



AJERMOUN Hamza

DATA SCIENTIST / ENGINEER

Étudiant en Master Data & Intelligence Artificielle à Hetic - Paris, je recherche une alternance combinant modélisation de données et ingénierie IA. Mes projets m'ont permis de travailler sur des systèmes mêlant NLP, OCR, LLM et traitement distribué, avec un fort accent sur le déploiement et la valorisation des données. Autonome, curieux et orienté solution, je suis motivé à contribuer à des projets data concrets, de bout en bout.

CONTACT

📍 Île-de-France, France

✉ hamzaajermoun06@gmail.com

☎ +33758049726

🌐 [HAMZA AJERMOUN](#)

FORMATIONS

🎓 Master Big Data & Intelligence Artificielle:

Hetic, Paris — 2024 à 2025

🎓 Licence Mathématiques Appliquées

Université de Reims, 2020 - 2021

🎓 Licence Fondamentale en Mathématiques et Applications

Université Hassan 2, 2019 - 2020

🎓 DEUG - Mathématiques & Applications

Université Hassan 2, 2016 - 2018

LANGUES

Français - Bilingue

Anglais - Avancée

Arabe - Langue maternelle

SOFT SKILLS

- Rigueur
- Autonomie
- Esprit d'analyse
- Travail en équipe

PROJETS

OCR Matching System - Streamlit & NLP Hetic, Paris

- Extraction d'informations depuis des reçus scannés (OCR via Mistral AI)
- Matching flou avec relevé bancaire (fuzzy matching : montant, date, commerçant)
- Application déployée sur Streamlit avec génération de fichiers CSV enrichis

LLM-Powered Document Assistant - LangChain & FAISS Projet personnel

- Développé une application interactive permettant de poser des questions à partir de documents PDF multiples.
- Implémenté l'extraction de texte, la recherche sémantique et la génération de réponses avec LangChain, FAISS et Google Generative AI.
- Déployé l'application sur Streamlit avec une interface conviviale.

Sentiment Classification Pipeline - NLP avec BERT Hetic, Paris

- Collecte et prétraitement des données.
- Utilisation de BERT pour l'analyse de sentiments.
- Entraînement de modèles supervisés.

ETL Pipeline - Python, PostgreSQL & Docker Hetic, Paris

- Développement d'un pipeline ETL local pour nettoyer, transformer et charger des données CSV dans une base PostgreSQL.
- Automatisation du pipeline avec un script Python.
- Architecture légère, reproductible sur tout environnement via un Dockerfile.

COMPÉTENCES

Langages & Outils Data : Python, R, SQL, Pandas, Excel, Talend, SAS

Machine Learning & Deep Learning : TensorFlow, Keras, Scikit-learn, PySpark, OpenCV

Modèles : BERT, LSTM, CNN, ResNet50, LLM, NLP

Big Data & Cloud : Hadoop, Spark, Kafka, Docker, AWS, Azure, Google Cloud

Bases de données : MongoDB, Elasticsearch, PL/SQL

Visualisation : Power BI, Tableau, Seaborn, Plotly

DevOps & MLOps : LangChain, FAISS, Git, Streamlit