AYOUB EL HOUDRI *Data Scientist*

✉ ayoubelhoudri@gmail.com   📞 +33 7 83 22 67 03   📍 Paris, France   🔗 y-aoub.github.io   ⌨ github.com/y-aoub

in linkedin.com/in/elhoudriay

SUMMARY

Data Scientist with experience in developing and implementing artificial intelligence solutions. Proficient in machine learning, deep learning, and generative AI, with a focus on contributing to all stages of a project, from research and prototyping to deployment and ongoing improvement. Committed to delivering effective, data-driven solutions through collaboration and continuous learning.

EXPERIENCE

AI Engineer (Permanent), *Waalaxy*   Sep 2024 – Present | Montpellier, France
- Built a multi-agent system to create personalized LinkedIn messages, contributing to a 35% increase in response rates.
- Developed a vector-based search engine for prospect identification, utilizing a database of 700 million LinkedIn profiles.
- Developed pipelines to ensure robust and reliable performance of the models in production.

Tools: *Python (LangGraph, Unsloth), Langfuse, MongoDB (Atlas Vector Database), GCP (Vertex AI, BigQuery)*

Data Scientist and Data Engineer (Freelance), *AEH*   Nov 2022 – Present | Paris, France
- Designed and deployed machine learning models and generative AI workflows for clients across various industries.
- Implemented automated systems for data extraction, transformation, and continuous data warehouse maintenance.
- Improved the reliability and performance of production models for high-traffic services.

Tools: *Python (LangChain, PySpark, Scikit-learn, MLflow), Snowflake, Pinecone, GCP (Cloud Run, Vertex AI)*

Data Scientist (Internship), *Median Technologies*   May 2023 – Oct 2023 | Nice, France
- Developed a GAN-based model for 3D lung segmentation on CT scans, achieving 96% precision.
- Improved a tumor detection pipeline, reducing false negatives by 30%.
- Contributed to a tumor classification system validated by two radiologists to support diagnostic accuracy.

Tools: *Python (PyTorch, MONAI, OpenCV), PostgreSQL, Apptainer (ex-Singularity), 3D Slicer*

Data Scientist (Internship), *Karmen*   Jun 2022 – Sep 2022 | Paris, France
- Developed a data extraction pipeline for financial documents, reducing manual data entry by 75%.
- Built and deployed a fraud detection system for tabular data, achieving 82% accuracy.
- Implemented monitoring for production machine learning models to detect data drift during inference.

Tools: *Python (PyTorch, PyTesseract, OpenCV, XGBoost), Snowflake, Grafana, Dataiku*

Machine Learning Engineer (Internship), *Digimind Labs*   May 2021 – Aug 2021 | Berlin, Germany
- Developed a synthetic 3D data generator to augment the training dataset.
- Trained the Pixel2Mesh model for single-view 3D reconstruction tasks.
- Assisted in deploying models to production, ensuring integration with operational workflows.

Tools: *Python (Tensorflow, OpenCV), Blender API, AWS (Amazon S3, AWS SageMaker)*

EDUCATION

Research Master's Degree in Computer Science – Artificial Intelligence and Complex Systems, *CY Cergy Paris University / ENSEA / ETIS Laboratory (CNRS)*   2022 – 2023 | Cergy, France

Engineering Degree in Applied Mathematics and Computer Science – Data Science, *CY Tech*   2020 – 2023 | Cergy, France

SKILLS

Programming Languages: Python, R, SQL, Bash, C

Machine Learning and AI: PyTorch, Scikit-learn, XGBoost, Hugging Face Transformers, LangChain, LangGraph

Monitoring and Versioning: MLflow, Langfuse, Docker, Git, GitHub Actions

Data Engineering: Pandas, Polars, NumPy, SciPy, PySpark, OpenCV, Apache Airflow

Cloud and Deployment: Google Cloud Platform, Kubernetes, FastAPI, Flask

Databases: Snowflake, MySQL, PostgreSQL, MongoDB, Pinecone

LANGUAGES

French: Native   **English:** Fluent (C1)   **Spanish:** Beginner