

Profil

Étudiant en Master 2 Mathématiques et Applications – Ingénierie Statistique et Financière, voie Data Science, à l’Université Paris-Dauphine PSL, je recherche un **stage de fin d’études** dans le domaine de l’**intelligence artificielle appliquée**. Passionné par les modèles **statistiques**, l’apprentissage automatique et leur mise en œuvre concrète, je souhaite contribuer à des projets innovants en tant que **Data Scientist** ou **Applied AI Researcher**. Je m’investis pleinement dans les projets alliant **modélisation mathématique** et **application terrain**. Rigoureux, curieux et motivé par la **recherche appliquée**, je vise à approfondir mes compétences théoriques et pratiques, avec l’ambition de préparer une future **thèse en intelligence artificielle**.

Expériences professionnelles

ArtDanjProduction

AUTO-ENTREPRENEUR – REMOTE

FÉV 2022 — PRÉSENT

- Administration de mon **homelab** (serveurs, bases de données, stockage, sauvegardes) utilisé comme support pour des projets de machine learning et d’expérimentation MLOps.
- Conception, développement et maintenance de projets web, data et cloud, en Python, TypeScript, Vue.js, Nuxt 3 et R.
- Maintenance et mise à jour des projets en production.
- Rédaction de documentation technique, articles de blog et mise en ligne sur mon site personnel ([arthurdanjou.fr](http://arthurdanjou.fr)).
- Apprentissage continu de **nouvelles technologies** (LLMs, MLOps, NLP, Agents IA) via projets personnels intégrés dans un environnement DevOps.

Sevetys

STAGE DATA ENGINEER – PARIS, FRANCE

JUIN — JUILLET 2025

- Définition et implémentation de règles de nettoyage pour les données clients et patients.
- Mise en production de **pipelines de data cleaning** quotidiens sur Azure à l’aide de PySpark, améliorant la qualité des données de **20 à 30 %** selon les cliniques.
- Automatisation d’un **rapport mensuel de qualité des données**, détaillant les taux de complétude, cohérence et exploitabilité par clinique.
- Travail en lien avec les équipes métier pour assurer la **pertinence** des règles de qualité.

Formation académique

Master 2 - Ingénierie Statistique et Financière, voie Data

UNIVERSITÉ PARIS DAUPHINE-PSL – PARIS, FRANCE

Cours clés : Machine Learning (théorie et pratique), Deep Learning, NLP, Reinforcement Learning, Clustering, Estimation non paramétrique, Learning Theory, Generative AI, Data Quality, Visualisation (R), SQL.

Projet transverse : Data Project – Projet complet de data science sous la forme d’un “Hackaton 3.0”.

Master 1 - Mathématiques et applications, voie Statistiques

UNIVERSITÉ PARIS DAUPHINE-PSL – PARIS, FRANCE

Cours clés : Statistical Learning (supervisé), Modèles Linéaires Généralisés (GLM), Monte Carlo, Analyse de données, Statistique non paramétrique, Séries temporelles, Optimisation numérique.

Projets réalisés :

- Simulation Monte Carlo** (R & LaTeX) : Exploration de divers estimateurs à l’aide de méthodes et d’algorithmes de Monte Carlo.
- Modèles Linéaires Généralisés** (R & LaTeX): Modélisation de la demande de vélos en libre-service à l’aide de la régression linéaire, avec évaluation de la performance et de l’adéquation du modèle.
- Apprentissage Statistique** (Python & LaTeX) : Détection du cancer du sein pour classifier les patients à l’aide de différentes techniques d’apprentissage supervisé (k-NN, réseau de neurones, régression logistique et Naive Bayes).

Licence de mathématiques

UNIVERSITÉ PARIS-SACLAY – ORSAY, FRANCE

- Modèle de ségrégation de Schelling** (Python & LaTeX) : Analyse statistique et probabiliste du modèle de ségrégation de Schelling. Simulation et visualisation des résultats à l’aide de Python.

Compétences

- Programmation** : Python, R, TypeScript, Java, Git et LaTeX.
- Librairies & Frameworks** : NumPy, Pandas, Scikit-learn, PyTorch, Matplotlib, Seaborn
- Base de données** : SQL, Redis
- Cloud & DevOps** : Proxmox, Docker, Azure, Linux, SysAdmin
- Analyse de données multivariées** : Analyse en Composantes Principales (ACP), Analyse des Correspondances (AC), techniques de clustering et analyse des corrélations.
- Apprentissage supervisé** : k-NN, régressions linéaires et logistiques, CNN, naïve Bayes
- Apprentissage non-supervisé** : clustering, réduction de dimension, k-means
- IA & Modèles modernes** : NLP, Transformers, LLMs, agents IA, embeddings, fine-tuning

Soft Skills

Solides compétences analytiques –  
Curiosité scientifique – Autonomie technique – Communication efficace –  
Travail collaboratif en équipe agile

Langues

- Français (langue maternelle)
- Anglais (B2/C1)
- Espagnol (A2)

Intérêts et Passions

IA et ses concepts, Mathématiques appliquées, Cloud personnel (homelab), Sports d’équipe

Portfolio

Tous mes projets académiques et personnels sont accessibles sur mon site : [arthurdanjou.fr](http://arthurdanjou.fr).  
Retrouvez également mon blog : <https://go.arthurdanjou.fr/writings>.