

**ANAS Yassine**

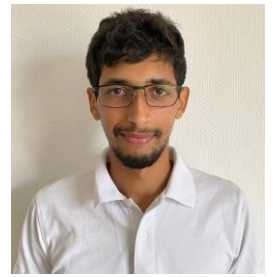
**Lieu :** France

**Tél :** 06 59 06 03 23

**Email :** yassine.anas71@gmail.com

**Portfolio :** <https://yassined4s.github.io/portfolio/>

# Data Scientist



## EXPERIENCES

### Projets personnels Data Scientist (Depuis Novembre 2024)

- Mise en place de Dashboards interactifs (PowerBI, Dash)
- Maîtrise des modèles génératifs, incluant OpenAI et Llama, avec expertise dans leur utilisation pour des cas pratiques
- Conception et implémentation de pipelines de traitement de données de masse en utilisant Azure Data Factory pour l'orchestration, et les notebooks Azure Databricks avec PySpark pour le traitement distribué et l'analyse avancée
- Développement de modèles d'apprentissage from scratch, sans recours à des bibliothèques spécialisées (Python)

### Stage de fin d'étude Data Scientist (Avril 2024- Octobre 2024)

Au sein du pôle « NEX'US », chez **ENEDIS**, Lyon, France.

- Structurer et exploiter une base de données pour optimiser le traitement des données (SQL : DBeaver)
- Analyser, nettoyer, traiter et visualiser des données afin d'améliorer la qualité des insights (Python, R)
- Concevoir, développer, implémenter, évaluer et interpréter des modèles prédictifs de machine learning pour optimiser le rattachement des clients au réseau (Python)
- Automatisation et industrialisation des pipelines analytiques avec des outils de versionning (Python, Git)

**Outils :** SQL, Python, R, QGIS et Git

### Stagiaire pédagogique Data Scientist (Septembre 2023- Février 2024)

Au sein du **HUG**, Genève, Suisse.

- Construction et évaluation d'un modèle de risque linéaire sur des données de rétinopathie afin d'améliorer le diagnostic (Python)
- Développer et évaluer un modèle de Machine Learning pour prédire la survenue d'une IRA améliorant la détection précoce (Python)
- Développement d'un classificateur de radiographies thoraciques basé sur des réseaux de neurones et évaluation du modèle avec AUC-ROC (Python)

**Outils :** Python

### Stagiaire de recherche Data Scientist (Janvier 2023- Mai 2023)

Université Claude Bernard, Lyon, France.

- Implémenter et comparer des modèles de Machine Learning pour des tâches de régression et de classification (Python)
- Interpréter et expliquer les modèles sélectionnés à l'aide d'algorithmes agnostiques pour renforcer la confiance dans les prédictions (Python)

**Outils :** Python

## FORMATION

**2022 - 2024 :** Master Statistiques Modélisation et Science des Données, Université Claude Bernard, Lyon, France.

**2018-2022 :** Licence de mathématiques, Université Claude Bernard, Lyon, France

## COMPETENCES

**Mathématiques :** Algèbre linéaire, analyse et statistiques

**Programmation :** Python (NumPy, Pandas, Pyspark, Scikit-learn, TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, Dash), R (ggplot2, dplyr, tidyr, Shiny)

**Base de Données :** Conception et exploitation de base de données avec SQL (DBeaver) et NoSQL (Cassandra)

**Analyse de Données :** Visualiser, nettoyer et traiter des données. Mise en place de Dashboards instructifs sur Power BI

**IA et Machine Learning :** Développement et évaluation de modèles prédictifs pour la classification supervisée, non supervisée, la régression, ainsi que de projets en génération d'IA (GenAI), dont une étude sur la Stable Diffusion.

**Cloud et Big Data :** Développement de pipelines de traitement de données massives avec ADF pour le stockage, Databricks et PySpark pour le traitement distribué, et orchestration optimisée des workflows

## SOFT SKILLS

Travail en équipe, autonome, rigoureux et proactif

## LANGUES

**Français :** Maternelle  
**Anglais :** Opérationnel

## LOISIRS

Sport, cuisine et voyage