



# Deep Learning pour l'analyse des données de consommation et environnement d'expérimentation

Georges HEBRAIL

EDF R&D et IRT SystemX

30 Novembre 2017

4<sup>ème</sup> journée scientifique de l'Institut RISEGrid

Centrale-Supelec

*Nota : cette présentation comprend quelques planches empruntées à G.Germaine et S.Lamarche*



# SOMMAIRE

1. **DONNÉES DE CONSOMMATION ÉLECTRIQUE**
2. **APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION**
  - CONSTRUCTION DE PROFILS
  - DÉCOMPOSITION PAR USAGES
3. **ENVIRONNEMENT D'EXPÉRIMENTATION**
4. **CONCLUSION**

# DONNÉES DE CONSOMMATION ÉLECTRIQUE

## ■ Des données de plus en plus précises

### □ Avant Linky

- Données semestrielles ou mensuelles
- Panels au pas 10 minutes

### □ Avec Linky

- Données journalières
- Données 10' ou 30'
- Disponibilité J+1

### □ Avec l'ERL

- Données 1" à 2"
- Disponibles en temps réel

## ■ Différents niveaux d'agrégation

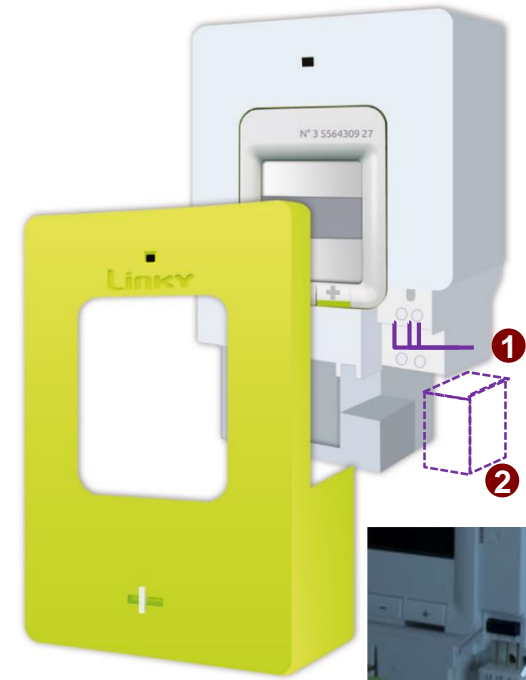
- **Consommation individuelle**
- Consommation agrégée : nœud du réseau, portefeuille de clients
- Consommation nationale



# DONNÉES DE CONSOMMATION ÉLECTRIQUE

## ■ Des données de plus en plus précises

- Avant Linky
  - Données semestrielles ou mensuelles
  - Panels au pas 10 minutes
- Avec Linky
  - Données journalières
  - Données 10' ou 30'
  - Disponibilité J+1
- Avec l'ERL
  - Données 1" à 2"
  - Disponibles en temps réel



Space dedicated to ERL

ERL (Radio Transmitter Local):  
Wireless communication



## ■ Différents niveaux d'agrégation

- **Consommation individuelle**
- Consommation agrégée : nœud du réseau, portefeuille de clients
- Consommation nationale

# ANALYSE DE DONNÉES DE CONSOMMATION

- **Analyse exploratoire**

- Comprendre les habitudes de consommation
- **Etablir des profils type de consommation**
- Relier les profils à des secteurs/usages

- **Analyse prédictive**

- Sur des données agrégées : prévision de consommation, ex. à 7 jours
- Sur des données individuelles :
  - Prévision de consommation
  - **Décomposition de la consommation par usages**

- **Quel est l'apport des nouvelles approches de Deep Learning ?**

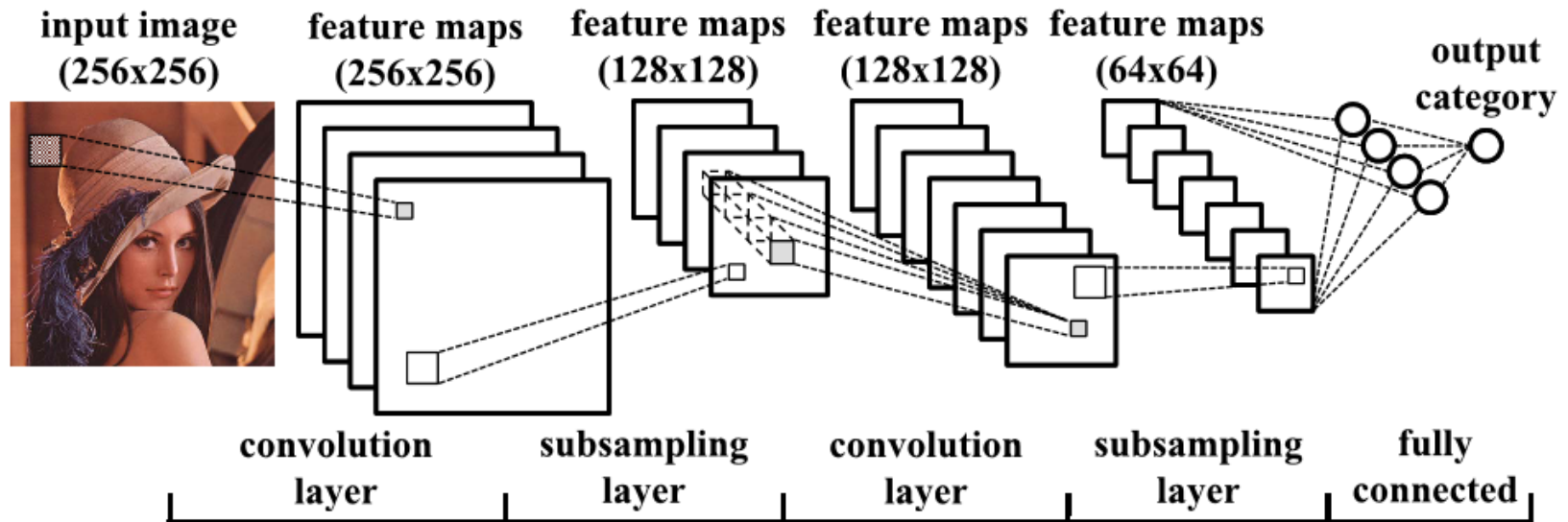


# SOMMAIRE

1. DONNÉES DE CONSOMMATION ÉLECTRIQUE
2. **APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION**
  - CONSTRUCTION DE PROFILS
  - DÉCOMPOSITION PAR USAGES
3. ENVIRONNEMENT D'EXPÉRIMENTATION
4. CONCLUSION

