iForest(X, t, \psi)$

Inputs: X - input data, t - number of trees, ψ - sub-sampling size

Output: a set of t *iTrees*

1: **Initialize** $Forest$
2: set height limit $l = ceiling(\log_2 \psi)$
3: **for** $i = 1$ to t **do**
4: $X' \leftarrow sample(X, \psi)$
5: $Forest \leftarrow Forest \cup iTree(X', 0, l)$
6: **end for**
7: **return** $Forest$

Isolation Forest

IV – Exercise Introduction