



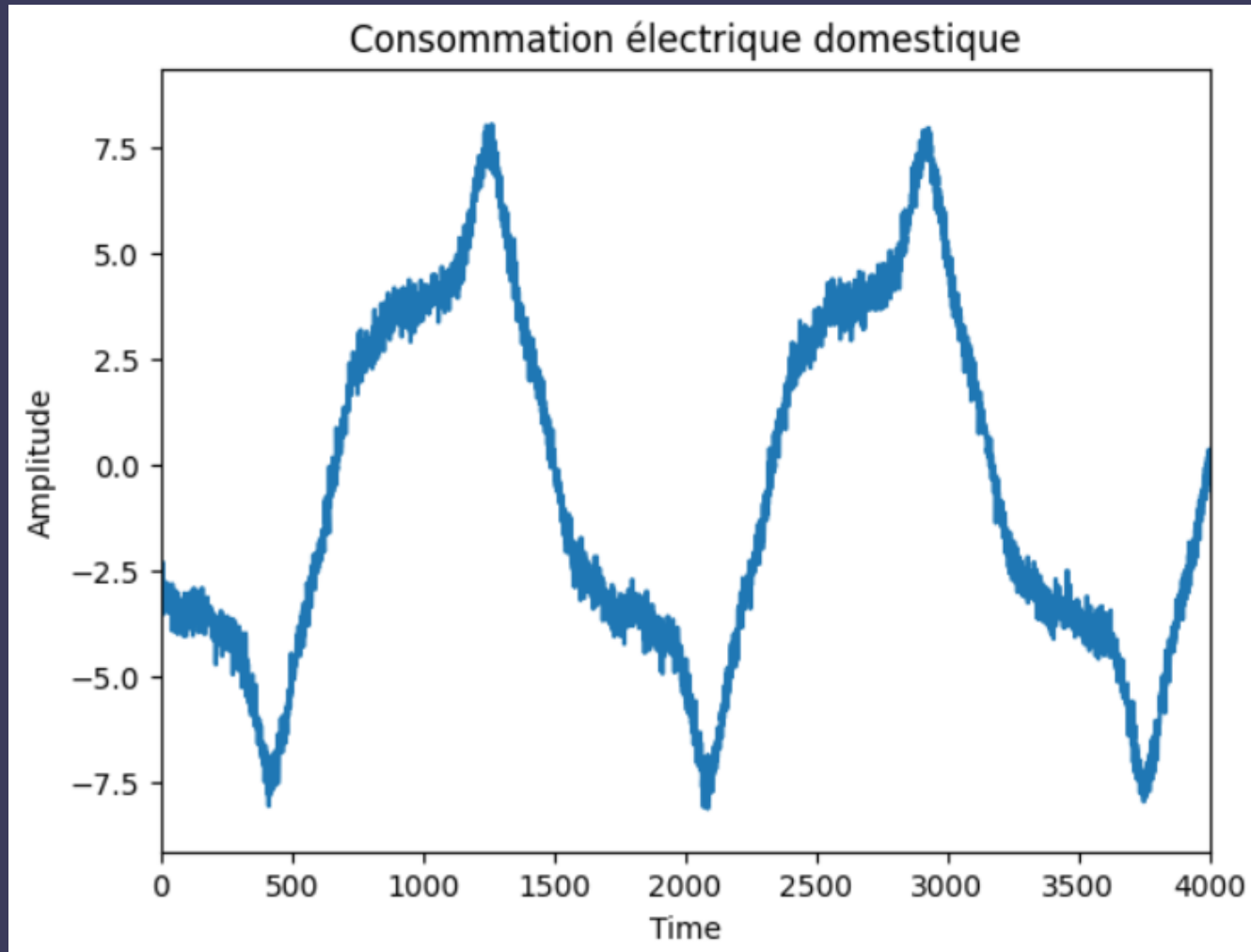
# Identification des appareils dans la consommation domestique

Projet: energy monitoring

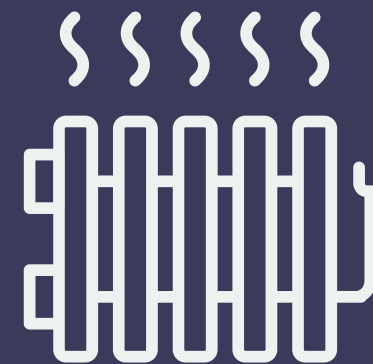
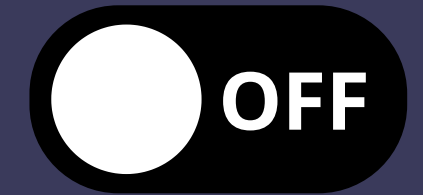
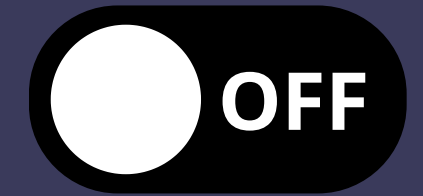
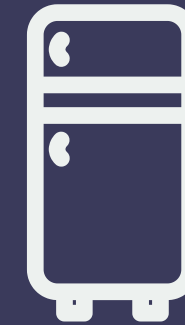
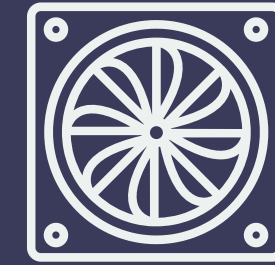
CASA NOVA - DAVID - PLOT - CHARDON



