

# Le machine learning : un atout pour une meilleure efficacité

## Applications à la gestion énergétique des bâtiments

Comment gérer de manière intelligente les consommations et productions d'énergie dans les bâtiments ? Les solutions à ce problème complexe pourraient venir du monde de l'apprentissage automatique ou « machine learning ». Celui-ci permet la mise au point d'algorithmes de contrôle avancés visant simultanément la réduction de la consommation d'énergie, l'amélioration du confort de l'utilisateur et l'adaptation à ses besoins.

**Antonio Ridi, Christophe Gisler, Jean Hennebert**

Dans les pays développés, les bâtiments représentent l'une des principales sources de consommation énergétique. En Europe et aux USA, ils sont à eux seuls responsables de 20 à 40% de la consommation énergétique totale, soit un pourcentage supérieur à ceux de l'industrie et du transport.

Dans un bâtiment, la moitié de la consommation provient du triplet chauffage, ventilation et climatisation (CVC). Viennent ensuite l'éclairage et les autres appareils à fonctionnement électrique. Dans les bureaux, le CVC, l'éclairage et les appareils électriques atteignent ensemble environ 85% de la consommation totale d'énergie [1]. Fait intéressant, les systèmes CVC représentent un quart de la consommation d'énergie totale dans les pays développés et, malgré l'amélioration générale des propriétés physiques des bâtiments, une augmentation de cette valeur est à prévoir. Ceci est principalement dû à l'augmentation de la population et à une demande croissante en termes de confort et de services des bâtiments.

Au cœur du problème ne se trouve pas uniquement la physique du bâtiment. Il faut également tenir compte du comportement des utilisateurs. En effet, l'efficacité énergétique des bâtiments peut se définir comme le rapport entre les consommations d'énergie et le confort des usagers. Si l'énergie produite et consommée peut se mesurer à travers différents capteurs distribués dans le bâtiment, le confort est par contre une notion beaucoup plus floue, dépendante des individus et de leurs activités. Dans ce

contexte, il faut aller plus loin que les systèmes actuels basés sur des consignes de températures, sur des minuteries ou encore sur une modélisation des propriétés physiques du bâtiment en fonction de la météo. Le but est donc d'offrir la possibilité de contrôler de manière automatique et intelligente les services d'un bâtiment (figure 1) en considérant son utilisation réelle et pas uniquement ses propriétés physiques.

L'utilisation d'un bâtiment, le comportement des utilisateurs et la notion de confort sont des événements de

nature stochastique avec des biais contextuels et individuels. Dans ce cadre, les systèmes auto-apprenants de type « machine learning » présentent un réel avantage.

Cet article résume dans un premier temps les principes et méthodologies du machine learning. Puis, différentes applications sont présentées, ainsi que divers projets de recherche dont le but est d'analyser la pertinence des approches « machine learning ». Certaines simplifications de ces approches qui ont déjà conduit à des solutions commercialisées sur le marché sont rapportées en conclusion.

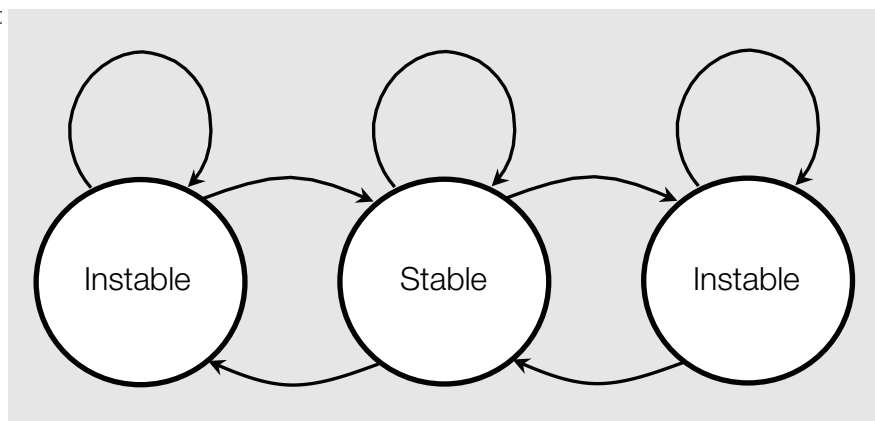
### Principes du machine learning

De façon simplifiée, le machine learning utilise des approches se basant sur la découverte d'équations mathématiques ne provenant pas de processus physiques, mais directement de l'analyse des données. Dans le cas des bâtiments, les données représentent typiquement des séries temporelles provenant de capteurs mesurant par exemple la tempéra-



**Figure 1** Les bâtiments intelligents peuvent être composés d'une grande variété de capteurs et d'actionneurs qui communiquent avec un système de contrôle central.