# DEEP LEARNING : RESEAUX CONVOLUTIFS

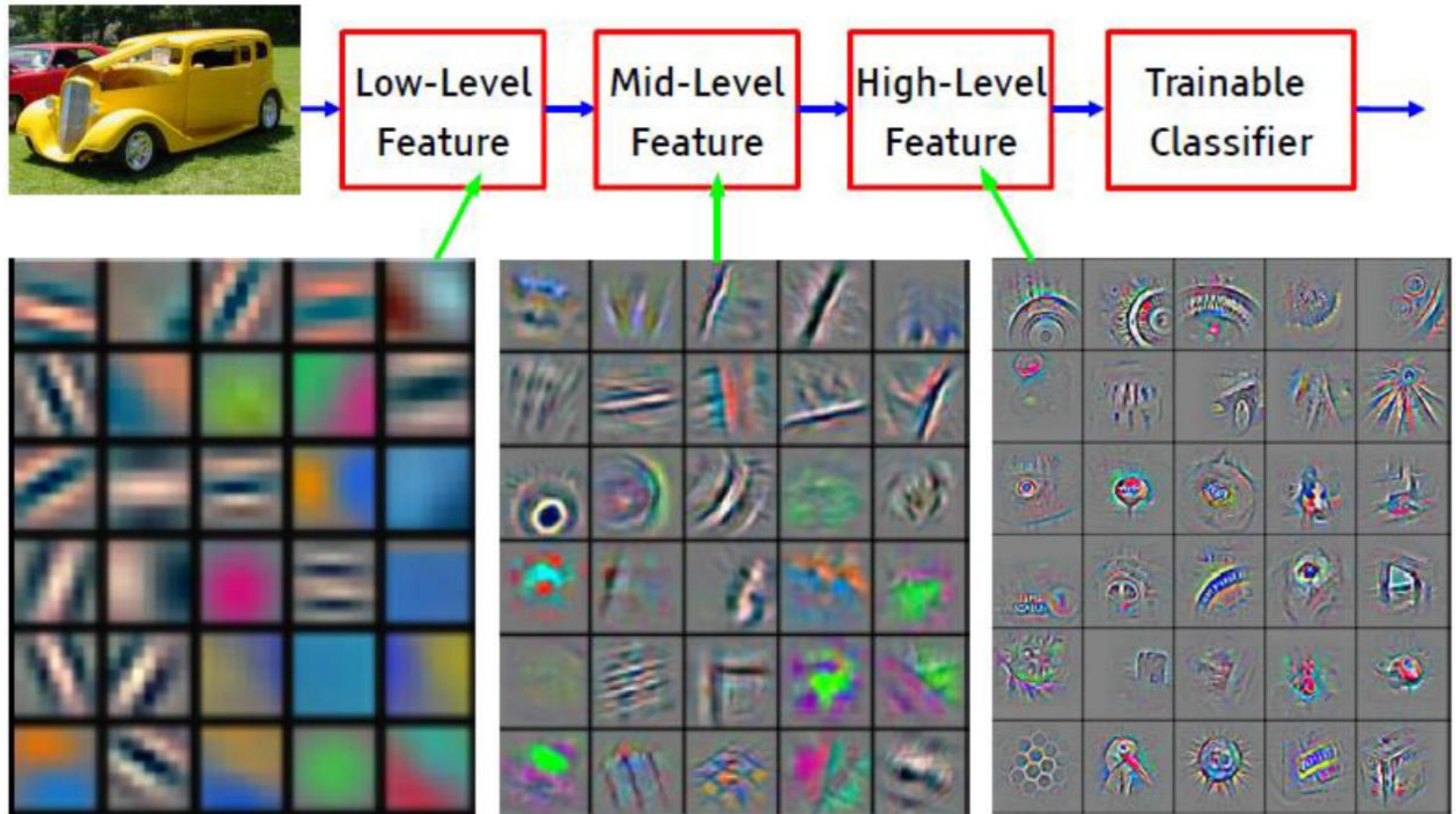
- L'intérêt des réseaux convolutifs est d'extraire des descripteurs en appliquant des transformations non linéaires (convolutions + fonctions d'activations) successives sur les données brutes
- Le réseau « apprend » par lui-même les *features* les plus discriminantes pour le problème à traiter



