Ayyoub BENRGUIG

Graduate Data science Engineering student seeking an entry-level opportunity in AI &Data science starting in september 2025

✉ ayyoub.benrguig@grenoble-inp.org 📞 0646532179 📍 Paris, France in ayyoub-benrguig 🔗 benrguig.com ⊙ Ayyoubbenrguig

PROFESSIONAL EXPERIENCE

Renault Group   02/2025 – 08/2025 │ Guyancourt, France
Data Scientist
- Developed a **RAG-based chatbot** using **GCP resources**. ⊡
- Built **interactive dashboards** for analyzing the composition of the vehicle fleet in France using **Looker Studio**, **Spotfire**, and the **Dash** library.
- Collected, cleaned, and structured a **multi-source database** on **Google Cloud Storage (GCS)** and **BigQuery**.
- Built a reusable data pipeline to integrate vehicle fleet datasets from various sources, hosted on GCS and queried via BigQuery
- **Skills required :** GCP, Vertex AI, RAG, LLM, BigQuery, NLP, Airflow, Python, SQL, Machine Learning, Looker Studio, Spotfire, Chatbot, GCS, Web Scraping

INTELCIA   04/2024 – 08/2024 │ Casablanca, Maroc
AI Engineer Intern
- Designed and deployed **custom chatbots** based on **Retrieval-Augmented Generation (RAG)** architecture for two major clients.
- Enhanced answer quality by integrating advanced techniques such as **Semantic Chunking**, **History-Aware Retriever**, **Multi-Query Retriever**, and **Query Decomposition**, reaching 92 **% answer relevancy**.
- Delivered a working demo and collaborated with cross-functional teams to align chatbot behavior with client needs.
- **Skills required :** LLMs, LangChain, Azure, OpenAI, Claude, Embeddings, Streamlit, LangSmith

Tecpap   06/2023 – 08/2023 │ Casablanca, Maroc
Computer Vision Intern
- Analyzed business needs to improve **Overall Equipment Effectiveness (OEE)** through defect detection on industrial printing lines.
- Developed and deployed **computer vision models** (CNN, GAN) for real-time error detection in printed materials, using advanced image processing techniques.
- Containerized the solution with **Docker** and orchestrated deployment via **Kubernetes**, ensuring scalability and maintainability.
- Contributed to the company's digital transformation by reducing manual inspection effort and increasing defect detection accuracy.
- **Skills required** : Python, C++, OpenCV, Deep Learning, Pytorch, TensorFlow, GAN, CNN, Docker, Kubernetes, AWS, Computer Vision

OCP Group   07/2022 – 08/2022 │ EL Jadida, Maroc
Data Science Intern
- Collected and analyzed operational data from industrial equipment to support predictive maintenance strategies.
- Built automated scripts for web scraping and database structuring using BeautifulSoup, Scrapy, and MySQL. Integrated cleaned datasets into predictive models
- Designed a **preventive maintenance plan** using machine learning models such as **SVM**, **logistic regression**, and **decision trees**, improving early failure detection.
- **Skills required** : Python, SQL, MySQL, BeautifulSoup, Scrapy, Logistic Regression, Decision Trees, SVM

EDUCATION

Ecole nationale supérieure d'informatique et de mathématiques appliquées (Ensimag)   09/2024 – Present │ Grenoble, France
Engineering Degree in Data science Engineering
Key Courses: Deep Learning, Reinforcement Learning, Multi-Agent Systems, Computer Vision, Advanced Natural Language Processing, Machine learning, GP-GPU Computing and High-Performance Computing (HPC).

Ecole Centrale Casablanca(ECC)   09/2022 – 08/2024
Engineering Degree in Data Science
Key Courses: Probability, Optimization, Statistics, Programming, Machine Learning, Graph Theory, Modeling, Operations research.
Grade: Ranked 5th in my cohort (5/180)

RELEVANT PROJECTS

Intelligent Document Parsing Pipeline ⊡   09/2024 – 11/2024
- Built a **complete automated pipeline** to extract structured financial data from complex PDF tables and generate formatted Excel reports.
- Combined **LLM prompt engineering**, **OCR fallback**, and **GPT-based post-correction** to ensure accurate extraction of semi-structured content.
- Integrated tools like **Azure OpenAI**, **Vertex AI**, and **OpenPyXL** for end-to-end data automation and reporting.
- **Tech stack**: Azure OpenAI, Vertex AI, Gemini Pro, Python, Pandas, JSON, OpenPyXL, OCR, Excel Automation

Voicebot Development   10/2023 – 03/2024
- Designed and implemented a **voicebot** using a **Retrieval-Augmented Generation (RAG)** architecture to answer client inquiries in the banking sector.
- Integrated advanced techniques such as **Semantic Chunking** and **Query Decomposition**, achieving **87% answer relevancy**.
- Enabled **speech-to-text** and **text-to-speech** capabilities using **OpenAI Whisper**, and built an interactive front-end with **Streamlit**.
- **Tech stack**: RAG, LangChain, LLMs, OpenAI Whisper, Azure, NLP, Streamlit

Image Segmentation with Swin U-Net Transformer ⊡   09/2023 – 01/2024
- Applied **data augmentation using GANs** to enrich the training dataset for medical and industrial images.
- Built and trained a **Swin U-Net Transformer** model for high-precision image segmentation tasks.
- Compared model performance with classical ML approaches (**KNN**, **SVM**, **Random Forest**) and baseline CNN models like **U-Net**.
- **Tech stack**: Swin U-Net, GAN, U-Net, KNN, SVM, Random Forest, Transformers, Computer Vision, Image Processing

SKILLS AND LANGUAGES

Data Science / AI Tools: Python, Pandas, Numpy, Scikit-Learn, Matplotlib, Seaborn, Pytorch, TensorFlow, LangChain, Transformers, Computer vision, Image Processing.
Data Engineering & Software: SQL, BigQuery, Vertex IA, Dataiku, GCP, AWS, Kafka, CI/CD, Docker, Git, REST APIs, Flask, FastAPI.
Soft Skills: Problem Solving, Leadership, Analytical Thinking, Teamwork, Adaptability, Autonomy, Agility, Continuous Learning, Agile Methodologies, Multicultural Communication.
Languages: English (B2), French (DALF C1), Arabic (Mother Language)