

Data Scientist Gen AI & Big Data

#### Profiles

in https://www.linkedin.com/in/yacine-berkani-a66189244/

https://github.com/yacineberkani

#### Summary

Ingénieur IA junior, diplômé d'un Master 2 Big Data (Université Paris 8), spécialisé en IA générative et architectures agentiques. Fort de deux stages opérationnels (RAG, fine-tuning de LLM, pipelines NLP, micro-services Python/FastAPI, CI/CD, Docker), je maîtrise Python, PyTorch/TensorFlow, LangChain et Hugging Face. Capable de passer rapidement de la R&D à la production, je suis créatif, autonome et animé par l'excellence pour déployer des solutions GenAI à forte valeur ajoutée.

#### Experience

# Stage Data Scientiste, NLP GEN AI | Assystem EOS

Février 2025 - jusqu'à présent

- Intégration d'une API pour Modèles SysML v2 avec une Base de Données de Spécifications
- Extraction des exigences à partir des cahiers de charges (RAG, LLM, NLP).
- Traduction des exigences en langage formel (Fine Tuninge, LLM)

## Stage Data Scientist, Gen AI | Laboratoire Paragraphe Paris 8

Mars 2024 - Août 2024

- Scraping d'Articles à partir d'un Prompt en Langage Naturel (beautifulsoup, request, NER, SpaCy)
- Génération d'étude comparative automatique d'articles scientifiques (RAG, Fine Tuninge, LLM, Hugging Face)

#### Stage Data Analyst, Traitement Visualisation | IRIS

Avril 2023 - Juillet 2023

- L'extraction et le nettoyage des données avec Excel, la transformation des données.
- Création de tableaux de bord interactifs avec Power BI pour faciliter la prise de décision.

#### Education

## Master Informatique Option Big Data | Université Paris 8

Septembre 2023 - Présent

Maîtrise avancée des technologies Big Data, Techniques d'apprentissage automatique approfondies, Expertise spécifique en réseaux de neurones, Probabilités et statistiques pour la théorie de l'information

## Master 1 Recherche opérationnelle | UMMTO

Septembre 2022 - Septembre 2023

Maîtrise des méthodes d'optimisation (avec et sans contraintes ) programmation dynamique, algorithmes d'optimisation, (descente du gradient, simplex). compétences en simulation Monte Carlo.

# Licence mathématique et informatique | UMMTO

Septembre 2019 - Septembre 2022

Maîtrise des algorithmes et des structures de données, Création, gestion et optimisation de bases de données. Analyse de données et modélisation statistique appliquées à la science des données.

## **Projects**

# Détection de Tumeurs Cérébrales à partir d'IRM

Detection\_de\_Tumeurs

Utilisation de modèles de Machine Learning et de Deep Learning pour déterminer la classe d'appartenance d'une imagerie par résonance magnétique (IRM)

# Détection de Fatigue au Volant par Deep Learning

https://github.com/yacineberkani/yacineberkani-Fatigue-Driving-Detection

détection de fatigue au volant basé sur l'apprentissage profond et la fusion multi-indices, MTCNN E-MSR Net

## Prédire le prix de l'immobilier

Prédiction des prix des propriétés immobilières en fonction de leurs caractéristiques. Surface habitable, Nombre de pièces, en utilisant la régression

#### Skills

• • • • •	Python	SQL	R	JAVA
Scala	C	C++	VBA	Scripting Shell
Hadoop	Spark	ELK	<b>Google Cloud Platform</b>	Amazon web service
Talend	PowerBI	Looker Studio	<b>Machine Learning</b>	Deep Learning
Reinforcment Learning	Pandas	NumPy	Scikit-learn	PyTorch
Keras	TensorFlow	Transformers	NLP	LLM
Fine tuning	RAG	Mysql	PostgreSQL	MariaDB
MongoDB	Cassandra	Git	Github	Gitlab
Languages				

Français

**Anglais** . . . . . Kabyle • • • •