Source: Y.LeCun



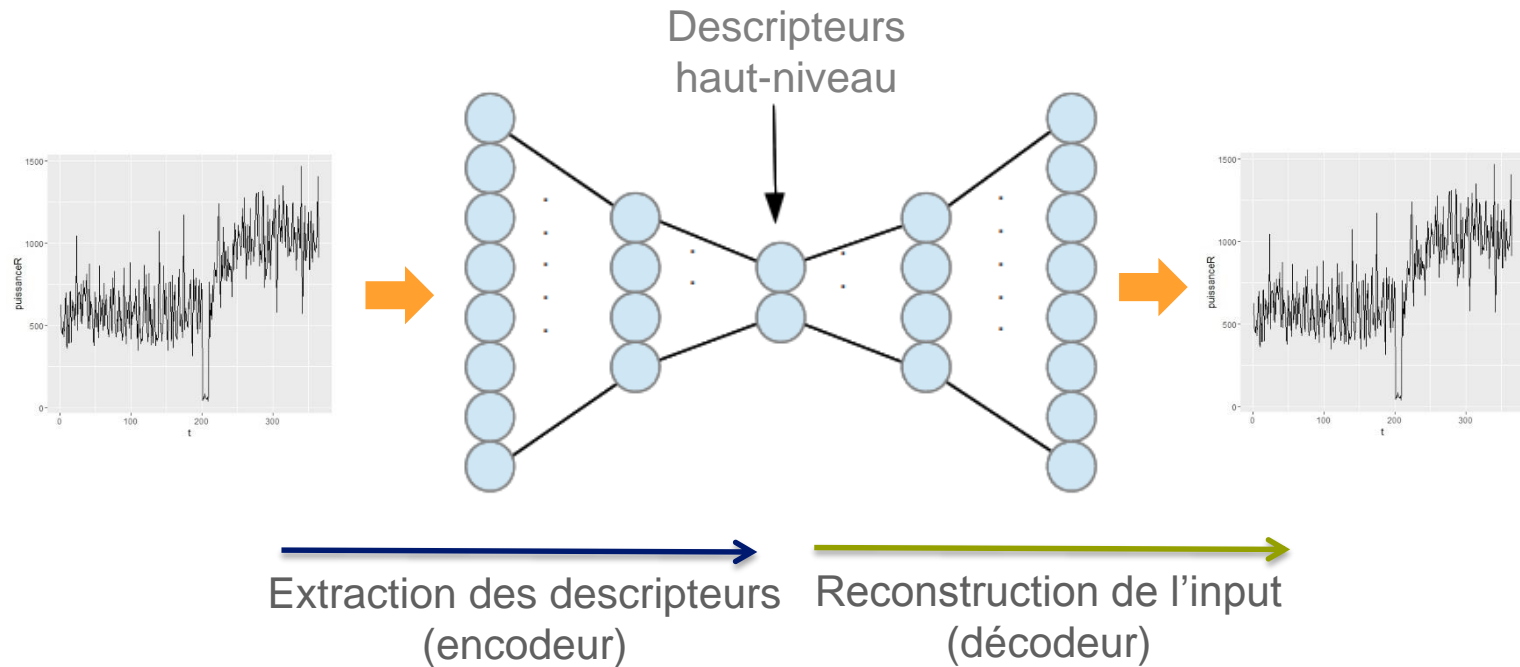
# DEEP LEARNING : RESEAUX CONVOLUTIFS



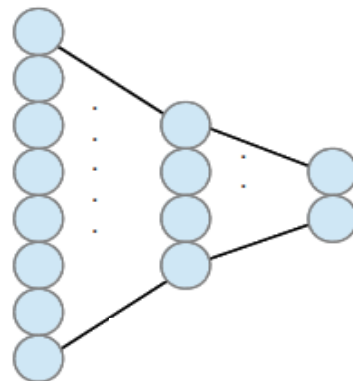
Source: Y.LeCun



# DEEP LEARNING : AUTO-ENCODEURS



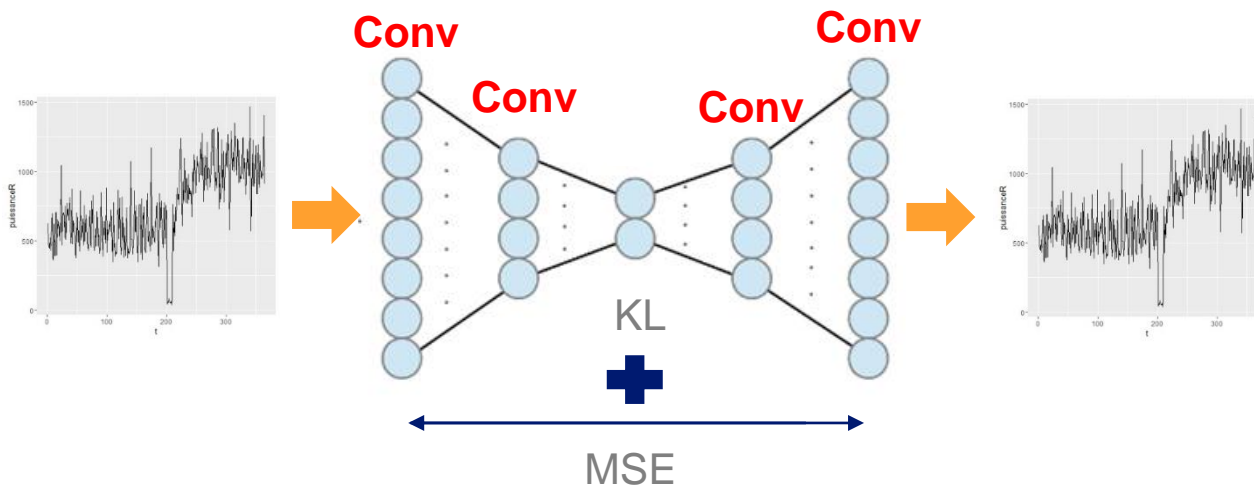
L'auto-encodeur est  
« coupé » en son milieu,  
pour ne garder que la partie  
« descripteurs »



*Features* caractérisant les  
courbes de consommation

# AUTO-ENCODEURS CONVOLUTIFS

- Autoencoder composé de couches convolutives, avec downsampling à chaque couche (maxpooling)
- Un autoencoder variationnel possède une double fonction de coût : la distance entre le vecteur d'entrée et de sortie (MSE), et une pénalisation forçant la distribution des variables latentes à se rapprocher d'une gaussienne (KL)



