



## Abdourahmane Ndiaye

Data engineer/scientist

📍 69000 Lyon

☎ 07 66 23 28 94

✉ abane.ndiaye@gmail.com

🌐 [linkedin.com/in/abdourahmane-ndiaye-116781253/](https://www.linkedin.com/in/abdourahmane-ndiaye-116781253/)

🌐 <https://github.com/Abdouragit>

### COMPETENCES TECHNIQUES

#### Data science & framework IA :

Machine learning, NLP, RAG, Langchain, openai, mistral, Ollama

**Versioning et CI/CD :** Git, Github, Azure Devops

**Containerisation :** Docker

**Base de données (No)SQL/Vectorielle :** SQLITE, MySQL, PostgreSQL, FAISS

**Langages :** Python, SQL, R

**Reporting :** PowerBI, Tableau, Streamlit, Dash

**Cloud :** GCP et Azure (Data factory, App service, ML Studio)

### COMPETENCES GENERALES

Méthodologie SCRUM

Programmation orienté objet

Analytics et reporting

Database modelisation

Anglais (fluent)

### INTERET

- Lecture
- Dessin
- Psychologie et sciences cognitives

## EXPERIENCE

### MACHINE LEARNING/ DATA ENGINEER

04/2024 - En cours

JEMS, Lyon

- Création d'un chatbot pour l'intranet de Jems en utilisant les documents internes (**RAG** - retrieval augmented generation). Application **Dockerisé** et déployé sur **Azure app service** via **Container Registry**.
- Création d'une application permettant d'analyser les impacts (**performances et coûts**) de différentes étapes d'une architecture **IA/LLM**. Restitution et analyse des résultats sur **Power BI**. Synthèse sur les méthodes de chunking.
- Développement d'algorithme permettant de classifier des projets parmi des fichiers via en utilisant les LLMs.
- Recherche et test de méthode d'extraction d'information de PDF encodés sans OCR.
- Création d'une application permettant d'exploiter une base donnée de dossiers de compétences en utilisant les LLMs.
- Gestion de projet via **Azure Devops** et mise à disposition Livrables documentés.

**Langages & Outils :** Python, Azure Devops, Azure app service, VScode, git, Langchain, Openai, Docker

### Développeur d'applications Rshiny

06/2023 - 08/2023

Cabinet Legal Quantum, Lyon

- Débogage d'applications Rshiny, réorganisation du code et ajouts de fonctionnalités.
- Gestion des bases de données liées aux applications.
- Vérification et modification des accès pour la sécurité des données de l'entreprise.

**Langages & Outils :** R, Rshiny, Rstudio, SQL, phpMyAdmin, Git

## FORMATION

### Master 2 SISE (Statistique et informatique pour la science des données

09/2023 - 09/2024

ICOM, Université Lumière Lyon II

**Contenu des cours :** Deep learning for computer vision - Machine learning sous python - programmation statistique sous R - Data visualisation, BI et Cloud - NoSQL : MongoDB, Redis - Fouille de données massives - Technologies Big Data - Fouille du web et analyse des réseaux sociaux - Text mining - Ateliers : Azure, Oracle, Talend, Nifi, Tableau, Qlik

### Certification GCP Professional Data Engineer

En cours

## PROJETS

#### Analyse et estimation de valeurs foncières sur python :

- Nettoyage de données - Statistiques descriptives - Cartographie des prix des biens fonciers - Application Dash pour analyser les données - Estimation de valeurs foncières.
- [https://github.com/Abdouragit/Land\\_value\\_estimation](https://github.com/Abdouragit/Land_value_estimation)

#### Application d'analyse de données scappées du Web

- Scraping des données Pôle emploi et de l'Apec via **Selenium**.
- Processus **ETL** python et creation d'une base de donnée **Sqlite3**.
- Création d'une application d'analyse **Streamlit** dans un container **Docker**.
- [https://github.com/Abdouragit/Text\\_mining](https://github.com/Abdouragit/Text_mining)