

# Anis Guechtouli

✉ guechtoulianiss7@gmail.com ☎ 0555648000 📍 Mahelma, Alger, Algérie 🌐 /in/anis-guechtouli/ 📺 /GuechtouliAnis

## PROFIL

Data Scientist avec une expérience concrète en **automatisation (IA, NLP)**, **analyse visuelle** et **optimisation de données**. Capable de développer des solutions de bout en bout : collecte, traitement, modélisation et interface utilisateur (**Streamlit**, **PowerBI**). À l'aise pour transformer des besoins métier en produits data opérationnels.

## EXPÉRIENCE

**Stagiaire Analyste de Données – CAAR** 02/2023 – 07/2023 | Alger, Algérie

- Construit un tableau de bord Streamlit pour une agence d'assurance auto, explorant ventes, sinistres et KPIs clients
- Interrogé des données Oracle via SQL pour alimenter l'analyse visuelle et les tendances
- Ajouté filtres dynamiques (temps, catégorie, répartition) et mot de passe pour exploration sécurisée et interactive
- Permis l'analyse de risques comme la fréquence des accidents par âge, améliorant la segmentation client

**Stagiaire Data Scientist – Djezzy** 11/2024 – 06/2025 | Alger, Algérie

- Développé un **chatbot RH interne** gérant **300+** requêtes récurrentes, montrant un fort potentiel d'automatisation
- Conçu et implémenté un pipeline **RAG** complet via **Flask API**, incluant le prétraitement de documents, le découpage en segments, et la génération de Q&A synthétiques stockées dans **ChromaDB**
- Affiné **Qwen3 4B** et **LLaMA3.2 3B** avec **Unsloth** pour améliorer précision et latence des réponses

**Business Analyst - SIG Service** 06/2025 – 08/2025 | Alger, Algérie

- Supporté projets de migration d'Odoo 8 vers Odoo 13 avec **Apache NiFi**
- Débogué et maintenu des flux NiFi pour des pipelines de données à fort volume
- Analysé schémas, remappé champs et assuré la cohérence entre sources **PostgreSQL**
- Rédigé **documentation technique complète** pour passation et audit, facilitant la collaboration équipes IT et métier

## ÉDUCATION

**Licence Systèmes Informatiques** 11/2020 – 07/2023

Université d'Alger 1 - Benyoucef Benkhedda

**Master Analyse et Sciences de Données** 09/2023 – 06/2025

Université d'Alger 1 - Benyoucef Benkhedda

## PROJETS

**Système de Reconnaissance de l'Iris**

- Conception d'un système biométrique basé sur **SIFT + OpenCV** pour la reconnaissance par iris.
- Pipeline complet : prétraitement, extraction de caractéristiques, matching visuel.
- **Déploiement d'une interface Streamlit** permettant de charger une image et identifier l'individu.

**Application d'Optimisation – Sac à Dos (Knapsack)**

- Modélisation d'un problème d'optimisation sous contrainte via **programmation mixte entière** (GEKKO).
- Développement d'une **interface Streamlit interactive** permettant aux utilisateurs de tester différents scénarios d'allocation.
- Mise en avant de la logique décisionnelle et de la visualisation dans des situations à ressources limitées.

**Analyse Exploratoire des Données dans le Commerce de Détail**

- Réalisé une **analyse exploratoire** sur un dataset retail de **9 994 ventes**, identifiant des problèmes clés comme des **produits non rentables** et des **remises excessives** dans certains états.
- Créé des **visualisations**, un **tableau de bord** et un **rapport détaillé** pour communiquer clairement les résultats et orienter les **décisions stratégiques**.

**Modèle Expected Goals (xG)**

- Collecte de données gratuites StatsBomb, création d'un dataset de **12 millions de lignes**.
- Prétraitement et nettoyage avec **PySpark** pour garantir qualité et cohérence.
- Ingénierie de **7 caractéristiques** clés : distance, angle, pied préféré...etc
- Modèle de **régression logistique** atteignant **~98 % matching rate** avec le modèle de StatsBomb et **R<sup>2</sup> de 0,83**.

## COMPÉTENCES

- **Langues** : Arabe (langue maternelle), Anglais (avancé), Français (professionnel)
- **Programmation** : SQL, Git, Python (PySpark, Pandas, Matplotlib, Scikit-learn, LangChain, Streamlit)
- **Gestion des données** : Apache NiFi, Apache Spark, SQLite, Oracle, PostgreSQL, DBeaver
- **Visualisation et reporting** : Excel, PowerBI, Tableau
- **Soft Skills** : Rédaction technique, Documentation, Présentation (PowerPoint, Canva, Markdown)