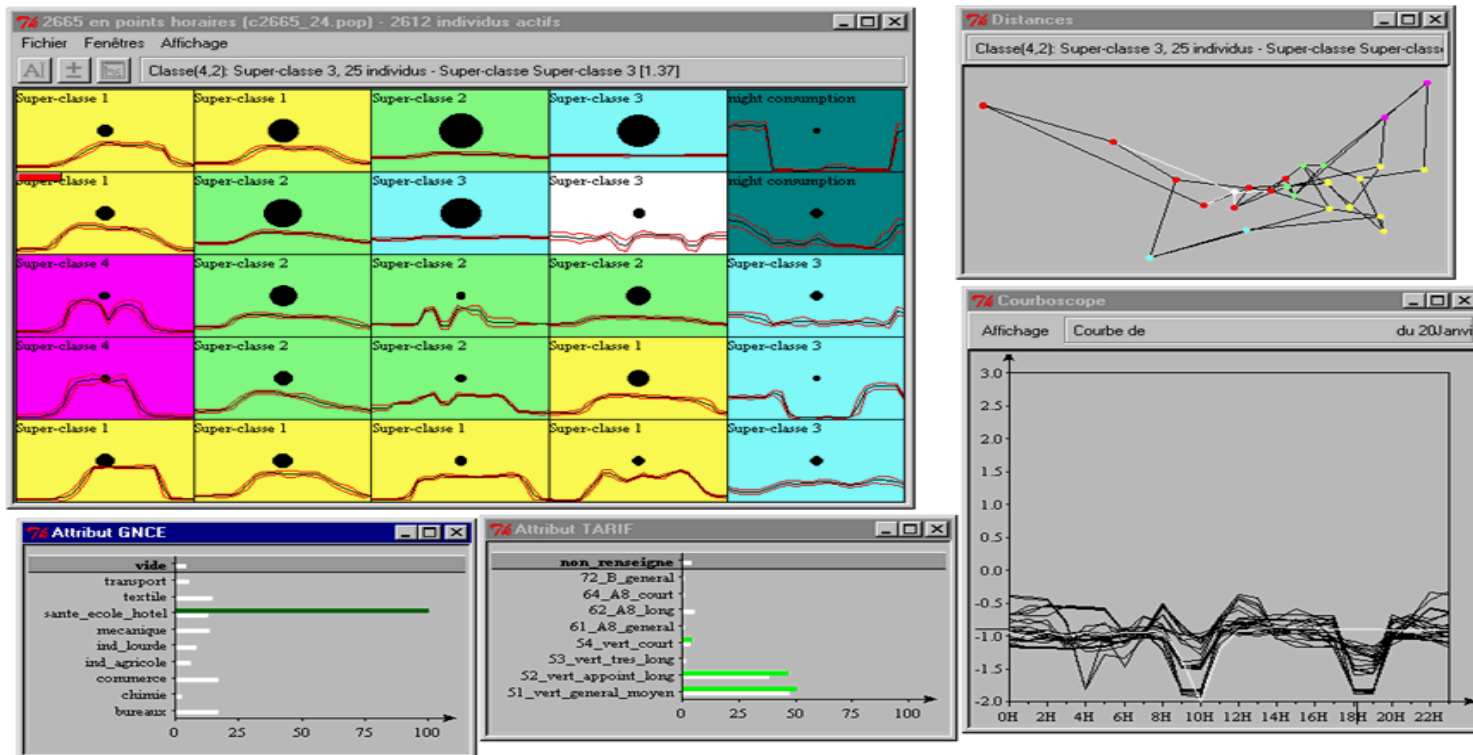
# SOMMAIRE

1. DONNÉES DE CONSOMMATION ÉLECTRIQUE
2. APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION  
CONSTRUCTION DE PROFILS  
DÉCOMPOSITION PAR USAGES
3. ENVIRONNEMENT D'EXPÉRIMENTATION
4. CONCLUSION

# CLUSTERING DE CONSOMMATIONS INDIVIDUELLES

## ■ Approche classique

- Méthode de clustering (k-means, modèle de mélange, Kohonen, ...)
- Prétraitement des courbes : normalisation, ACP, ...



# CLUSTERING DE CONSOMMATIONS INDIVIDUELLES

- **Nouvelles approches**

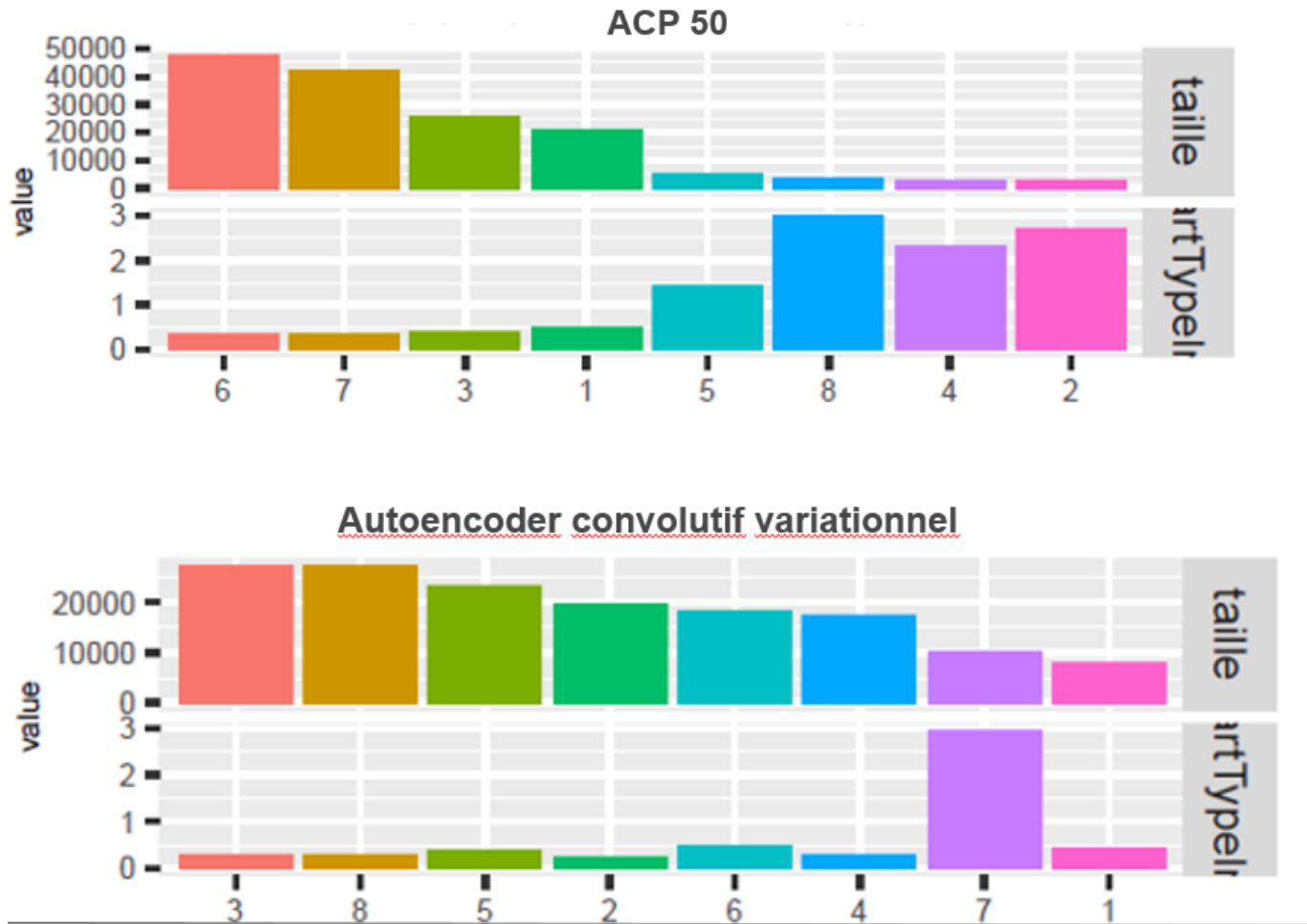
- Utilisation d'autres distances : ex. DTW, Teeraratkul et al. 2017
- Utilisation de features obtenus à l'aide d'un auto-encoder variationnel