# I. Introduction:



Identification  
des appareils  
allumés



# I. Introduction:

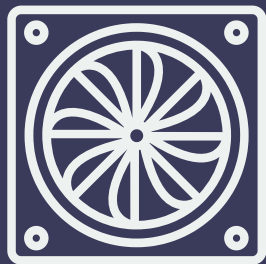


## Méthodes mises en place:

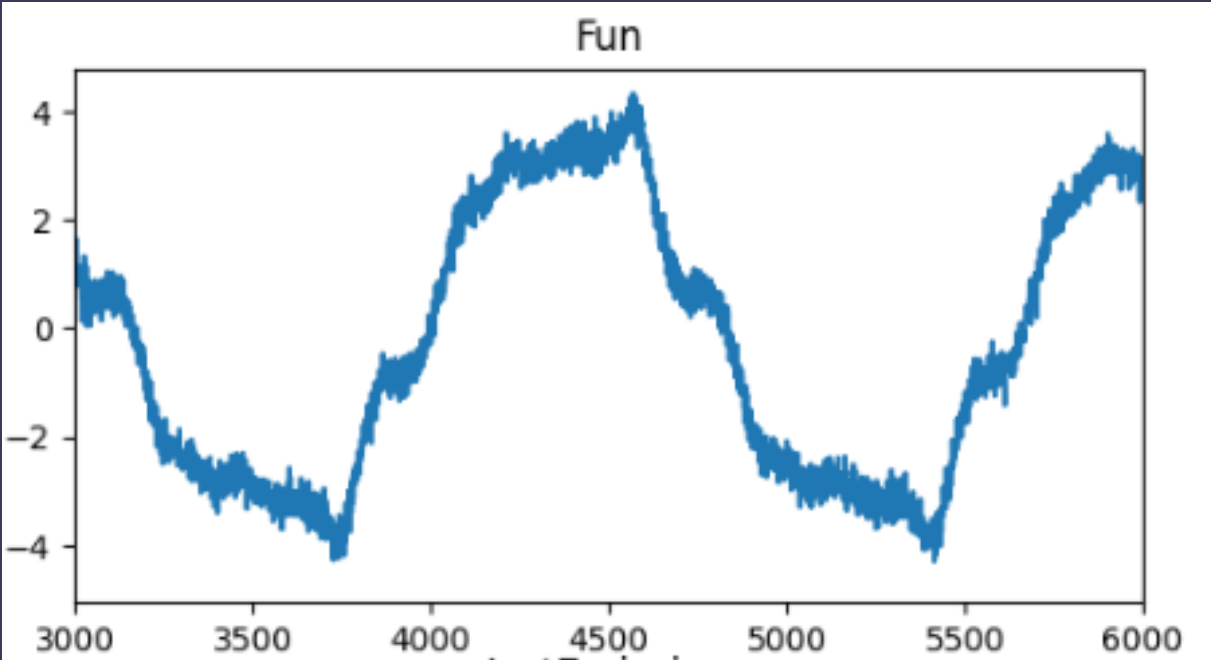
- Régression logistique multiple
- Support Vector Machine
- Convolutional Neural Network