EIA-FR [5]



**Figure 2** Exemple de modélisation basée sur les modèles de Markov cachés pour la reconnaissance d'activité avec trois états. Une telle topologie est utilisée pour modéliser la détente, le travail, le repos, l'absence de la maison, le passage du lit aux toilettes.

ture d'une pièce, la présence de personnes ou encore l'ouverture de portes et fenêtres.

Le principe consiste, d'une part, à choisir des modèles mathématiques suffisamment riches pour représenter chacun des processus sous-jacents et, de l'autre, à découvrir les paramètres de ces modèles qui expliquent au mieux les données observées. Le terme général de « machine learning » englobe des terminologies liées, voire équivalentes, comme les approches basées sur les données (data-driven), les algorithmes à auto-apprentissage ou encore l'analyse intelligente des données.

Dans le cadre du bâtiment, les approches de type « machine learning » ont fait l'objet de recherches récentes et de nombreux projets sont référencés dans la littérature [3]. Leurs avantages sont multiples :

- Une compréhension des processus physiques sous-jacents n'est pas forcément nécessaire.
- Une configuration nouvelle d'un bâtiment peut être traitée dès que les données d'observations sont disponibles, ce qui rend plus facile l'installation et la maintenance de ces systèmes de contrôle auto-apprenants.
- Ces modèles sont en mesure de capturer non seulement la physique du bâtiment, mais aussi le comportement stochastique des utilisateurs ou des notions plus floues comme le confort.

Cependant ces approches présentent également quelques désavantages, notamment le besoin général d'une grande quantité de données pour construire les estimateurs, ce qui implique un temps de mise en service relativement long.

Trois sous-domaines de machine learning peuvent être mentionnés :

- les problèmes de « prédiction » dans lesquels l'objectif consiste à calculer un estimateur d'une valeur physique dans le futur, et ce, en fonction des valeurs passées et actuelles ;
- les problèmes de « classification » dont le but est de calculer à partir d'un ensemble d'observations la probabilité d'appartenance du système à une catégorie donnée (état, classe) ;
- le « clustering » (partitionnement de données) dont l'objectif est de découvrir des cas similaires dans un vaste ensemble d'observations.

Ces deux premiers sous-domaines sont illustrés ci-après par différentes applications dans le cadre du bâtiment.

### Reconnaissance d'activités et prédiction de l'occupation

La reconnaissance automatique des activités en utilisant des données temporelles acquises à partir de capteurs de présence a fait l'objet de différents travaux (voir par exemple les références [4,5]). Ridi et al. [5] ont utilisé des « modèles de Markov cachés » pour modéliser les activités. Dans cette approche, les lieux d'activités sont associés aux états du modèle en distinguant les zones de transition (couloirs, portes) des zones de stabilité (local, partie d'une pièce) pour lesquelles les statistiques des observations remontant des capteurs sont très différentes (figure 2). Après entraînement, le système a modélisé les fonctions de densité de probabilité d'émission des observations dans les différents états, ainsi que les probabilités de transiter d'un lieu d'activité à un autre. En phase d'utilisation, le modèle permet de découvrir les activités et donc de labéliser l'usage qui est fait des différentes pièces du bâtiment en fonction du temps.

Lorsque le nombre de capteurs de présence est suffisant, des performances d'identification des activités supérieures à 95% ont été rapportées.

L'étape qui suit la reconnaissance d'activités humaines dans les pièces est l'élaboration de modèles statistiques permettant de prédire la probabilité de présence dans les pièces en fonction de l'heure, du jour de la semaine et, de façon générale, du calendrier. Cette prédiction permet le contrôle intelligent du chauffage en réduisant les consommations et en préservant le confort. Des essais pilotes ont démontré le potentiel et la viabilité commerciale de telles approches, comme dans le cadre du projet PreHeat mené par Microsoft Research [6].