- **Premières expérimentations**

- Jeu de données de 150K courbes à pas journalier sur 1 an (365 jours)
- Apprentissage d'un auto-encoder variationnel avec typiquement un codage à 20-50 neurones
- Comparaison du clustering avec données brutes, ACP à 50 facteurs

→ **Formation de classes beaucoup plus homogènes et d'une classe 'poubelle'**

# CLUSTERING DE CONSOMMATIONS INDIVIDUELLES





# SOMMAIRE

1. DONNÉES DE CONSOMMATION ÉLECTRIQUE
2. APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION  
CONSTRUCTION DE PROFILS  
DÉCOMPOSITION PAR USAGES
3. ENVIRONNEMENT D'EXPÉRIMENTATION
4. CONCLUSION

# DÉCOMPOSITION DE LA CONSOMMATION PAR USAGES

- **Décomposition par usages**

- Désagrégation de la courbe de charge (consommation)
- NILM pour « *Non-Intrusive Load Monitoring* »
- NIALM pour « *Non Intrusive Appliance Load Monitoring* »

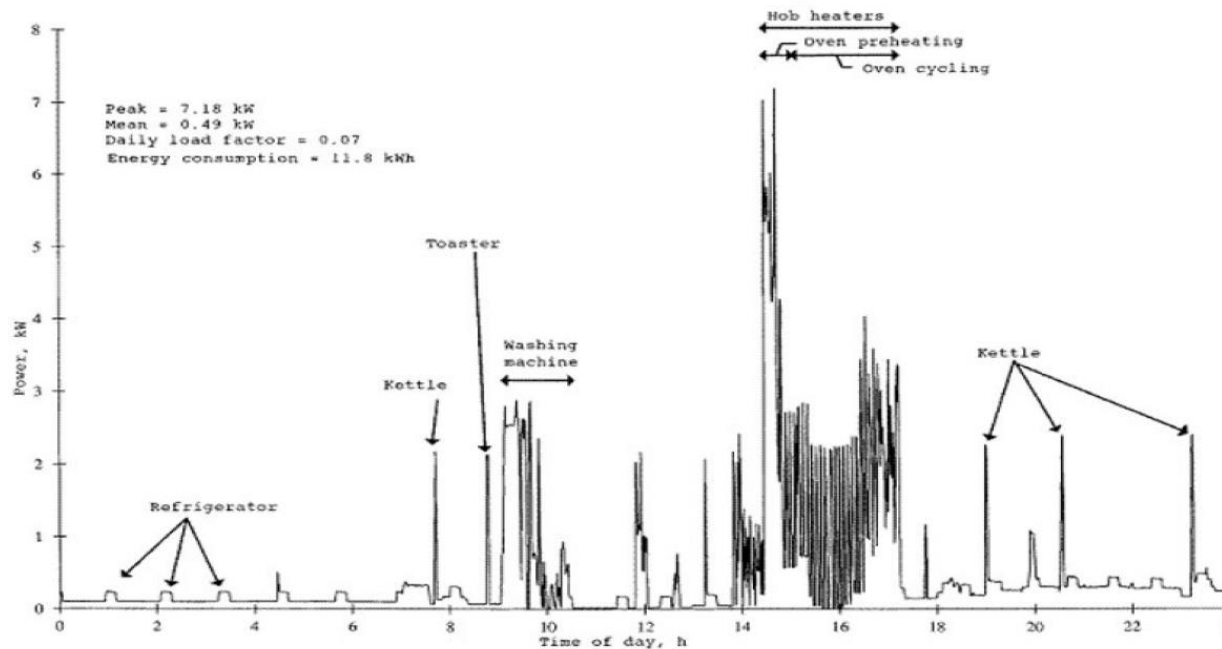
## DECOUVRIR LES EQUIPEMENTS ENERGIVORES