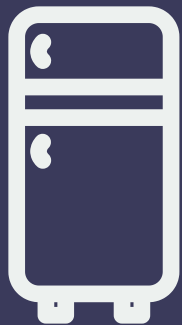
# A. Jeu de données initial



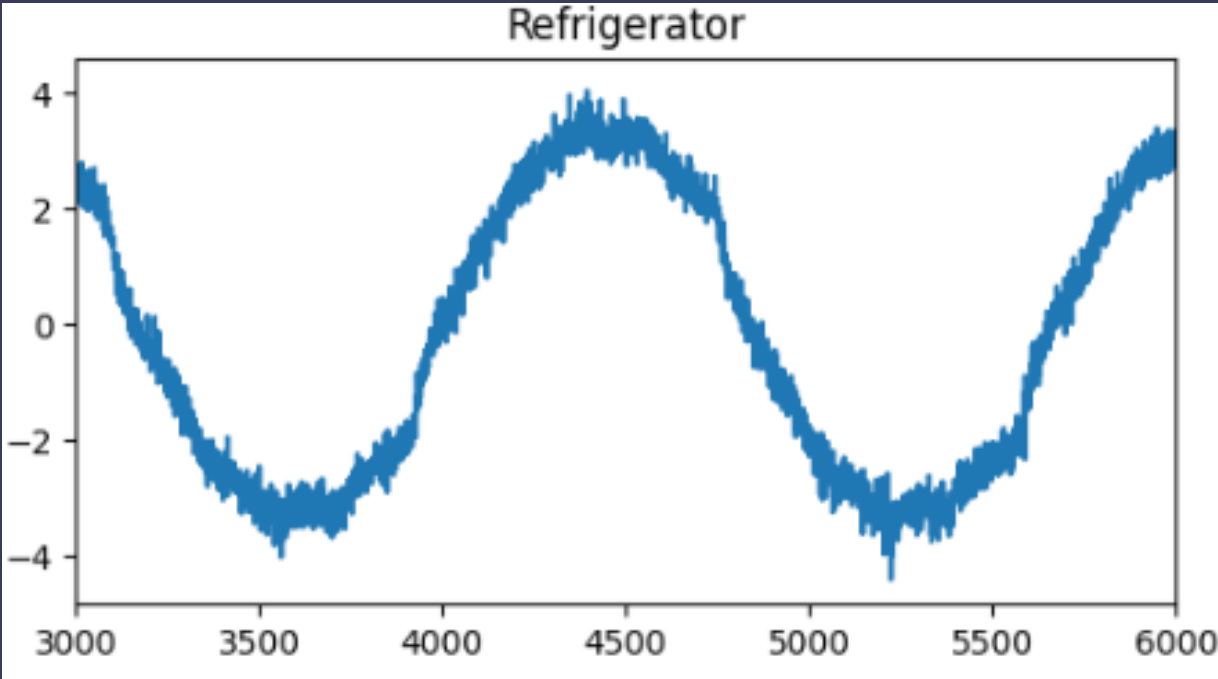
Fan



60 signaux



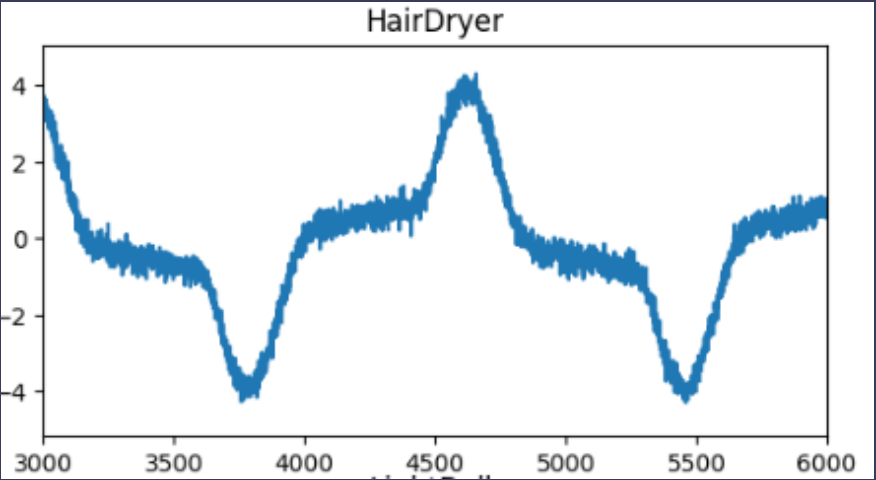
Refrigerator



60 signaux



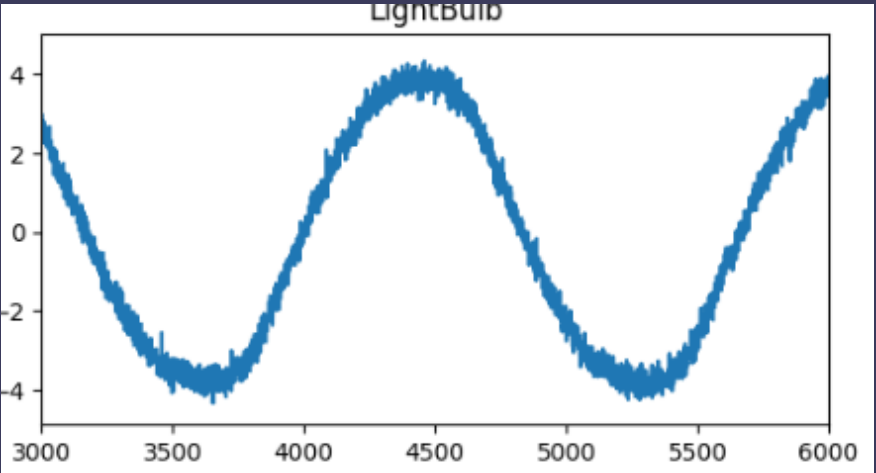
HairDryer



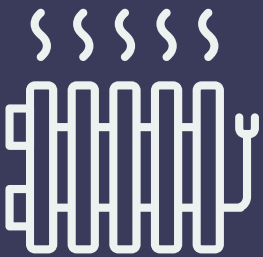
