

Proposition de stage

Segmentation et décomposition de données temporelles guidée par la physique

Mots clés : data, apprentissage automatique, génération de données synthétiques à partir de modèles physiques, qualité de l'air intérieur, consommation d'énergie au sein de bâtiments, comportements, algorithmes, R/Python

Pour candidater : envoyer un CV et une lettre de motivation à ionescu@u-pec.fr, allou.same@univ-eiffel.fr

Ce stage s'inscrit dans une double problématique : la maîtrise de la consommation d'énergie et l'amélioration de la qualité de l'air dans les espaces de vie. Aujourd'hui, les gens passent une grande partie de leur temps à domicile ou sur leur lieu travail ; il devient donc essentiel d'améliorer la compréhension de la dynamique d'évolution de la consommation d'énergie et de la qualité de l'air dans ces environnements clos. Dans cette optique, le stage vise à mettre en œuvre des algorithmes d'apprentissage automatique pour analyser conjointement des données environnementales, énergétiques et thermiques dans le but d'en extraire des indicateurs reflétant la dynamique comportementale des occupants de bâtiments. Une bonne compréhension du comportement des occupants est en effet cruciale pour développer des stratégies prédictives efficaces permettant de réguler la température intérieure, d'optimiser la ventilation et de limiter l'exposition aux polluants.

L'analyse de données temporelles, notamment la segmentation et la décomposition de mesures liées à la consommation d'énergie et aux concentrations de divers polluants dans l'air constitue le cœur de ce stage. En s'appuyant sur des méthodes d'apprentissage automatique, le travail consistera à traiter ces données multivariées dans le but de produire des résultats compatibles avec les connaissances sur la dynamique d'évolution des systèmes physiques et environnementaux sous-jacents. L'accent sera mis sur les points suivants, qui seront abordés en fonction de l'avancement du stage :

- Collecte de données : le candidat recruté aura pour mission initiale de collecter et de s'approprier les bases de données existantes concernant la consommation d'énergie et la qualité de l'air intérieur. Il devra les harmoniser sous un format unifié et les contextualiser. Ces bases pourront inclure à la fois des données réelles et des données synthétiques générées à partir de modèles physiques simplifiés qui ont déjà été développés dans le cadre de cette étude.
- Etude bibliographique méthodologique : le candidat mènera dans un second temps une étude bibliographique sur l'extraction de comportements dynamiques par analyse conjointe de données relatives à la qualité de l'air et celles liées à la consommation d'énergie.
- Formalisation et développement d'algorithmes : il s'agira dans un troisième temps de développer des algorithmes d'apprentissage automatique, notamment pour la segmentation et la décomposition de séries temporelles. Ces algorithmes seront appliqués aux données collectées lors de la première étape, en veillant à assurer leur cohérence avec les connaissances physiques existantes.

Les développements informatiques seront réalisés sous R/Python.

Environnement de travail : le stage se déroulera au sein des laboratoires CERTES (Université Paris-Est Créteil Val de Marne) et GRETTIA (Université Gustave Eiffel)

Profil du candidat : niveau M2 (informatique, statistique, science de données, mathématiques, physique), avec de bonnes bases en programmation

Durée du stage : 5 à 6 mois (démarrage en mars 2025)

Rémunération : environ 600 euros + prise en charge partielle des frais de transport