

# Anis Guechtouli

✉ guechtoulianiss7@gmail.com ☎ +213 555648000 📍 Algiers, Algeria 🌐 /in/anis-guechtouli 🗣 /GuechtouliAnis

## PROFILE

Data Scientist with hands-on experience building **LLM-based chatbots**, vision pipelines, and large-scale ML models. Delivered an **HR assistant at Djezzy** using **LangChain**, **RAG pipelines**, and fine-tuned models via **Unsloth**. Also developed a **PySpark**-based xG model on **12M+ rows** and built vision projects with **OpenCV**.

## EXPERIENCE

**Data Analyst Intern – CAAR** 🌐 02/2023 – 07/2023 | Algiers, Algeria

- Built a **Streamlit** dashboard for a car insurance agency, enabling exploration of **sales**, **sinistres**, and **client KPIs**
- Queried structured data from **Oracle DB** using **SQL** to power visual analytics and trend analysis
- Added **dynamic filters** (time, category, distribution) and password protection to support secured, interactive exploration
- Enabled **analysis of risk patterns** such as accident frequency by driver age, supporting more informed client segmentation

**Data Scientist Intern – Djezzy** 🌐 11/2024 – 06/2025 | Algiers, Algeria

- Developed an internal **HR chatbot** that handled **300+ recurring employee queries**, showing strong automation potential
- Designed and implemented a **full RAG pipeline** using Flask API and React, including document preprocessing, chunking, and synthetic Q&A generation stored in ChromaDB
- Fine-tuned** Qwen2.5 3B, Qwen3 4B, and LLaMA3.2 3B using **Unsloth** to improve response accuracy and latency

**Business Analyst – SIG Service** 🌐 06/2025 – 08/2025 | Algiers, Algeria

- Supporting **two large-scale data migration** projects (~100GB each) from Odoo 8 to Odoo 13 using **Apache NiFi**
- Debugging and maintaining NiFi flows for high-volume data pipelines
- Analyzing schemas, remapping fields, and ensuring consistency across **PostgreSQL** sources
- Authored **full technical documentation** for handover and audit readiness, streamlining collaboration with engineering and business teams

## EDUCATION

**Bachelor's Degree in Computer Science**, 09/2021 – 07/2023 | Algiers, Algeria  
*University Of Algiers 1 - Benyoucef Benkhedda* 🌐

**Master's Degree in Data Science and Analytics**, 09/2023 – 06/2025 | Algiers, Algeria  
*University Of Algiers 1 - Benyoucef Benkhedda* 🌐

## PROJECTS

**Iris Recognition System** 🌐

- Designed a biometric system for iris recognition using **SIFT** and **OpenCV**.
- Built a complete pipeline: preprocessing, feature extraction, visual matching.
- Deployed a **Streamlit** interface for image upload and individual identification.

**Exploratory Data Analysis in Retail** 🌐

- Performed **exploratory analysis** on a retail dataset of **9,994 sales**, identifying key issues such as **unprofitable products** and **excessive discounts** in specific states.
- Built **visualizations**, a **dashboard**, and a detailed **report** to communicate findings and guide strategic decisions.

**Expected Goals (xG) Model** 🌐

- Collected and preprocessed **12 million** lines of data using **PySpark** for scalable data cleaning and transformation
- Engineered spatial and contextual features including **shot distance**, **angle**, **footedness**
- Trained a **logistic regression model** achieving  **$R^2 = 0.83$**  and a **~98% match** with StatsBomb's official xG values

**Telecom Churn Prediction** 🌐

- Developed a **Streamlit app** with real-time churn predictions and confidence scoring.
- Trained and compared **6 ML models**, reaching an **accuracy of 80.48%** and **ROC AUC = 0.84** with **Gradient Boosting**.
- Engineered **optimized features** through encoding, correlation checks, and custom flags.

## SKILLS

- Languages** : Arabic (Native), English (Advanced), French (Intermediate)
- Programming** : SQL, Git, Python (PySpark, Pandas, Numpy, Matplotlib, Plotly, Streamlit)
- AI/ML Tools**: LangChain, Unsloth, ChromaDB, OpenCV, Scikit-learn
- Data Engineering**: Apache NiFi, Apache Spark, PostgreSQL, Oracle, SQLite
- Visualization Tools**: Excel, PowerBI, Tableau
- Soft Skills**: Technical writing, Documentation, Presentation (PowerPoint, Canva, Markdown)