60 signaux



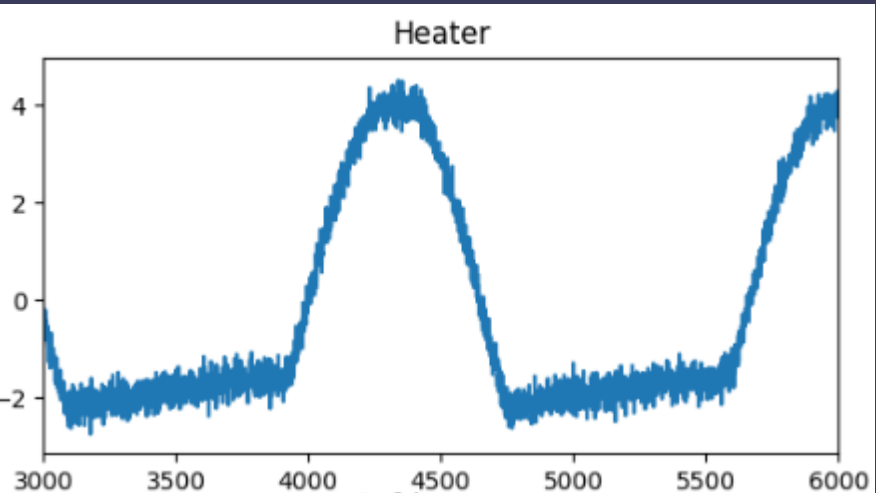
LightBulb



60 signaux

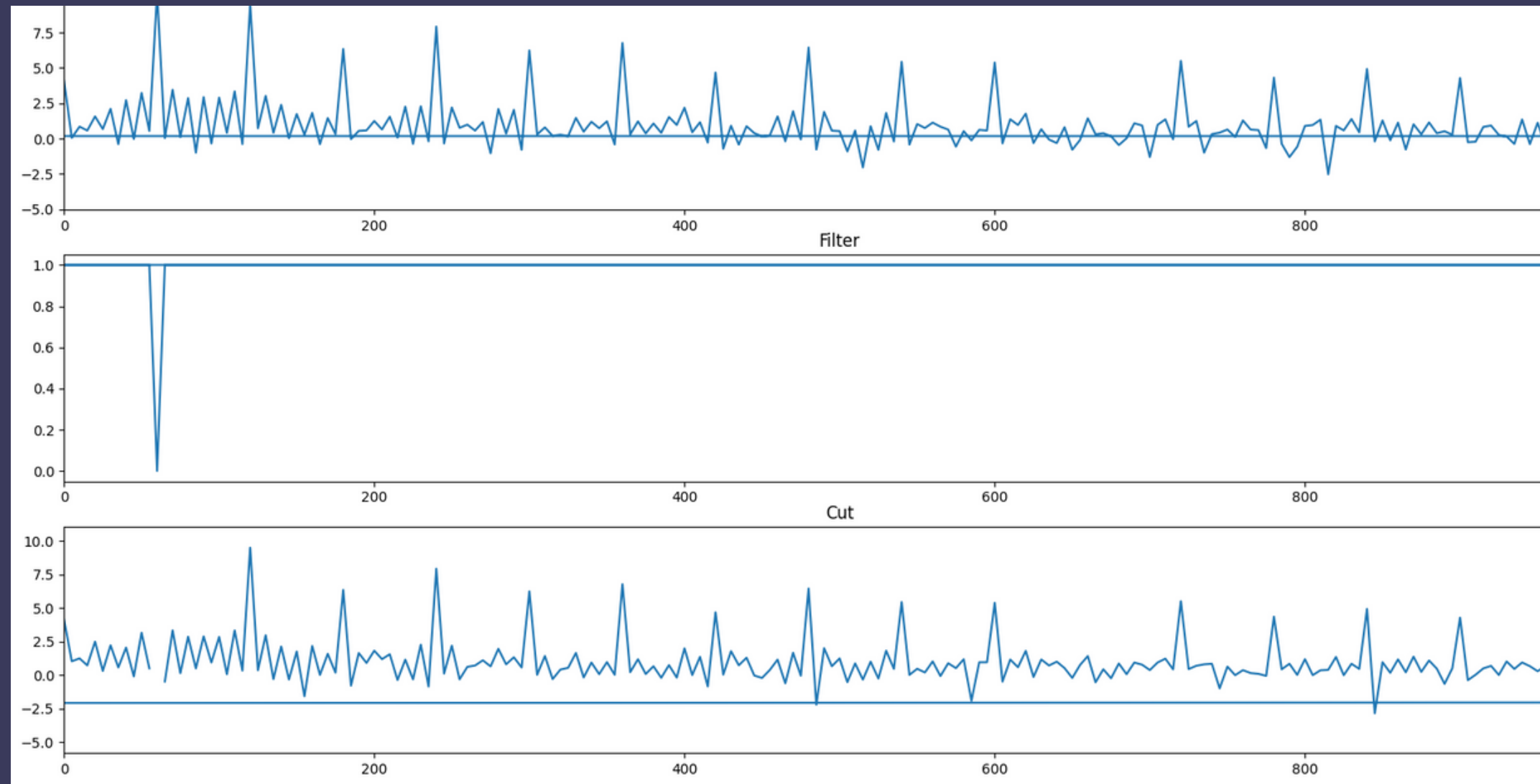


Heater



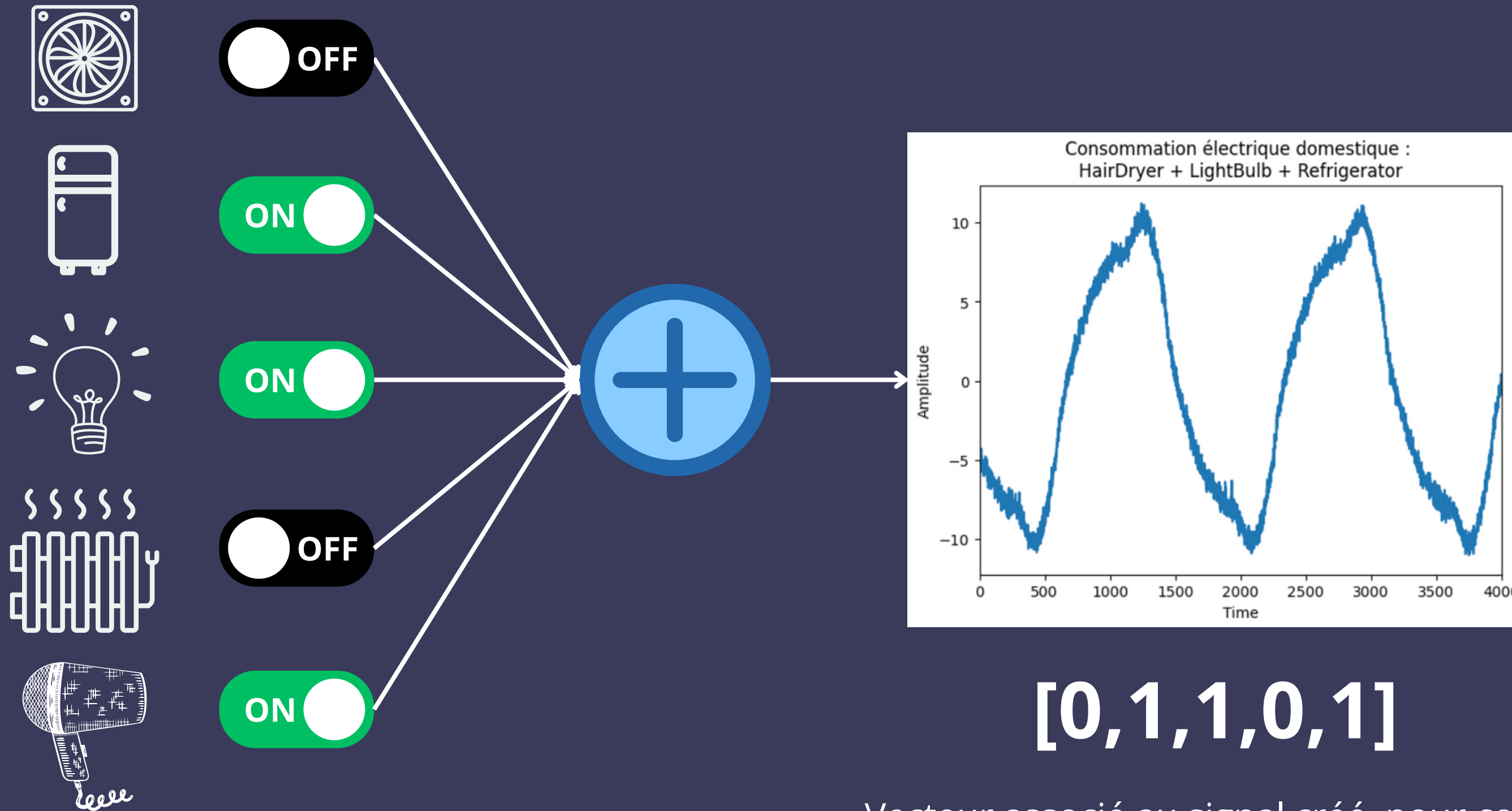
60 signaux

# B. Suppression du 60Hz



Représentation fréquentielle du signal initial (haut), du filtre et du signal filtré (bas)