# DÉCOMPOSITION DE LA CONSOMMATION PAR USAGES

## ■ Grande variété des problèmes

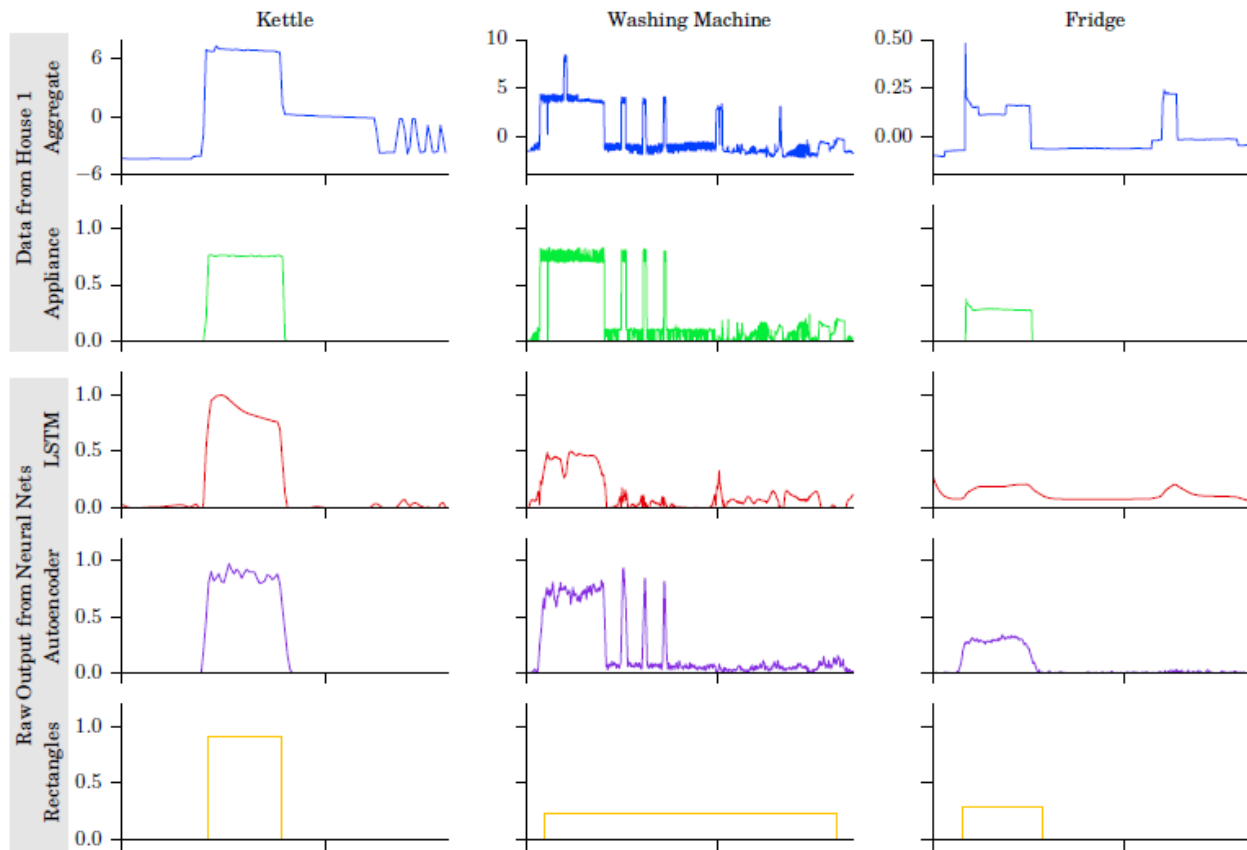
- Détection de la présence d'un type d'appareil dans un local
- Détection du fonctionnement d'un appareil au moins une fois dans une journée
- ...
- Retrouver la consommation à chaque instant de chaque appareil
- Pas de temps des données en entrée
  - Annuel, mensuel, journalier, 30', 10', 1'', 1KHz, 1MHz



Household electrical consumption example

# DÉCOMPOSITION DE LA CONSOMMATION PAR USAGES

- Travaux de **Jack Kelly** : "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes",  
Jack Kelly, William Knottenbelt, <https://arxiv.org/abs/1404.0284>, 2014

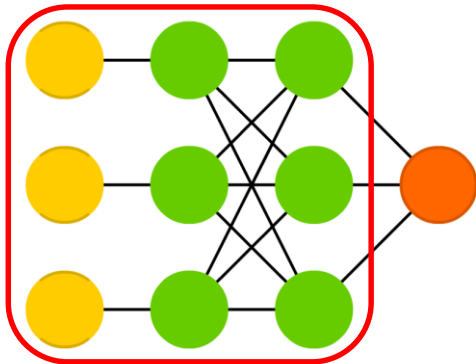


# DÉCOMPOSITION DE LA CONSOMMATION PAR USAGES

Architectures de réseaux expérimentées par Jack Kelly :

**Input** : courbe de charge globale,  
échantillonnée au pas de temps 6s

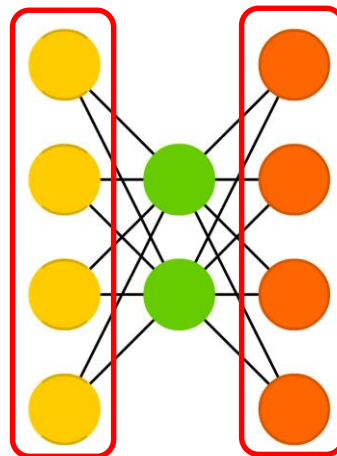
Régression (rectangle)



**Filtres convolutifs**

**Output** : timestamp de début, timestamp de fin et puissance moyenne de l'appareil cible

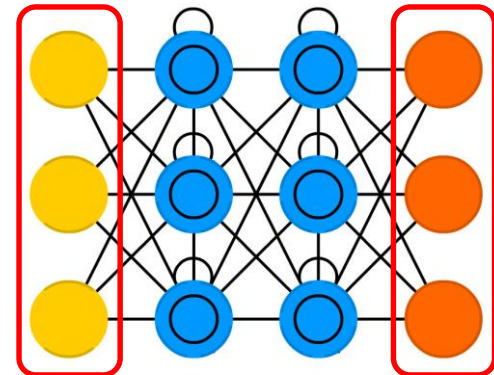
Auto-encoder



**Filtres convolutifs**

**Output** : courbe de charge de l'appareil cible

LSTM (réseau récurrent)



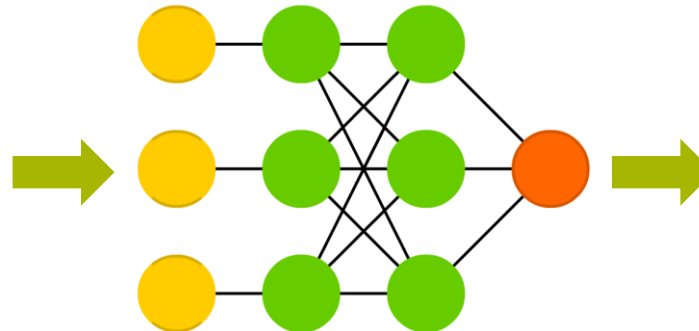
**Filtres convolutifs**

**Output** : courbe de charge de l'appareil cible

# DÉCOMPOSITION DE LA CONSOMMATION PAR USAGES

- Travaux réalisés en 2017 avec Centrale-Supelec Metz
- Un **réseau de neurones convolutif** par appareil
- Evaluation sur le **centre d'une fenêtre temporelle glissante** de la **présence ou l'absence de l'appareil électrique recherché** (cycle complet de déclenchement de l'appareil électrique cible)
- Deux appareils testés : le **four** et le **lave-vaisselle**
  - Pour le four, le taux de reconnaissance (AUC) est supérieur à 80% ; sur un même cas d'application, les performances obtenues sont équivalentes à celles d'une startup spécialisée en désagrégation
  - Les performances obtenues pour le lave-vaisselle sont moins bonnes (entre 60 et 70%) ; cela est principalement lié à la complexité des signatures de ce type d'appareils.