### Classification de signatures électriques

Des approches de type « machine learning » s'appliquent également à l'analyse automatique de la charge électrique dans les bâtiments. Le système décrit dans les références [7-9] utilise des capteurs embarqués dans les prises qui mesurent la consommation toutes les 10 secondes afin d'identifier les équipements utilisés, comme un téléviseur, une machine à laver ou un réfrigérateur. Formellement, le problème est de type « classification » via une estimation des vraisemblances des observations temporelles de consommation électrique étant donné une catégorie donnée d'appareils.

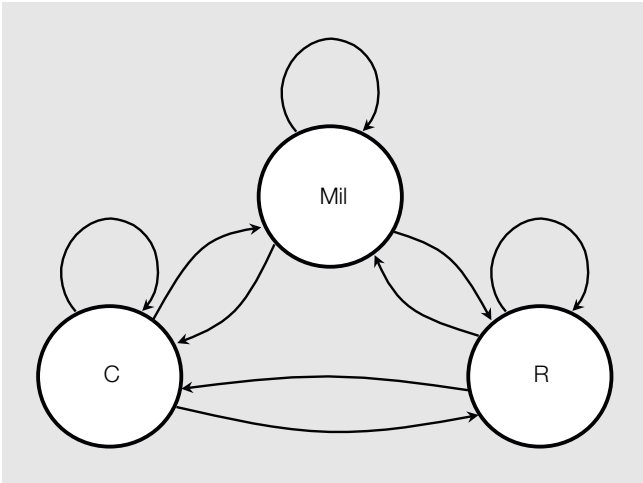
Dans ce travail, les technologies utilisées se basent sur les « mélanges de Gaussiennes » et les modèles de Markov cachés (encadrés) suivant des concepts très similaires aux approches utilisées dans la biométrie et la reconnaissance de la parole. Des performances d'identification d'équipements supérieures à 90%

#### Techniques de modélisation

##### Mélanges de Gaussiennes

La loi normale (ou loi Gaussienne) est généralement bien adaptée pour modéliser les observations faites sur un phénomène naturel en offrant un calcul simple des densités de probabilité. Les mélanges de Gaussiennes permettent de calculer des densités de probabilités en tenant compte d'une somme pondérée de lois Gaussiennes. Les mélanges de Gaussiennes peuvent donc être comprises comme une généralisation permettant de modéliser les observations faites sur un phénomène naturel plus complexes, ayant plusieurs modes de génération des observations.

EIA-FR [5]



**Figure 3** Second exemple de modélisation basée sur les modèles de Markov cachés avec les trois états possibles de la variable ergodique « saison » et les transitions possibles d'un état à un autre : saison de chauffage (C), saison de refroidissement (R) et saison intermédiaire (Mil).

ont été rapportées. Ces technologies permettent d'analyser finement l'utilisation de l'énergie électrique dans les bâtiments et de déterminer indirectement les activités réalisées dans les pièces et leurs taux d'occupation.

D'autres approches se basent sur l'analyse des consommations mesurées globalement au niveau du compteur intelligent avec comme objectif la « désagrégation » de la consommation globale afin de détailler les factures d'électricité [10].

### Modélisation des saisons

Dans cette troisième application, un nouveau modèle de prévision des saisons est introduit avec comme objectif un meilleur contrôle de la production d'énergie. Les approches rapportées dans la référence [5] sont basées sur des modèles à états tels que les modèles de Markov cachés, capables de détecter finement les changements de saisons à partir de données météorologiques. Lors de l'élaboration d'un système de contrôle avancé destiné aux services CVC des bâtiments, la variable « saison » représente une bonne méthode d'adaptation du comportement du système de contrôle aux conditions météorologiques. Une définition simple basée uniquement sur un calendrier ou sur la température extérieure n'est cependant pas satisfaisante car les caractéristiques du bâtiment et son utilisation doivent aussi être considérées.

