Idriss Khattabi

E-mail: drisskhattabi6@gmail.com | Phone: +212 6 35 24 57 97
Portfolio: drisskhattabi6.github.io/id-kh
LinkedIn: linkedin.com/in/idriss-khattabi-b3a266235/
GitHub: github.com/drisskhattabi6

About me

Data Scientist with a Master's in AI & Data Science, specialized in building machine and deep learning models, building data pipelines, predictive modeling, AI agent development, and LLM integration. Proficient in Python, SQL, with hands-on experience in feature engineering, model selection, hyperparameter tuning, and performance evaluation. Analytical, and passionate about creating robust models that solve real-world problems.

Education

M.S.T in Artificial Intelligence and Data Science	Oct 2023 – July 2025
Faculty of Sciences and Technologies of Tangier	Tangier, Morocco
BS in Computer Science and Mathematics	Oct 2019 – July 2023
Faculty multidisciplinary of Larache	Larache, Morocco

Experience

AI / Data Engineer Intern, TNA Consulting – Casablanca   Feb 2025 – July 2025

- Designed and developed an AI SaaS platform to support financial institutions in interpreting and leveraging official central bank publications & documents across Africa.
- Engineered an automated and scalable **Data pipeline** for scraping & collecting documents from websites, cleaning, transforming & chunking documents, and loading into a Vector database.
- Built an intelligent **chatbot** using RAG Technique capable of answering user questions on African Central Bank documents.
- Used Tools : **Python, Marker, Streamlit, Apach Airflow, ETL Pipeline, Flask, Ollama, Faiss, RAG, LLM, Selenium, Docker**

Research Assistant, Faculty of Poly-disciplinary – Larache   March 2023 – June 2023

- Designed and implemented an optimization system to improve the consistency of pairwise comparison matrices in the **AHP**, leveraging **GA** and **PSO**, with Developing a Django web application that makes it easier to generate AHP matrices.
- Co-authored **three** research papers with **Dr. Chakir Tajani**, papers examples :
  - A Metaheuristic Approach to Improve Consistency of the Pairwise Matrix in AHP
  - Generating a Set of Consistent Pairwise Comparison Test Matrices in AHP using PSO

Soft Skills

Analytical thinking and problem-solving – Fast learning – Teamwork and collaboration – Agility and Adaptability – Curious for new technologies

Technical Skills

- **Programming & Scripting:** Python, SQL, Java, JavaScript, C, Bash, Solidity
- **Data Analysis & Visualization:** NumPy, Pandas, Matplotlib, Seaborn, Plotly, Power BI, Power Query, Excel
- **Big Data & Distributed Systems:** Apache Spark (PySpark), Apache Kafka, Hadoop, Apache Airflow
- **Machine & Deep Learning & NLP:** Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLflow, NLTK, SpaCy
- **Generative AI & LLMs:** RAG, Prompt Engineering, Fine-tuning, LangChain, LangGraph, LlamaIndex, n8n
- **Databases:** MySQL, SQL Server, PostgreSQL, MongoDB, Supabase, IPFS, FAISS, Chroma, Qdrant, Milvus
- **Web Development:** HTML, CSS, Tailwind CSS, Javascript, ReactJS/NextJs, Django, Flask, FastAPI, Streamlit
- **DevOps & Tools:** Linux, Git/GitHub, Docker, Microsoft Azure, Selenium

Languages

Arabic: Native proficiency | **English:** Professional (B2) | **French:** Professional (B2)

Certifications

Associate Data Scientist in Python – DataCamp   **Associate AI Engineer for Data Scientists –** DataCamp
Microsoft Azure Fundamentals (AZ-900) – DataCamp   **Associate Data Engineer –** DataCamp

Projects

Age, Gender & Emotion Detection using CNN   [GitHub Repo]

- Detecting person Age, Gender, and Emotional state from facial images using CNN.
- Tools Used: Python, OpenCV, Tkinter, TensorFlow/Keras, CNN Algorithm

Real-time Twitter Post Sentiment Analysis   [GitHub Repo]

- Created a real-time system to predict sentiment from Twitter posts by streaming posts using Kafka and predicting in real-time PySpark using Machine Learning, and visualizing data in real time
- Tools Used: PySpark, Kafka, Docker, Django, MongoDB, ChartJS

Language Translation   [GitHub Repo]

- Developed a language translation system (English → Arabic) by fine-tuning the BERT model and deploying it using FastAPI.
- Tools Used: Python, FastAPI, Hugging Face, Pytorch, BERT

Data Analysis & Machine Learning Desktop Application   [GitHub Repo]

- A Python desktop app built with CustomTkinter that enables users to analyze qualitative and quantitative datasets and apply multiple machine learning algorithms through an intuitive GUI.
- Tools Used: Python, Tkinter / CustomTkinter, Scikit-learn, Pandas, Matplotlib

Fake News Detection   [GitHub Repo]

- Fine-tuning BERT model to detect fake news articles, and Deploying the trained model via a Flask web app for real-time prediction.
- Tools Used: Python, Flask, Pytorch, BERT

Obstacle Detection & Alert System for the Visually Impaired   [GitHub Repo]

- A real-time obstacle detection and audio alert system designed to assist visually impaired individuals using deep learning and automation tools.
- Tools Used: Python, Flask, PyTorch, MobileNetV2, CNN, OpenCV, Node-RED, FFmpeg

AI Scraper   [GitHub Repo]

- An interactive tool to scrape, process, and structure web data using local or cloud-based LLMs.
- Tools Used: Python, Selenium, Crawl4AI, Streamlit, Ollama, Sambanova, Pandas

Text To SQL Converter   [GitHub Repo]

- An interactive tool to convert user text to SQL query using AI Agent, with getting data from SQL DB and Visualizing using Interface.
- Tools Used: Python, LangGraph, Streamlit, Pandas

Moroccan Meal Recipes Chatbot   [GitHub Repo]

- An AI chatbot that uses RAG to deliver authentic Moroccan recipes with ingredients, steps, and optional images, supporting both local and cloud LLMs.
- Tools Used: Python, Streamlit, LangChain, Ollama, Sambanova, Faiss DB

Web Development with Django

- Marketplace Website - Chat With LLM - Blog Website - Note App - Room Chat App