Input : courbe de charge globale, échantillonnée au pas de temps 1s



1. détection de la présence ou l'absence de l'appareil cible (classification)
2. Si appareil détecté, estimation de la puissance consommée de l'appareil (régression)



# SOMMAIRE

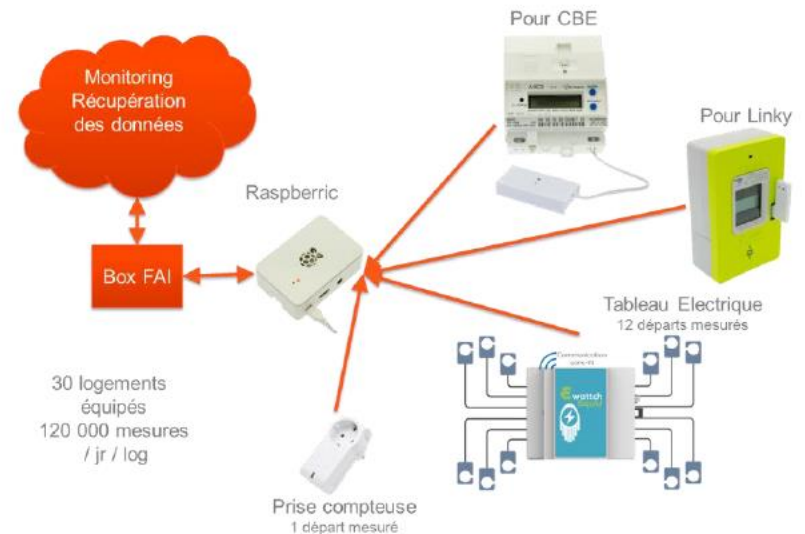
1. DONNÉES DE CONSOMMATION ÉLECTRIQUE
2. APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION  
CONSTRUCTION DE PROFILS  
DÉCOMPOSITION PAR USAGES
3. **ENVIRONNEMENT D'EXPÉRIMENTATION**
4. CONCLUSION

# ENVIRONNEMENT EXPÉRIMENTAL

- Nécessité de disposer de données réelles pour l'évaluation

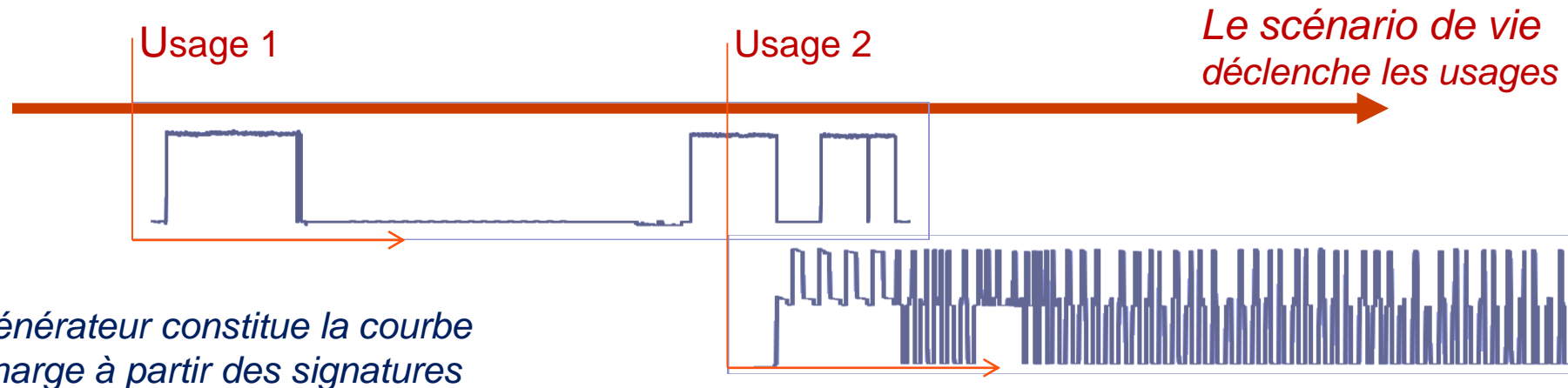
Instrumentation de 17 maisons au pas seconde, avec du sous-comptage par usage (les usages varient selon les maisons).

- Lave-Linge
- Lave-Vaisselle
- Sèche-Linge
- Four
- Plaque de cuisson
- Réfrigérateur
- ECS