Afin de modéliser les saisons, les signaux des CVC, l'ouverture des fenêtres, les stores, les températures intérieure et extérieure, ainsi que l'ensoleillement sont considérés. Un module séparé est également utilisé pour calculer la constante de temps du bâtiment (caractérisée par son inertie thermique) incluse en tant que paramètre dans le vecteur d'observations. Avec ce modèle, la

variable « saison » peut prendre les valeurs d'états suivantes (figure 3) :

- « saison de chauffage », qui intervient lorsque le chauffage est nécessaire pour éviter que la température intérieure ne devienne inférieure à la température de confort optimal ;
- « saison de refroidissement », qui intervient lorsque le refroidissement est nécessaire, que ce soit par un refroidissement mécanique ou par des moyens tels que le refroidissement nocturne passif ou la protection solaire passive ;
- « saison intermédiaire » (ou « mi-saison »), qui est la saison la plus critique. Au cours de cette saison, le bâtiment peut avoir besoin tantôt de chauffage, tantôt de refroidissement.

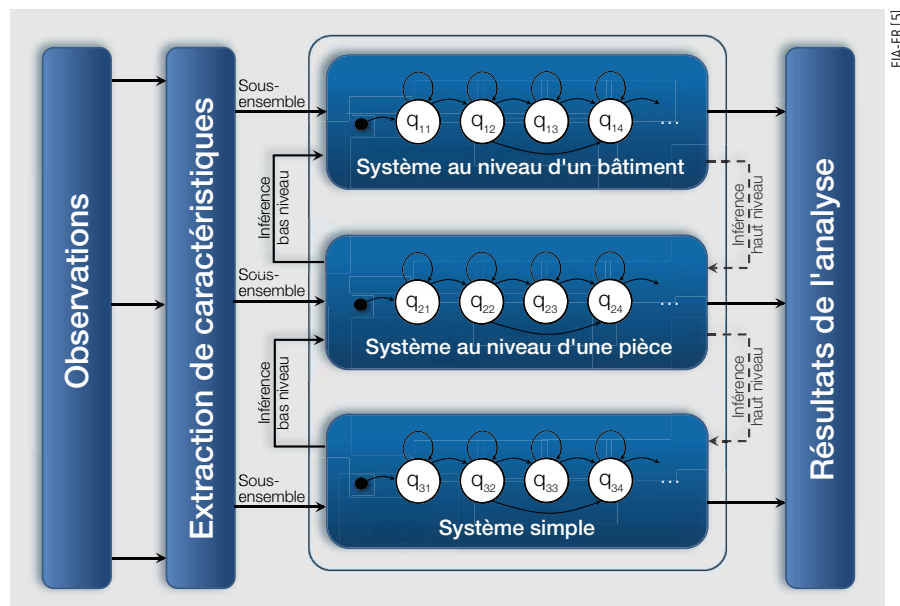
Un essai pilote de cette modélisation est en cours de réalisation dans le cadre du projet Green-Mod. Le pilote a lieu dans le bâtiment expérimental LESO-PB situé sur le campus de l'EPFL.

### Projet Green-Mod

Le but du projet Green-Mod est de modéliser des flux de mesures hétérogènes dans les bâtiments pour réduire les consommations d'énergie tout en préservant le confort humain. Soutenu par la fondation Hasler, il réunit des spécialistes de l'EPFL en physique du bâtiment et des experts de la HES-SO Fribourg en « machine learning » pour apporter des solutions innovantes dans ce contexte.

L'équipe de recherche Green-Mod vise une approche globale et unifiée pour la modélisation des signaux mesurés dans un bâtiment. Comme illustré dans la figure 4, le modèle se base sur différents systèmes de modélisation organisés en couches. Les observations qui proviennent des différents appareils de mesure et capteurs sont en premier lieu enregistrées et traitées à travers un module d'extraction de caractéristiques afin de les rendre exploitables. Les caractéristiques sont ensuite modélisées en tenant compte de leur origine et de l'objectif de modélisation. Les systèmes simples visent la modélisation de signaux typiquement en provenance d'un seul capteur, par exemple la classification des signatures électriques. Les systèmes au niveau d'une pièce visent la modélisation de signaux hétérogènes mesurés dans un local, signaux qui sont potentiellement complétés par les résultats des systèmes simples. De façon similaire et hiérarchique, des systèmes plus globaux, au niveau d'un étage ou d'un bâtiment, peuvent également être utilisés.