## II. Création de données réalistes:

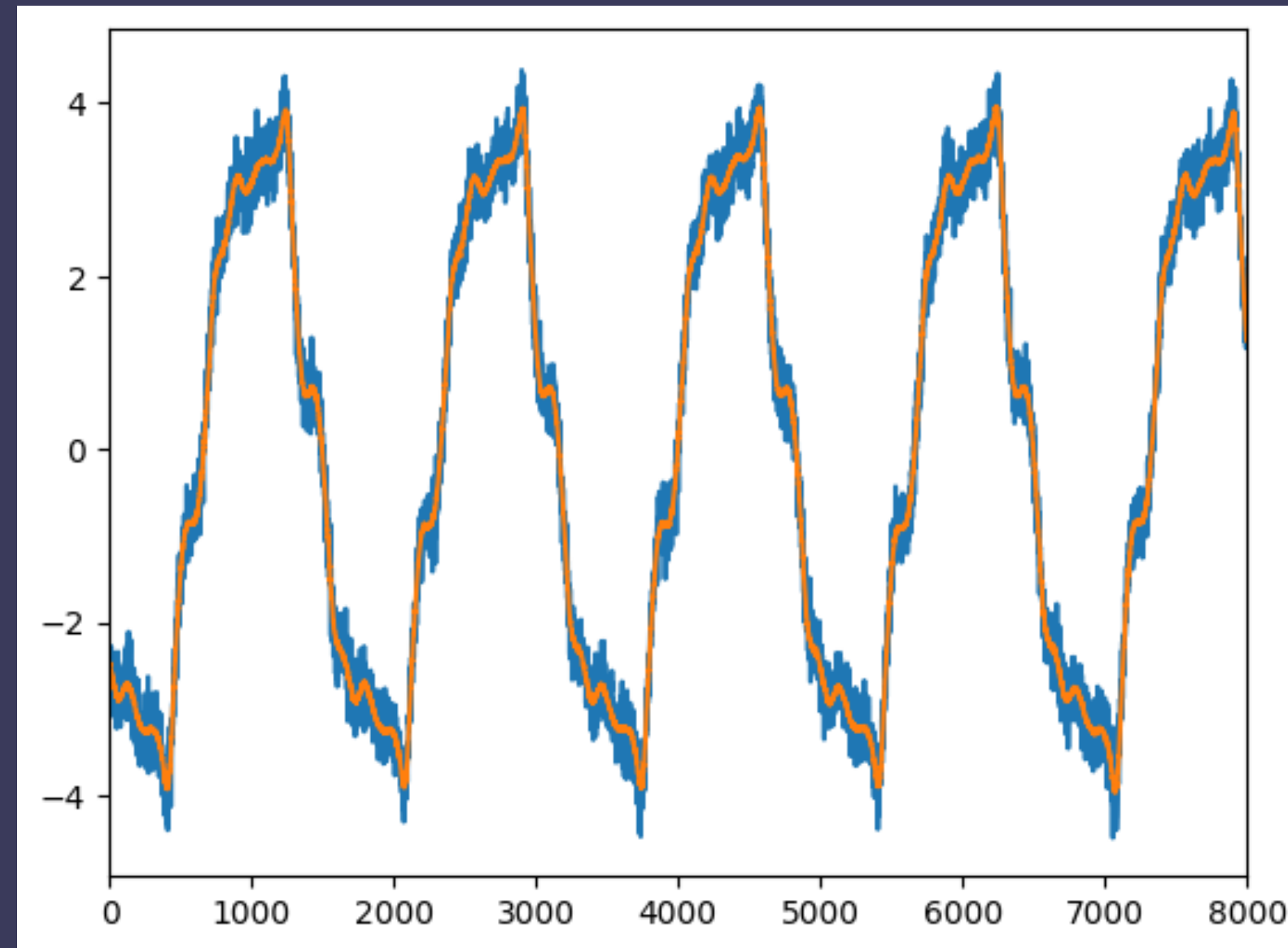


Choix aléatoire

Vecteur associé au signal créé, pour connaître les appareils allumés

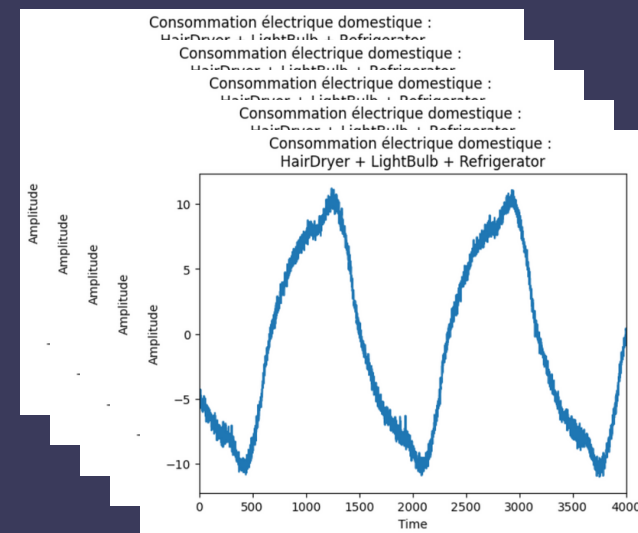
On obtient un jeu de données contenant 300 signaux

## II. Création de données réalistes: Bruit additionnel



Signal initial de ventilateur (orange) et signal bruité (bleu)

# III. Méthodes d'identification:



Consommation électrique

→ Dimension  
reduction

Input x

Création des  
datasets de training/  
validation  
(70% training/  
30%validation)



**Entraînement d'un  
modèle  
d'identification**

**[0,1,1,0,1] ... [0,1,0,0,1]**

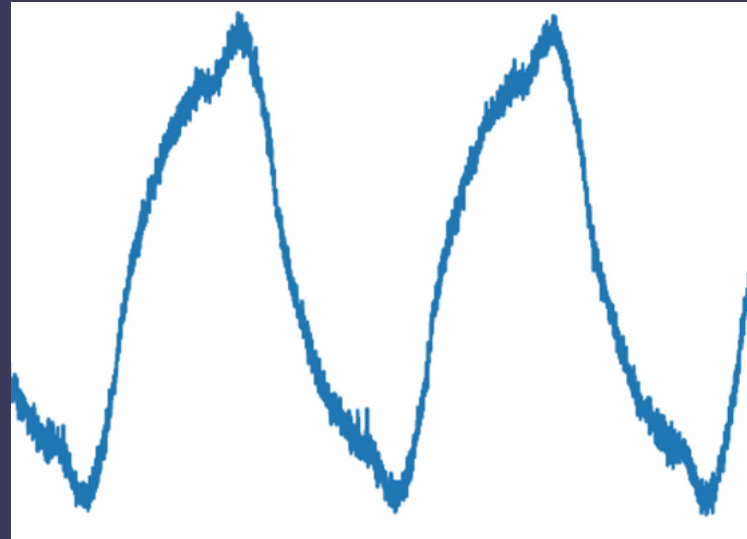
Vecteurs associés pour indiquer les  
appareils allumés