# ENVIRONNEMENT EXPÉRIMENTAL

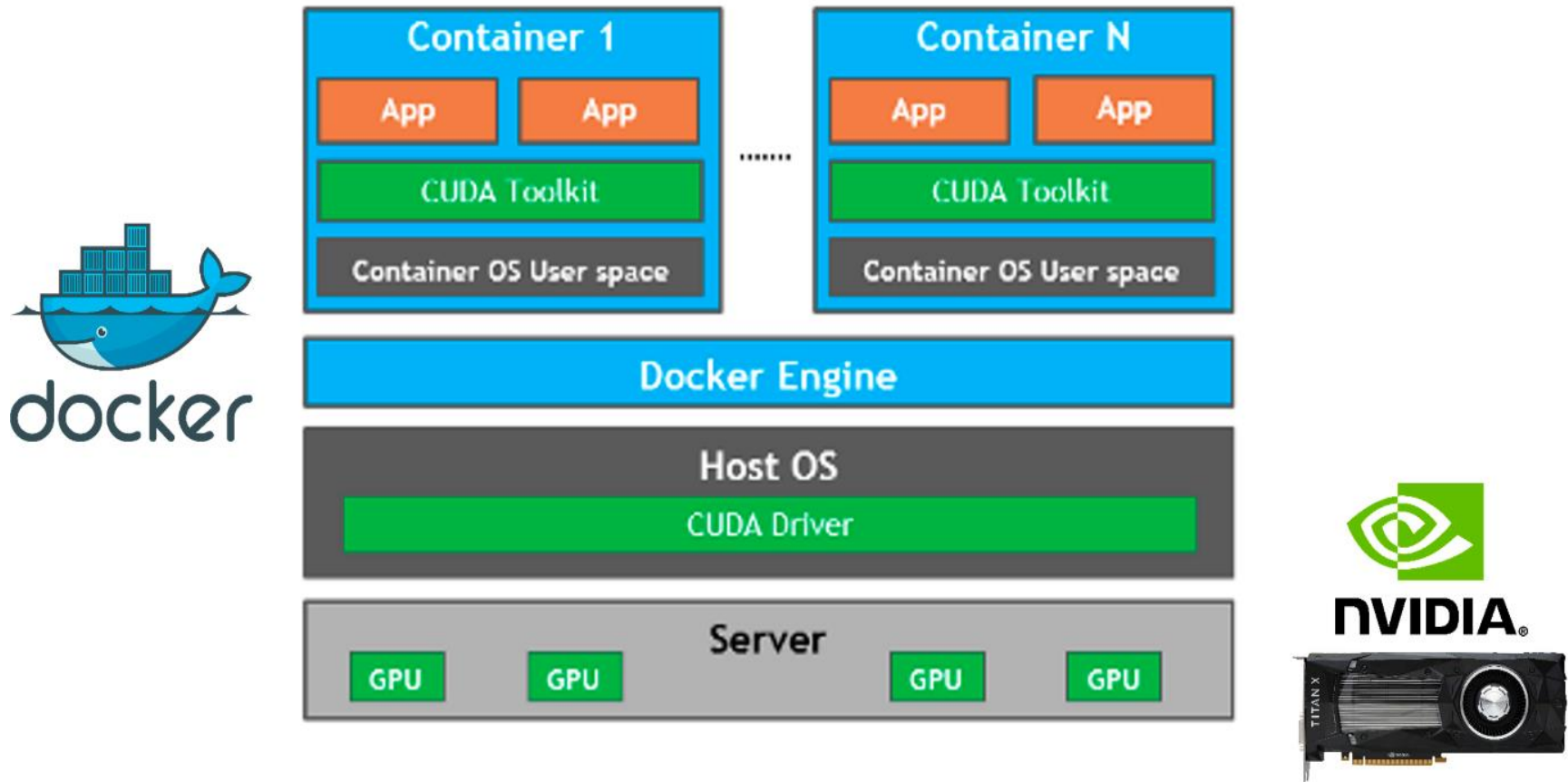
- Nécessité de disposer de données nombreuses pour **l'apprentissage**
- Générer des courbes de charge réalistes
  - Basées sur une simulation de comportements humains (enquête emploi du temps)
  - Utilisant des signatures d'appareil au pas seconde



- Générer des courbes d'apprentissage intensif
  - Succession rapide et superposition de signatures

# ARCHITECTURE TECHNIQUE

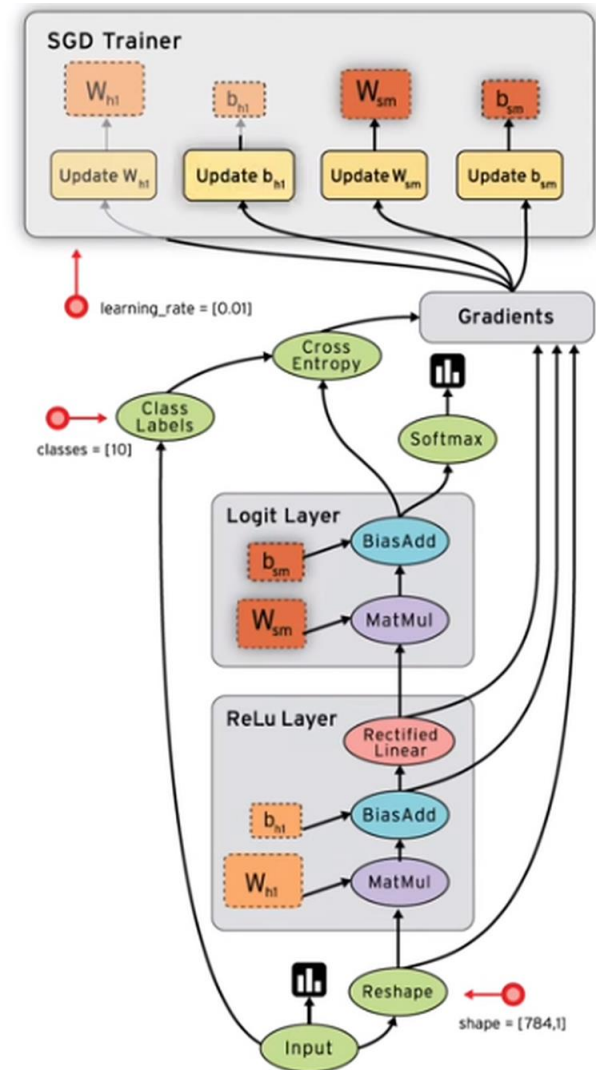
Environnement de travail déployé avec Docker, calculs effectués sur GPU Nvidia



# ARCHITECTURE TECHNIQUE



- **Framework de calcul numérique** développé par Google
- Flux de données représentés **sous forme de graphes**
- **node** = opérateur mathématique
- **edge** = tableau multidimensionnel (tensor)
- **Hardware agnostique** (CPU, GPU, TPU...)
- **API Python et C++**



# ARCHITECTURE TECHNIQUE



- **Librairie high-level** de réseaux de neurones
- Langage **Python**
- Fonctionne avec **Theano** ou **Tensorflow**
- Offre un grand nombre de briques « sur étagère » pour des **prototypages rapides**
- Offre un **accès direct à Theano ou Tensorflow**, permettant la création d'architectures customisées



# SOMMAIRE

1. DONNÉES DE CONSOMMATION ÉLECTRIQUE
2. APPROCHES DE DEEP LEARNING SUR DONNÉES DE CONSOMMATION  
CONSTRUCTION DE PROFILS  
DÉCOMPOSITION PAR USAGES
3. ENVIRONNEMENT D'EXPÉRIMENTATION
4. **CONCLUSION**

# CONCLUSION

- **Premiers travaux sur le Deep Learning pour les données de consommation**
  - Construction de profils par clustering
  - Décomposition par usages
  - ... mais aussi prévision de consommation, ...
  
- **On bénéficie sur les séries temporelles des travaux sur les images**
  - Simplification des approches sur les images
  - Des outils open source à disposition
  
- **Mais ...**
  - Une mise en œuvre technique assez complexe
  - Un protocole expérimental rigoureux à mettre en place