L'objectif du projet est de saisir la nature temporelle et spatiale des signaux correspondant aux différents niveaux



**Figure 4** Modélisation basée sur les états.

EIA-FR [5]



d'utilisation d'un bâtiment en incluant, du plus granulaire au plus global, un équipement spécifique, une zone dans une pièce, un local, un étage et le bâtiment dans sa globalité. Largement basée sur l'utilisation de modèles de Markov cachés, l'approche permet de détecter automatiquement les états les plus probables liés aux modes d'utilisation qui, à leur tour, seront utilisés comme entrées pour le paramétrage des algorithmes de contrôle intelligents.

### Perspectives

Les concepts du machine learning font leur apparition dans le monde du bâtiment intelligent avec comme objectif l'élaboration d'algorithmes de contrôle avancés permettant, d'une part, la réduction de la consommation d'énergie et, de l'autre, l'amélioration du confort de l'utilisateur en s'adaptant à ses besoins. Ce domaine fait l'objet d'une activité croissante des équipes de recherche. Il reste encore de nombreux défis à relever, tels que l'optimisation des algorithmes, la validation des modèles à travers des essais pilotes ou la mise au point d'approches moins gourmandes en calculs et en données. Néanmoins, des applications commerciales émergent qui reprennent certains principes simplifiés des technologies présentées dans cet article, telles que Nest [11,12] ou encore Smart Control [13].

### Références

- [1] L. Pérez-Lombard, J. Ortiz and C. Pout: A review on buildings energy consumption information. *Energy and Building*, Vol. 40, pp. 394-398, 2008.
- [2] G. Bovet et J. Hennebert: Le Web des objets à la conquête des bâtiments intelligents. *Bulletin SEV/AES 10s/2012*, pp. 15-18, 2012.
- [3] M. Chan, D. Estève, C. Escriba and E. Campo: A review of smart homes: Present state and

### Techniques de modélisation

#### Modèles de Markov cachés

Les modèles de Markov cachés peuvent également être considérés comme une généralisation des modélisations par mélanges de Gaussiennes ou par loi normale. Ils tiennent compte d'une potentielle évolution temporelle (par états) des processus de génération des observations. Lorsque le processus est dans un état donné, les lois de probabilité d'émission sont supposées être stationnaires. Le processus peut alors transiter vers un autre état où les lois de probabilité d'émission sont différentes.

### Zusammenfassung

#### Maschinelles Lernen: Ein Glücksfall für die

#### Steigerung der Energieeffizienz

#### Anwendungen für das Energiemanagement von Gebäuden

Wie können Energieverbrauch und -erzeugung in Gebäuden auf intelligente Weise gesteuert werden? Lösungen für dieses komplexe Problem könnten aus der Welt des maschinellen Lernens, des «Machine Learning», kommen. Das Machine Learning stützt sich auf Lösungsansätze, die auf der Entdeckung mathematischer Zusammenhänge beruhen, die sich nicht aus physikalischen Prozessen, sondern direkt aus der Datenanalyse ableiten lassen. Bei Gebäuden bestehen letztere typischerweise aus Zeitreihen, die von Sensoren geliefert werden. Diese erfassen zum Beispiel die Temperatur eines Raums, die Anwesenheit von Personen oder auch das Öffnen von Türen und Fenstern.

So ermöglicht das Machine Learning ein Energiemanagement, das zwar auf den physikalischen Eigenschaften des Gebäudes und auf Witterungsverhältnissen beruht, aber auch das Verhalten und den Komfort der Nutzer berücksichtigt – also auf Parametern, die von deren Tätigkeiten abhängen und folglich schwieriger zu ermitteln sind. Daher erfolgt zunächst eine Modellierung der Tätigkeiten der Gebäudenutzer auf Grundlage der von den Anwesenheitssensoren erfassten Daten. So kann die unterschiedliche Nutzung der einzelnen Gebäuderäume in Abhängigkeit von der Zeit registriert werden. Dieser Schritt ermöglicht anschliessend eine Prognose der Anwesenheit von Nutzern in den Räumen in Abhängigkeit von der Uhrzeit und dem Tag, die als Basis für eine intelligente Steuerung der Heizungsanlage genutzt werden kann.

Auf dem Machine Learning basierende Lösungsansätze finden zudem Anwendung bei der Erkennung von in Gebäuden genutzten Geräten mittels Sensoren in den Steckdosen oder bei der Modellierung der Jahreszeiten (Heiz-, Kühl- oder Übergangsperioden).

