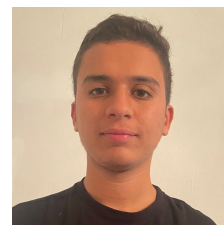


Rachid Sahli



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

Formation

IUT de Paris - Université Paris Cité

Paris, 75016

BUT Science des Données, parcours **exploration et modélisation statistique**

Sept 2022 – Juin 2025

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

Lycée François Villon

Paris, 75014

Baccalauréat STMG, Mention Très Bien

Sept 2021 – Juin 2022

Options : Mention européenne, théâtre, ping-pong

Expérience professionnelle

Statisticien

Montrouge, 92120

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

Sept 2022 – 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). J'étais en charge de comparer la couverture de deux bases de sondage. Mes principales responsabilités incluaient :

- L'analyse de la couverture et de la représentativité de la base Résil par rapport au recensement de la population et l'évaluation des gains en couverture.
- La confrontation de plusieurs techniques d'appariement, sur leurs indicateurs statistiques, mais aussi en caractérisant leurs divergences.
- La réalisation de contrôles approfondis sur des échantillons pour arbitrer entre les méthodes et qualifier ainsi que quantifier les erreurs d'appariement.
- Le test de nouvelles règles de décision de la constitution de la base Résil et la mesure de leur impact sur la qualité de couverture.

Outils utilisés : R, RShiny, SQL

Projets

Classification d'images MNIST-2

- Développement d'un modèle de classification binaire pour distinguer les chiffres 2 et 7 du dataset MNIST.
- Utilisation de la descente de gradient et d'opérations numériques sur les images pour extraire des caractéristiques permettant de passer d'une image à un seul nombre.
- Optimisation d'un seuil afin de minimiser l'erreur de classification.

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

Prédiction de la production de charbon aux États-Unis

github.com/serie_temp

- Analyse de la saisonnalité : calcul des coefficients saisonniers, décomposition de la série et ajustement.
- Modélisation et prévision de la production pour 2023 via trois méthodes de prévision (tendance + saisonnalité, Hot-Winters, ARMA), comparaison des erreurs sur l'année 2022.
- Conception et déploiement d'une application RShiny pour visualiser les résultats.

Outils utilisés : R (Forecast, Tseries) & Rshiny

Compétences et Centres d'Intérêt

Techniques : Python (Pandas, Numpy, Machine Learning), R et RShiny, SQL, SAS, Git, Excel, L^AT_EX

Langues : Français (natif), Anglais (C2), Espagnol (C1), Arabe (C1)

Centres d'intérêt : Robotique, Vélo, Cinéma, Programmation