Output y  
prediction



# III. Méthodes d'identification:

Input vector : dimension 20000



Réduction de  
dimension

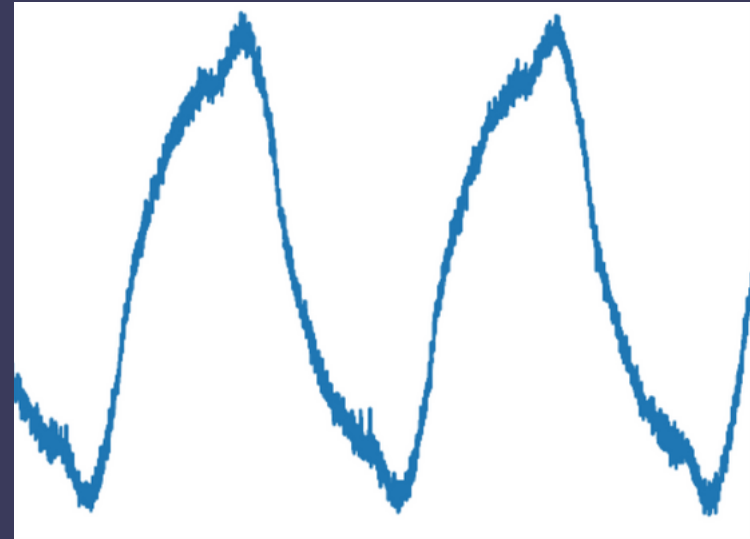


Features



# III. Méthodes d'identification:

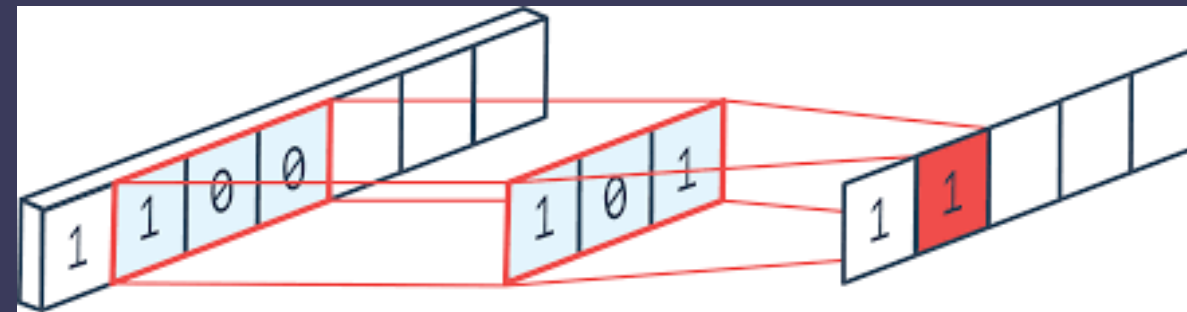
Input vector : dimension 20000



Réduction de  
dimension

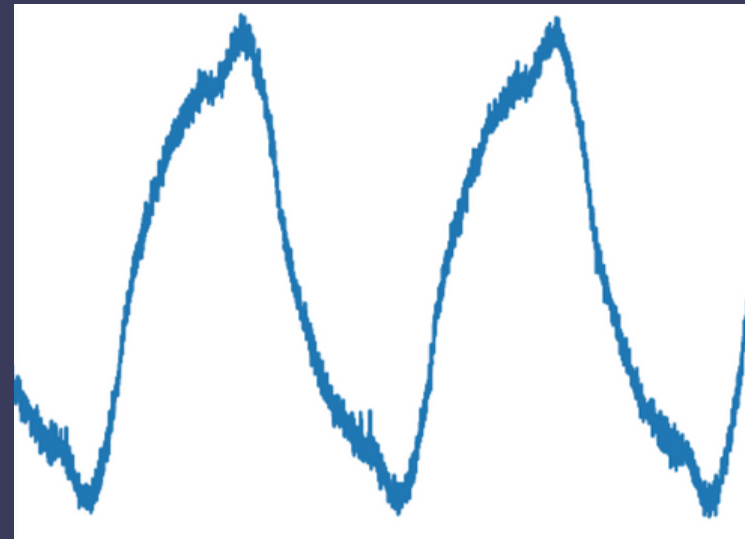


Features



# III. Méthodes d'identification:

Input vector : dimension 20000



