



EDF R&D : PROPOSITION DE STAGE (6 MOIS)

Estimation du volume de production photovoltaïque d'une entreprise par Machine Learning

Le contexte :

Ces dernières années, de nombreux clients de EDF (particuliers et entreprises) ont fait installer des panneaux photovoltaïques. Ce déploiement soutenu s'accompagne de nouveaux défis.

En effet, EDF doit notamment anticiper la consommation future de son périmètre de clients, afin de provisionner l'énergie correspondante. Cette activité est réalisée par la Direction Sourcing Economie et Finance (DSEF), qui doit régulièrement s'adapter aux évolutions de la société.

Les clients peuvent revendre cette production photovoltaïque (ils sont alors connus en tant que producteurs, avec un contrat dédié), ou choisir de l'autoconsommer. Ces auto-consommateurs ne sont pas toujours connus par EDF : ils n'ont pas de contrat dédié, et n'injectent pas d'énergie sur le réseau électrique.

Pourtant cette information est capitale pour EDF. En effet, grâce à leur production photovoltaïque, ces entreprises soutirent au réseau une quantité d'énergie plus faible que prévue lors des jours de production solaire. In fine, la consommation est partiellement surestimée, ce qui représente un surcoût pour DSEF.

Au sein de la R&D d'EDF, l'équipe « Analyse des Consommations Clientèle, Offres de Fourniture » regroupe une vingtaine d'ingénieurs-chercheurs de profils data science, informatique et économie. Leur objet d'étude est les courbes de charges (courbes de consommation d'électricité à un pas de temps fin) des clients de EDF.

L'objectif de ce stage est de produire un algorithme permettant, à partir d'une courbe de charge d'une entreprise, de caractériser le volume d'énergie auto-consommé par cette entreprise.

Le stage :

Lors de précédents travaux, un algorithme permettant de déterminer si une entreprise dispose de panneaux solaires pour de l'autoconsommation a été développé. Cette tâche relève de la classification d'une courbe de charge, et les performances de cet algorithme sont très satisfaisantes.

L'objectif du stage proposé est de développer un algorithme permettant d'estimer le volume d'énergie autoconsommé par une entreprise dont on sait qu'elle dispose de panneaux solaires en autoconsommation. Le but est d'estimer la courbe de production photovoltaïque d'une entreprise, à partir de sa courbe de charge. On appelle cela la désagrégation d'une courbe.

Cet algorithme reposera sur deux sources de données : d'une part la courbe de charge de l'entreprise, et d'autre part la météo connue la plus proche de cette même entreprise. La donnée météorologique d'intérêt s'appelle l'irradiance (en W/m^2), qui donne l'intensité locale du flux solaire. Ce stage s'intéresse spécifiquement aux entreprises de France continentale, mais une extension de l'étude sera menée pour les entreprises en outre-mer, dans certains territoires (Corse, Martinique, Guadeloupe, Réunion, Guyane).

L'organisation prévisionnelle du stage est la suivante :

- **Prise en main du sujet**
 - **Etat de l'Art** : publications pertinentes pour le sujet, que ce soit sur l'estimation de production photovoltaïque grâce au Machine Learning ou Deep Learning, ou bien le lien physique entre irradiance et production photovoltaïque (modélisation physique).
 - **Prise en main des infrastructures de travail de l'équipe** : bases de données SQL, serveur de calcul Python, mise à niveau à Git.
 - **Familiarisation à l'analyse de courbes de charge** et au marché d'affaires EDF.

- **Analyse de l'existant** : étude des travaux déjà réalisés par l'équipe autour du PV et de la désagrégation de courbes de charge.
- **Implémentation d'un modèle simple (baseline), et conception d'un protocole expérimental de mesure systématique des performances**
 - Implémentation d'une première méthode volontairement simple d'estimation de la production PV (sans Machine Learning).
 - Conception d'une démarche expérimentale : mesurer les performances avec des métriques pertinentes pour la tâche de désagrégation. Cette démarche expérimentale devra être pensée pour pouvoir tester de nouveaux modèles.
- **Implémentations de nouveaux modèles et de nouvelles idées**
 - À la lumière de l'état de l'art et des discussions avec les autres chercheurs de l'équipe, proposer de nouvelles idées (modèles, features pertinents).
 - Implémenter les idées les plus prometteuses
 - Quantifier l'apport individuel de chaque nouvelle proposition, en comparaison à la baseline et aux modèles précédents.
- **Restitution exhaustive des travaux du stage**
 - La/le stagiaire devra produire un code modulaire, documenté, et conformes aux bonnes pratiques de programmation.
 - Nous attendons une comparaison méthodique (benchmark) des différents modèles proposés.
 - Une analyse critique des performances (et plus spécifiquement des erreurs) devra être menée : dans quelles situations l'algorithme se trompe, pourquoi, et dans quelle mesure ?

Informations complémentaires :

La R&D propose ce stage de fin d'études, à destination d'étudiants en écoles d'ingénieurs ou Master 2, spécialisés en Data Science / Statistiques / Deep Learning.

Compétences : L'étudiant(e) sera amené(e) à mettre en œuvre et/ou acquérir des compétences en :

- Machine Learning (régressions, modèles à base d'arbres, gradient boosting),
- Modélisation de séries temporelles,
- Deep Learning (RNN, CNN, Transformers),
- Bases de données relationnelles (SQL),
- Lecture et synthèse d'articles de recherche.

Les développements seront réalisés en Python. Une expérience démontrée avec des projets de Machine Learning serait appréciée. Une expérience en Deep Learning serait également reçue positivement.

Sur le plan humain, la personne retenue devra être curieuse, rigoureuse, et force de proposition.

Dates : à partir de février/mars 2024, pour une durée de **6 mois**.

Lieu du stage : EDF Lab Paris-Saclay – Recherches et Développement, 7 Bd Gaspard Monge, 91120 Palaiseau.

Contacts :

Eva GIROUSSE (Ingénieure-chercheuse), mail : eva.girousse@edf.fr

Nathan ETOURNEAU (Ingénieur Chercheur), mail : nathan.etourneau@edf.fr

Dominique VOIRIN (Chef de Projet), mail : dominique.voirin@edf.fr

Merci d'envoyer un C.V et un court texte détaillant votre motivation à ces e-mails.

Horaires : 35 h / semaine

Gratification : Selon grille EDF

Bibliographie :

- D. L. Donaldson et D. Jayaweera, « Effective solar prosumer identification using net smart meter data », *International Journal of Electrical Power & Energy Systems*, vol. 118, p. 105823, juin 2020, doi: [10.1016/j.ijepes.2020.105823](https://doi.org/10.1016/j.ijepes.2020.105823).
- W. Kanchana et J. G. Singh, « Detection of Invisible Solar PV Installations in Power Distribution System: A Review », in *2022 International Conference and Utility Exhibition on Energy, Environment and Climate Change (ICUE)*, oct. 2022, p. 1-8. doi: [10.1109/ICUE55325.2022.10113521](https://doi.org/10.1109/ICUE55325.2022.10113521).
- F. Kabir, N. Yu, W. Yao, R. Yang, et Y. Zhang, « Estimation of Behind-the-Meter Solar Generation by Integrating Physical with Statistical Models », in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, oct. 2019, p. 1-6. doi: [10.1109/SmartGridComm.2019.8909743](https://doi.org/10.1109/SmartGridComm.2019.8909743).