

WENDY SAN

+33679247833 ◊ 92600, Asnières-sur-Seine, France ◊ wen2104@hotmail.fr

Proactive et force de proposition, je recherche un poste de Data Scientist à partir de décembre 2024. Je possède une solide formation en statistiques, en mathématiques appliquées et en modèles de machine learning fondamentaux, complétée par une expérience pratique en tant que Data Scientist chez EDF R&D.

FORMATION

Diplôme d'ingénieur, ENSTA Paris, IP Paris, 91120 Palaiseau Sept 2021- Avril 2024
Mathématiques appliquées / Data science. Cours pertinents :

- Machine Learning, Deep learning, reinforcement learning, probabilités & statistiques
- Optimisation différentiable et discrète
- Recherche opérationnelle (LP, MILP)

Classe préparatoire MPSI/MP, Lycée Condorcet, 75009 Paris Sept 2018 - Juil 2021

EXPERIENCE

Stagiaire Data Scientist (R&D) Avril 2024 - Oct 2024
EDF Lab *Palaiseau, France*

- Apprentissage de représentations de séries temporelles pour le clustering de courbes de charge
- Implémentation en pytorch de modèles à base de d'autoencodeurs, de contrastive learning, utilisation du modèle de fondation MOMENT.

Stage de recherche Mai 2023 - Août 2023
Université de Tokyo, Campus Hongo *Tokyo, Japon*

- Inférence d'une équation aux dérivées partielles du mouvement collectif de *Bacillus subtilis* en suspension à partir d'une vidéo (Toner-Tu-Swift-Hohenberg)
- Apprentissage supervisé (régression Ridge/LASSO et processus gaussien) pour l'inférence de termes candidats à l'EDP

PROJETS

Projet CAPSTONE Jan 2024 - Avril 2024

Projet de groupe en collaboration avec Capgemini Invent pour la désaggrégation d'une courbe de charge. Implémentation d'un modèle de machine-learning pour désagréger la partie thermo-sensible et la partie régulière d'une courbe de charge. Comparaison de plusieurs modèles de prévisions (LSTM, XGBoost, Catboost) pour la consommation électrique et l'émission de CO₂. et SHAP-values.

COMPÉTENCES

Langages de programmation	Python (pandas, pytorch, scikit-learn etc...), R, Julia (CLPEX, Gurobi)
Visualisation	Python (plotly), PowerBI
Autre	Dataiku, SQL, Git, LaTeX, Microsoft Office (Word, Excel, etc)
Langues	français, anglais (bilingue), espagnol - japonais (intermédiaire)

CERTIFICATIONS

Dataiku	Core Designer Certificate, Advanced Designer Certificate, ML Practitioner
Anglais	Linguaskill (2023) (C1+ level), IGCSE English as a Second Language (2018) grade A.
Japonais	JLPT N2 (déc 2021).