

Youness Boumlik

Data Science & Artificial Intelligence Engineer

Casablanca, Morocco | +212 631467691
younessboumlik1949@gmail.com | Youness Boumlik
Younessboumlik | boumlikyouness.me

About: Recent State Engineer graduate of ENSA Khouribga in **Computer Science and Data Engineering**. Specialized in **Data Science**, **Artificial Intelligence** and data engineering, I design solutions combining Machine Learning, NLP, Computer Vision, Generative AI and software development. Passionate about building complete systems, I am looking for a **Data / AI Engineer** position to contribute to innovative data-driven projects.

EDUCATION

ENSA Khouribga 2021 – 2026
State Engineering Degree – Computer Science and Data Engineering (Très Bien, Highest Honors)

ALX Software Engineering Program 2023 – 2024
Backend Development specialization (12-month intensive program) Online

Khadija Oum Al Mouminine High School 2021
Baccalaureate in Mathematical Sciences A (Good Honors)

PROFESSIONAL EXPERIENCE

Data Science & AI Intern (Final Year Project) Feb. – Jul. 2026
BIM Pioneers – Casablanca

- **Design & Development:** Design and development of BIMFLUX, an AI platform combining an AutoCAD plugin, a Python FastAPI backend (multi AI engines), a PostgreSQL/PostGIS database and a Revit plugin in WPF/MVVM, with **Docker** containerization and a professional installer (**Inno Setup**) for multi-version deployment, automating the conversion of 2D MEP CAD drawings into 3D BIM models.
- **Classification (ML):** Developed a **linear SVM + TF-IDF** engine for the hierarchical classification of MEP equipment (Discipline/Sub-discipline/Type), reaching **92.4% accuracy** on the discipline. Training data labeled with an LLM: comparative study of several LLMs, **Qwen** (via **Ollama**) selected, with validation by domain experts.
- **NLP / NER:** Built a hybrid **spaCy + EntityRuler** engine to extract technical annotations (dimensions, flow rates, diameters), with an **F1-score of 99.6%**.
- **Geometric Detection:** Benchmark of 4 algorithms (**RandomForest** selected) for classifying HVAC duct networks, with *group-safe* cross-validation and strict data-leakage prevention, reaching **81.6% precision** and **94.8% ROC-AUC** on an independent test set.

Data Science Intern (Year Project) Jul. – Sep. 2025
Maghreb Steel – Casablanca

- **Data Engineering:** Complete **CRISP-DM** pipeline on **90k+ rows** of production data (cleaning, outlier handling, harmonization of redundant variables) and selection of **8 key variables** via **RFE**.
- **Modeling & Deployment:** Benchmark of **15+ algorithms** (Random Forest, XGBoost, LightGBM, CatBoost...) with hyperparameter tuning (GridSearch, Optuna, BayesSearchCV); final **Stacking** model reaching **R^2 = 0.731** and **MAE = 69.01 kg**, integrated into a **PyQt5** desktop interface for decision support.

Full Stack Developer Intern Jul. – Sep. 2024
Tsunami IT – Remote (Canada)

- Developed a real-time multiplayer Scrabble application with **Angular**, **ASP.NET API** and **Pusher**.
- Agile collaboration with **Azure DevOps**, Git version control and cloud deployment.

ACADEMIC AND PERSONAL PROJECTS

Big Data Climate Prediction & Hybrid AI Jan. 2026
Real-time alert system built on a streaming pipeline (**Kafka, Spark**). **GRU** ($R^2 = 0.85$) and **XGBoost** (F1-Macro = 0.67) models validated by an **ontology**. **Angular/FastAPI** application.

AI & LLM Automation (n8n + Google Gemini) Dec. 2025
HR process automation pipeline: a CV (PDF) submitted through a web form is extracted, then analyzed by **Google Gemini** (two parallel AI calls) to generate a profile summary sent to the recruiter on **Telegram**, and a **JSON** extraction of skills/experience via *prompt engineering*, stored in **Google Sheets**.

Facial Segmentation with Deep Learning Jan. 2025
Comparative study of three architectures (**U-Net**, **SegNet**, **PSPNet**) on the **LaPa** dataset (22k images, 11 facial classes), with **SegNet** achieving the best score, **0.85 mean Dice** (**0.75 mean IoU**) on foreground classes. Robustness analysis against facial occlusions (TensorFlow, Keras).

TECHNICAL SKILLS

Data Science & AI	Python, Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, XGBoost, LightGBM, OpenCV, spaCy
Generative AI & LLMs	Prompt engineering, RAG, instruction fine-tuning (FLAN-T5), Hugging Face Transformers, Ollama (Qwen), Google Gemini, Claude
Data Engineering, BI & Databases	Spark, Kafka, Hadoop, Hive, PostgreSQL, MySQL, SQL Server, MongoDB, Power BI
Development, APIs & Automation	FastAPI, Spring Boot, .NET Web API, Angular, React, PyQt5, REST API, n8n
Languages	Python, C#, Java, C, SQL, Shell Scripting, Scala
DevOps, Cloud & Tools	Git/GitHub, GitLab, Docker, Kubernetes, Azure
Methodologies	Agile/Scrum, CRISP-DM, UML, MERISE

CERTIFICATIONS

Certifications (Exams)	Data Scientist Associate (DataCamp) *(Nov. 2024)* • SQL Associate (DataCamp) *(Sep. 2024)*
Tracks & Courses	Associate AI Engineer (DataCamp) *(Nov. 2025)* • Data Analyst Career Track (365 Data Science) *(Nov. 2024)* • Power BI Fundamentals (DataCamp) • Data Analysis with Python (IBM) • Generative AI with Large Language Models (DeepLearning.AI & AWS, Coursera, in progress) • Retrieval Augmented Generation (RAG) (Coursera, in progress)

ADDITIONAL INFORMATION

Languages: Arabic (Native), French (Professional), English (Professional)
Soft Skills: Problem solving, Analytical mindset, Rigor
Interests: Graphic Design, Video Editing

Activities & Involvement:

- **Digital Marketing Lead** – ENACTUS (2022-23).
- **Organizing Committee** – Dataverse (2023).