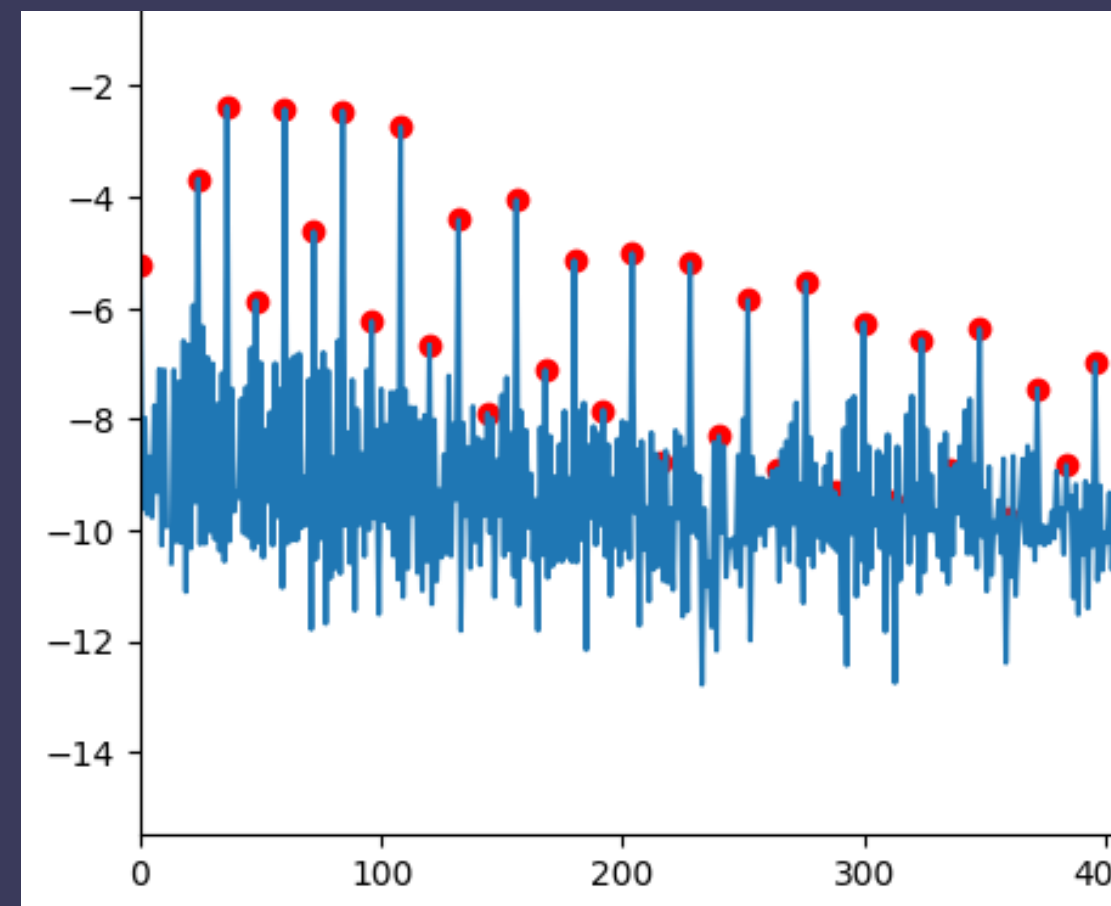
Réduction de  
dimension



Features



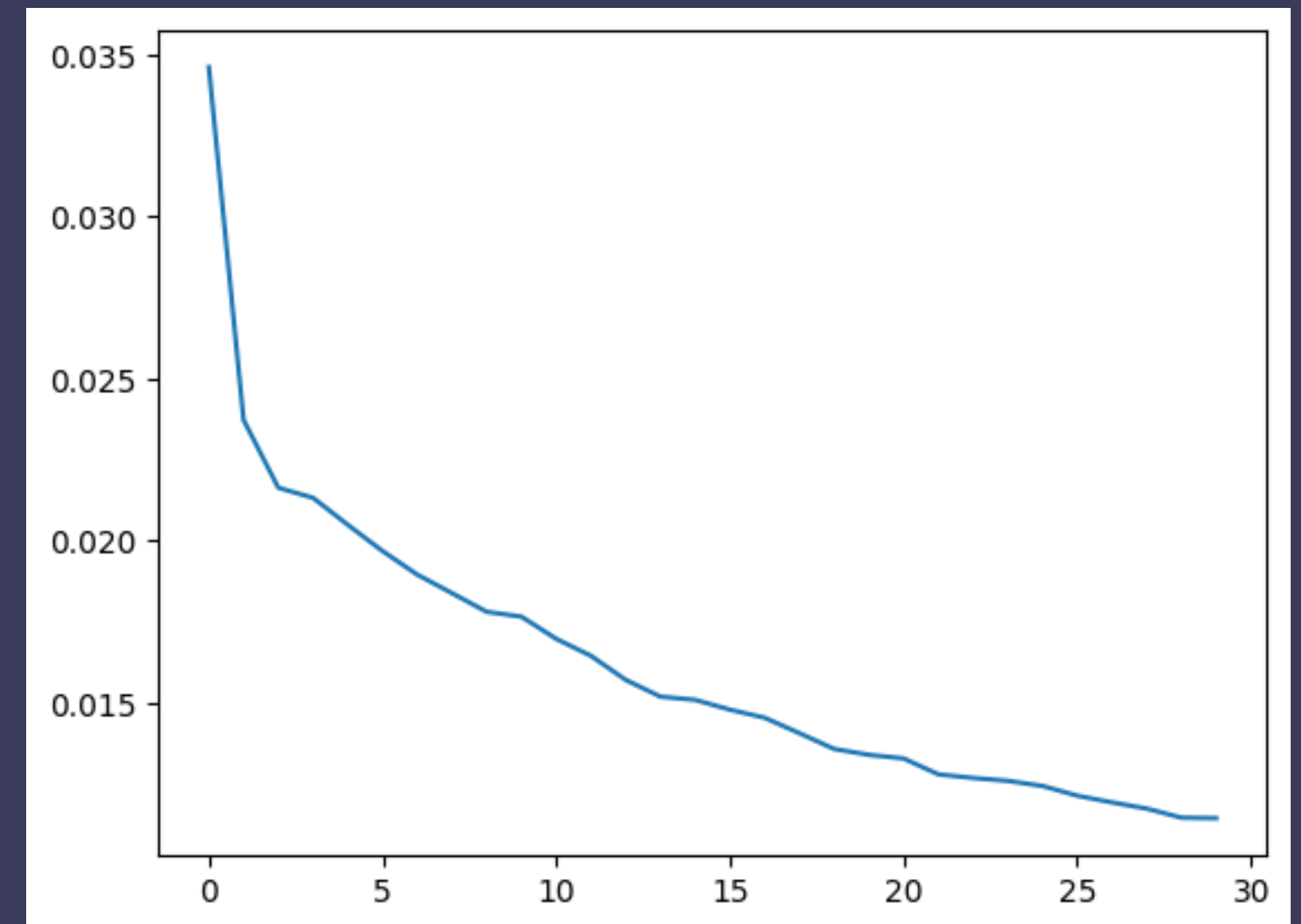
Harmonique



# III. Méthodes d'identification:

## Regression logistique multiple

- On effectue une PCA sur le signal temporel pour réduire les dimensions du problème
- On effectue une régression logistique sur chacune des catégories (5 en tout)
- Chaque régression prévoit la présence ou l'absence d'un type d'appareil



