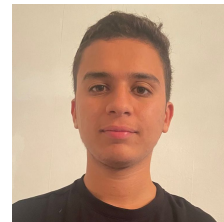


Rachid Sahli



Recherche alternance 36 mois – cycle d'ingénieur (rythme : 1 mois école / 1 mois entreprise)

✉ sahlirachid427@gmail.com ☎ 06 49 75 56 23 🔗 rachidsahli.github.io/site 📄 rachidsahli

Formation

Polytech Lille

Lille, 59000

Diplôme d'ingénieur, Informatique et Mathématiques appliquées

Sept 2025 – Sept 2028

Statistiques, Probabilités, Analyse, Deep Learning, Optimisation, Recherche opérationnelle, Programmation objet

IUT de Paris - Université Paris Cité

Paris, 75016

Bachelor Universitaire de Technologie, Science des Données

Sept 2022 – Juin 2025

Statistique inférentielle, paramétrique et non-paramétrique, Modélisation statistique, Algèbre linéaire, Analyse, Probabilités, Machine learning, Data mining, Programmation, Base de données

Expérience professionnelle

Data Scientist

Montrouge, 92120

Institut national de la statistique et des études économiques (INSEE)

Sept 2023 – Sept 2025

Alternance de deux ans au sein de la direction des statistiques démographiques et sociales, au sein de l'équipe Répertoire statistique des individus et des logements (Résil). Mes principales responsabilités incluent :

- Analyser la couverture et la représentativité d'une base de sondage par rapport au recensement de la population, et évaluer les gains en couverture.
- Confronter plusieurs techniques d'appariement, en comparant leurs indicateurs statistiques et en caractérisant leurs divergences.
- Modéliser (arbres de décision, régression logistique) pour prédire l'identification d'individus lors de l'appariement entre sources administratives.

Outils utilisés : R, RShiny, SQL

Projets

Classification d'images MNIST-2

- Développer un modèle de classification binaire pour distinguer les chiffres 2 et 7 du dataset MNIST.

Outils utilisés : Python (Scikit-learn, Pandas, NumPy, Matplotlib)

Prédiction de la production de charbon aux États-Unis

github.com/serie_temp 🔗

- Analyser la saisonnalité : calculer les coefficients saisonniers, décomposer la série et ajuster les données.
- Modéliser et prévoir la production pour 2023 via trois méthodes (tendance + saisonnalité, Holt-Winters, ARMA), puis comparer les erreurs sur l'année 2022.
- Concevoir et déployer une application RShiny pour visualiser les résultats.

Outils utilisés : R (Forecast, Tseries) & Rshiny

Prévision en temps réel de la consommation via l'API Enedis

- Traiter et analyser statistiquement des données énergétiques en temps réel.
- Concevoir des modèles de machine learning pour prédire la consommation énergétique.

Outils utilisés : Python (Pandas, Scikit-learn, Streamlit), Git, API REST

Compétences et Centres d'Intérêt

Techniques : Python (Pandas, Numpy, Scikit-Learn, OpenCV), C, R, SQL, FastAPI, Git, LaTeX

Langues : Français (natif), Anglais (B1), Espagnol (B2)

Centres d'intérêt : Cyclisme, Robotique, Architecture, Cinéma | **Bénévolat :** DataForGood, La Chorba