

Rida Al Hajj Ali

Beirut, Lebanon | +96181248452 | ridahajali8@gmail.com | [LinkedIn](#) | [GitHub](#) | [Medium](#)

Summary:

Junior Data Scientist and AI specialist Proficient in Python, SQL, machine learning, artificial intelligence, and data visualization, skills developed through both academic projects and hands on experience. Ready to use my skills to solve real-world problems.

Experience:

IBKS – AI and Data Analytics Specialist (On-site)

Full time:

October 2025 – Present

- Implemented an advanced natural language-to-SQL tool connected to multiple databases using a local LLM (Qwen 3.5-9B) and Vanna, Vector database (Qdrant), advanced database schema storing, SQL postprocessing and validation to enable dynamic data querying.
- Developed and optimized T-SQL queries to train, test the LLM for accurate sql generation
- Implemented a Feedback Loop feature, built by storing, analyzing, and retrieving the user-based feedbacks to improve the NL-to-SQL tool.
- Migrated AI workflows from Gemini to a fully local LLM setup using Llama.cpp, improving privacy, and scalability.
- Implemented an AI guardrail for safe inputs and outputs, built based on Nvidia NeMo Guardrails
- Finetuned different small language models for the invoices, using modern frameworks like unsloth and choosing the best model (like qwen2.5-coder-3b and phi-3.5-mini)
- Worked on Designing pipeline for accurate PDF-to-JSON tool, optimized for complex invoices

Internship:

June 2025 – October 2025

- Developed an AI-powered RAG chatbot capable of retrieving and synthesizing answers from documents (PDF, CSV, DOCX, PPTX, etc..) and images using OCR (Tesseract).
- Implemented FastAPI for the Chabot's backend services, ChromaDB for vector similarity search, Redis for low-latency response caching, Gemini as the LLM, and Mixbread-AI for embedding.

Siren Analytics – AI / Data Science Internship (On-site)

January 2025 – May 2025

- Contributed to a national project mapping quarries across Lebanon using satellite imagery.
- Developed and trained a YOLOv9 model to detect new quarry formations achieving high accuracy with 85% on segmentation and 80% for detection.
- Implemented a full training pipeline using ResNet50 to monitor quarry depth changes.
- Built real-time, chat-based voice AI assistants on Gradio and Telegram, supporting English, French, Spanish, and Lebanese Arabic.

Education:

University of Saint Joseph (USJ) – Metn, Lebanon

Bachelor's Degree in Data Science | GPA 3.5/4

September 2022 – May 2025

Skills:

Technical Skills: Machine Learning, Artificial Intelligence, Data Analysis, Computer Vision, Natural Language Processing, Prompt Engineering, Fine tuning

Programming Languages: Python, SQL, R, C++

Tools & Libraries: TensorFlow, PyTorch, Numpy, Docker, Git, Unsloth, HuggingFace, Pydantic, llama.cpp

Soft Skills: Teamwork, Critical Thinking, Ownership of Projects, Time Management , Documentation

Achievements:

Achieved 2nd among more than 100 participants in Google Developers hackathon done at Lebanese University, for building a [RAG chatbot and analyzing Real-Estate data](#)

Open Source:

Contributed to multiple repositories by identifying and reporting issues or assisting developers, particularly for top organizations such as Microsoft ([1,2](#)), Google DeepMind ([1](#)), LiveKit ([1](#)), and NumPy ([1](#)).

Languages:

Arabic: Mother Tongue | English: Proficient