

Paul LO

## Data Analyst BI – Alternance

✉ hsiaowenpaul@gmail.com

🏠 PARIS

📍 Île-de-France  
75013

☎ +33 767694454

### Mes réalisations

Application web de comparaison de villes développée sous Streamlit –

<https://projetville.streamlit.app>

e-portfolio : [drakoriz.github.io](https://github.com/drakoriz)

### Langues

Anglais

B2 – professionnel (meetings, rédaction et compréhension technique)

### Certificat

Cambridge English Certificate (CEC)

LEVEL B2, Sept 2022, Paris

### Centres d'intérêt

Art digital

Création visuelle sur Procreate

### Réseaux sociaux

in Paul LO

Étudiant en Big Data & Machine Learning à l'EFREI, je mets à profit une expérience en data science appliquée (Data Analyst – RATP Dev) et des projets académiques réussis. Je recherche une alternance 3j/2j.

### Diplômes et Formations

#### ● Cycle Ingénieur - Big Data et Machine Learning

De septembre 2025 à août 2028 EFREI – Villejuif

Statistiques et apprentissage automatique, Apprentissage profond et par renforcement, Modèles de langage de grande taille (LLM), Cloud computing et systèmes distribués

#### ● BUT - Science des Données

De septembre 2022 à juin 2025 IUT de Paris – Rives de Seine – Paris

Formation à la data science, couvrant l'ensemble de la chaîne de valeur des données : analyse, visualisation et conception d'outils décisionnels.

#### ● Baccalauréat Général

2022 Lycée Pierre-Gilles de Gennes – ENCPB – Paris

Enseignements de spécialités : Mathématiques et Physique-Chimie

Options : Mathématiques expertes et Section européenne en anglais.

### Expérience professionnelle

#### ● Data Analyst (alternance) – Pôle Data, Direction Digital & Data

De septembre 2024 à septembre 2025 RATP Dev Paris

- **Exploitation des données de transport et de maintenance** via *Python*, *SQL* et *Snowflake* pour concevoir des indicateurs de performance et alimenter des dashboards *Apache Superset* pour le réseau d'Angers.
- **Conception et déploiement** d'un *dashboard interactif Streamlit* (API intégrée, stockage *AWS S3*) pour le suivi des KPI du réseau d'Al'Ula (*Arabie saoudite*), présenté à la *Royal Commission of AlUla* et utilisé quotidiennement par l'équipe locale.
- Développement d'un **outil de pilotage des pièces de bus** (acheminement, stocks, demandes).

### Compétences

#### Langage & Data

Python, SQL, R, SAS, MongoDB

#### Analyse & Apprentissage Automatique

Classification, Séries temporelles, Régression linéaire, Analyse factorielle  
Estimation, Statistiques inférentielles, Arbres de régression CART, Clustering

#### Outils BI & Viz

R Shiny, Power BI, Tableau, Excel, Dataiku

#### Développement & Cloud

Django, Streamlit, Web Scraping, HTML, CSS, Outils Cloud (AWS), Git

#### Autres bibliothèques

Pandas, NumPy, Scikit-learn, TensorFlow, Keras, Plotly, Seaborn, Matplotlib

### Projet professionnels et universitaires

#### Université – Ciblage rachat sur données ( 2025 )

Développement d'un modèle de prédiction des clients susceptibles de racheter à partir de données CRM.

Évaluation de plusieurs méthodes de scoring (aléatoire, ciblée, profilée) et implémentation d'un modèle Random Forest avec feature engineering avancé.

Conception d'un dashboard Power BI pour la visualisation des résultats.

**Résultat** : Meilleur score de la promotion obtenu avec 95% de rachats contre le rachat moyen de 45%

#### Université – Apprentissage pour l'IA : Classification CIFAR-10 (2025)

Analyse et préparation de 50 000 images (format 32x32) ; entraînement de deux modèles d'apprentissage automatique (un modèle simple et un modèle convolutionnel) avec TensorFlow/Keras.

**Résultats** : le modèle convolutionnel obtient une précision de **76% sur les données de test** contre **48%** pour le modèle simple ; le modèle convolutionnel apprend mieux les images mais montre davantage de sur-apprentissage.