# III. Méthodes d'identification:

# CNN

Convolution 1D



Max Pooling



Convolution 1D



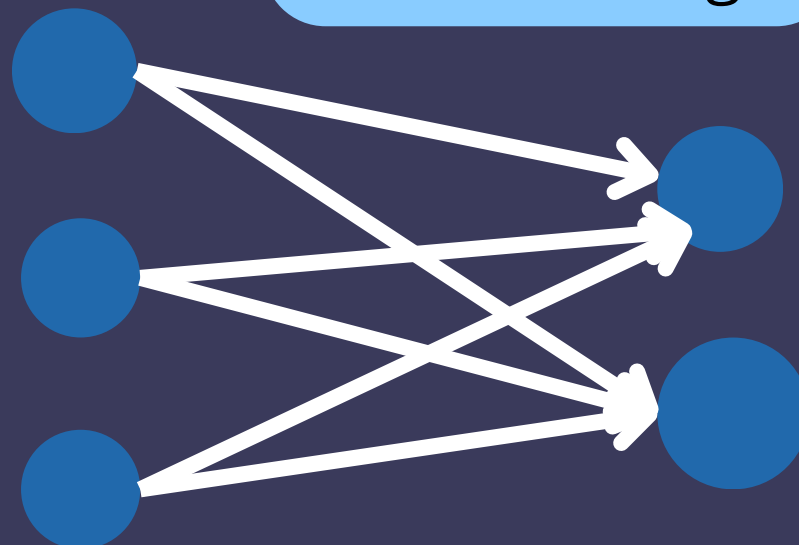
Max Pooling



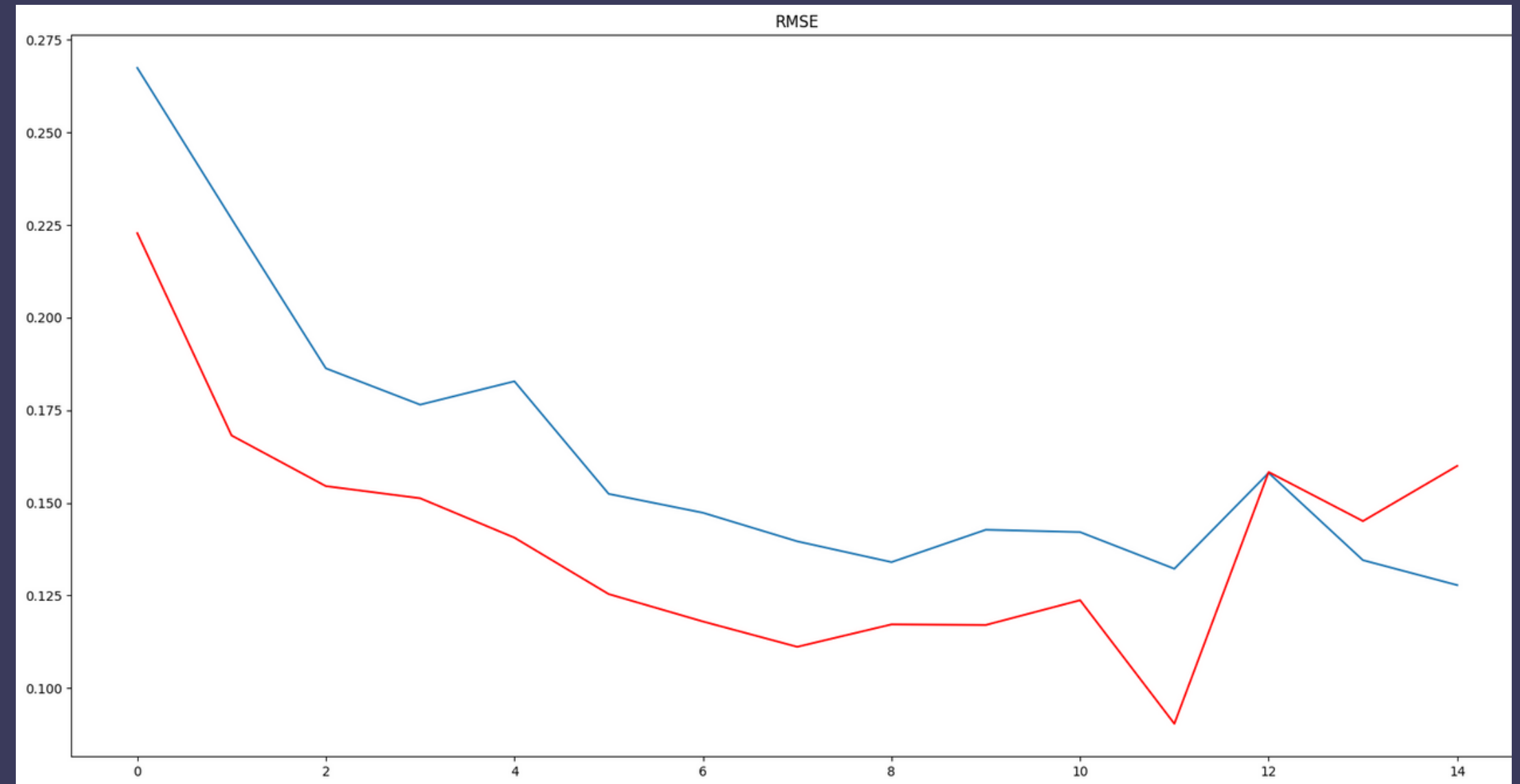
Convolution 1D



Max Pooling



[1 0 0 1 1]



RMSE for training set (red) and test set (blue)



# III. Méthodes d'identification:

# Support Vector Machine

- On effectue une extraction des harmoniques des signaux pour réduire la dimension du problème.
- Ensuite on utilise un algorithme de support vector machine sur ces données, nous permettant de détecter les divers appareils.

# IV. Conclusion:

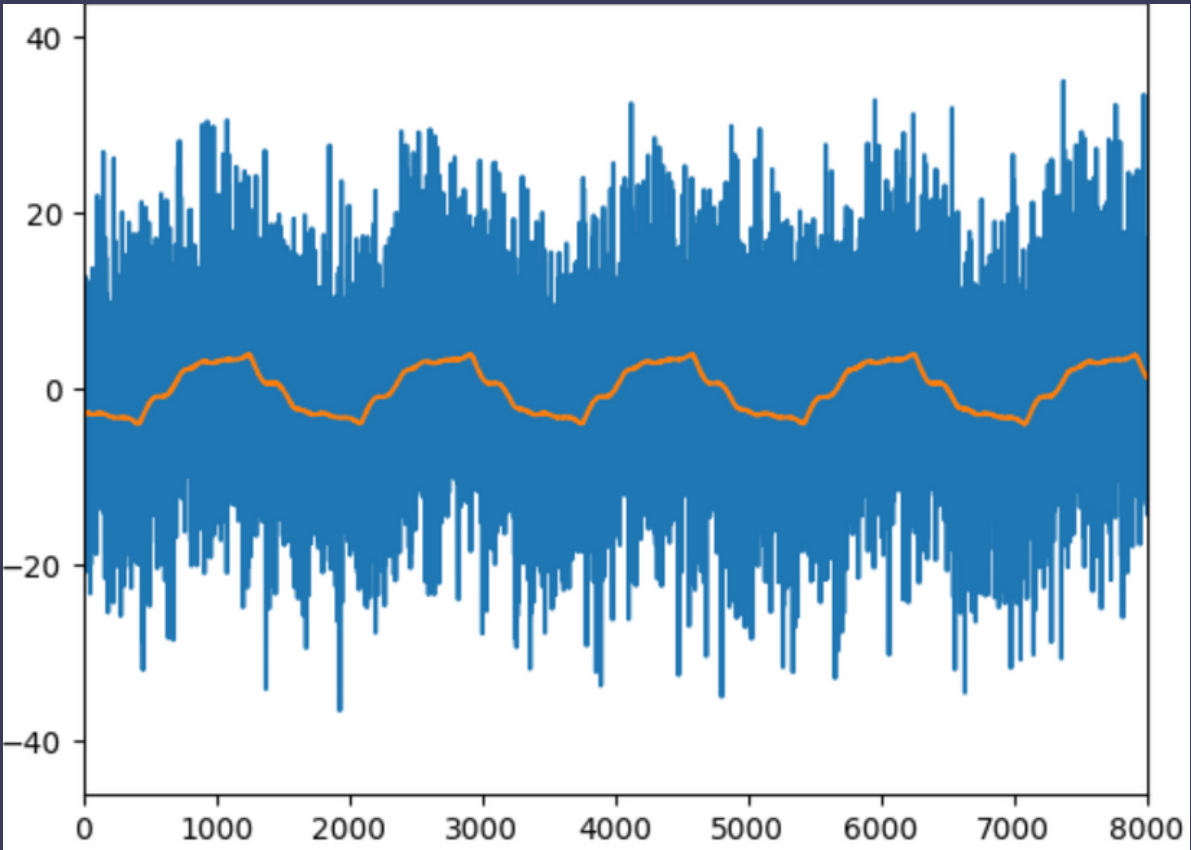
## Meilleurs résultats:

	feature	f1 score	accuracy
SVM	Time signal	0.85	0.87
	Harmonics	0.97	0.97
	Spectrum	0.55	0.55
CNN	Time signal	0.97	0.97
Multi Reg	Time signal + PCA (n=5)	0.65	0.74

# IV. Conclusion:

## Robustesse :

	feature	f1 score	accuracy
SVM	Time signal	0.55	0.54
	Harmonics	0.68	0.68
CNN	Time signal	0.7	0.7
Multi Reg	Time signal + PCA (n=5)	0.	0.5



## IV. Conclusion:

### **Pistes d'améliorations:**

- Augmenter le jeu de données
- Validation et tests avec jeu de données réaliste
- Combinaison de modèles