Durch die Kombination dieser Techniken können fortschrittliche Steuerungsalgorithmen entwickelt werden, die sowohl einen reduzierten Energieverbrauch als auch einen höheren Komfort des Nutzers und eine Anpassung an dessen Bedürfnisse zum Ziel haben. CHe

- future challenges. *Computer Methods and Programs in Biomedicine*, Vol. 91, pp. 55-81, 2008.
- [4] D. Cook: Learning setting-generalized activity models for smart spaces. *IEEE Intelligent Systems*, Vol. 27, pp. 32-38, Jan. 2012.
- [5] A. Ridi, N. Zarkadis, G. Bovet, N. Morel and J. Hennebert: Towards reliable stochastic data-driven models applied to the energy saving in buildings. *CISBAT'13, International Conference on Clean-tech for Smart Cities and Buildings*, Lausanne, 4-6 Sept. 2013.
- [6] J. Scott, A.J.B. Brush, J. Krumm, B. Meyers, M. Hazas, S. Hodges and V. Villar: PreHeat: Controlling Home Heating Using Occupancy Prediction. *Proceedings of the 2011 ACM Conference on Ubiquitous Computing, UbiComp'11*, Beijing, China, pp. 281-290, 2011.
- [7] C. Gisler, A. Ridi, D. Zufferey, O.A. Khaled and J. Hennebert: Appliance consumption signature database and recognition test protocols. *Proceedings of the 8th International Workshop on Systems, Signal Processing and their Applications, WoSSPA'13*, pp. 336-341, 2013.
- [8] A. Ridi, C. Gisler and J. Hennebert: Automatic identification of electrical appliances using smart plugs. *Proceedings of the 8th International Workshop on Systems, Signal Processing and their Applications, WoSSPA'13*, pp. 301-305, 2013.
- [9] A. Ridi, C. Gisler and J. Hennebert: Unseen appliances identification. *Proceedings of the 18th Iberoamerican Congress on Pattern Recognition (CIARP 2013)*, Havana, Cuba, 2013.
- [10] K. Carrie Armel, A. Gupta, G. Shrimali and A. Albert: Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, Vol. 52, pp. 213-234, Jan. 2013.
- [11] R. Yang and M. W. Newman: Living with an intelligent thermostat: Advanced control for heating

and cooling systems. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp'12*, Pittsburgh, PA, USA, pp. 1102-1107, 2012.

[12] <http://nest.com>.

[13] [www.smart-bits.net](http://www.smart-bits.net).

### Informations sur les auteurs

**Antonio Ridi** est collaborateur scientifique à l'École d'Ingénieurs et d'Architectes de Fribourg (HES-SO) et doctorant à l'Université de Fribourg. Il est titulaire d'un Master en ingénierie biomédicale de l'Université de Florence en Italie. Ses travaux de recherche portent sur le machine learning.

EIA-FR, 1705 Fribourg, [antonio.ridi@hefr.ch](mailto:antonio.ridi@hefr.ch)

**Christophe Gisler** est collaborateur scientifique à l'École d'Ingénieurs et d'Architectes de Fribourg (HES-SO) et doctorant à l'Université de Fribourg. Il est titulaire d'un Master of Science en informatique de l'Université de Fribourg. Ses travaux de recherche portent sur le machine learning.

EIA-FR, 1705 Fribourg, [christophe.gisler@hefr.ch](mailto:christophe.gisler@hefr.ch)

Prof. Dr **Jean Hennebert** est professeur à l'École d'Ingénieurs et d'Architectes de Fribourg (HES-SO), ainsi que chargé de cours à l'Université de Fribourg où il dirige des thèses de doctorat. Il est titulaire d'un Master en ingénierie électrique et télécommunications (FPMS Mons, Belgique), ainsi que d'un doctorat obtenu auprès de l'EPFL.

EIA-FR, 1705 Fribourg, [jean.hennebert@hefr.ch](mailto:jean.hennebert@hefr.ch)

Certains travaux mentionnés dans cet article, dont le projet Green-Mod, ont été financés par la Fondation Hasler et par la Haute École Spécialisée de Suisse Occidentale, HES-SO.