

Rayan Al Abdallah

☎ Telephone: +961 70925528 ✉ Email: rayan3abdallah@gmail.com 📍 Address: Bierut, Lebanon  [Rayan Al Abdallah](#)  [rayan abdallah](#)

AI/ML Engineer & Data Science Intern specializing in **LLMs, RAG systems, and NLP**, with hands-on experience building **end-to-end AI pipelines**, vector databases, embeddings, and applied AI solutions. Skilled in designing real-world AI applications across education, healthcare, wearables, and VR.

Experience

Data Science Intern — Auvia Group (Oct 2025 – Present)

- Working on applied AI/ML projects across **LLMs, NLP, data engineering, and predictive analytics**.
- **Proactive LLM System:** Building an autonomous RAG-based agent with LangChain + vector DBs for proactive information retrieval and contextual reasoning.
- **Smartwatch for Students:** Supporting ML-based student activity monitoring and dashboard visualizations.
- **Neuroscience (EEG/BCI):** Exploring signal preprocessing and baseline modeling for cognitive patterns.
- **Tech Stack:** Python, LangChain, OpenAI API, Chroma/FAISS, scikit-learn, TensorFlow, Hugging Face, Pandas, Git.

Projects

BookBuddy — AI Document Assistant (2025)

A full RAG-powered AI system built completely from scratch for conversational document Q&A.

Tech Stack: Flask (real-time streaming), custom UI (no frameworks), Ollama (llama3.1:8b), ChromaDB, Sentence-Transformers

Key Contributions:

- Implemented PDF/TXT/DOCX ingestion with intelligent chunking.
- Built vector embeddings + semantic search using ChromaDB.
- Added real-time streaming responses and persistent chat history.
- Enabled multi-document Q&A and automatic summarization.
- Demonstrated full LLM engineering: ingestion → embeddings → search → LLM reasoning.

AI & ML in Cancer Imaging — University Project (Nov 2024–Feb 2025)

- Studied early breast cancer detection using the Breast Cancer Wisconsin dataset.
- Trained and compared Logistic Regression, Random Forest, and a simple Neural Network using accuracy, precision, recall, and training time.
- Presented results with references to real-world research and discussed clinical trade-offs.

Medical Insurance Costs — Linear Regression (Nov 2024–Feb 2025)

- Ran Simple and Multiple Linear Regression to estimate charges from demographic/health features.
- Preprocessed with one-hot encoding, feature selection, and train/test split (pandas, scikit-learn, Matplotlib).
- Evaluated using R^2 and MSE and visualized model behavior to support pricing insights.

Auto-Advising System for University Course Planning — Senior Project (Feb 2025–May 2025)

- Designed and developed a web-based platform with HTML, CSS, JavaScript, PHP, and MySQL.
- Integrated OpenAI's GPT API to suggest courses based on academic progress and prerequisites.
- Created role-based dashboards with secure authentication and real-time advising logic.
- Focused on reliability, correctness of prerequisites, and reducing manual review.

Sentiment Analysis for User Reviews — College Project (Feb 2025–May 2025)

- Classified Trustpilot reviews using lexicon methods (VADER), a fine-tuned BERT model, and GPT-based assessment.
- Built text preprocessing with NLTK and spaCy (tokenization, lemmatization, POS tagging).
- Compared pipelines, visualized distributions, and extracted key phrases via TF-IDF and bigrams.

Core Skills

Programming: **Python, JavaScript, Java, SQL, Dart, PHP**

AI/ML: **scikit-learn, TensorFlow, PyTorch, Hugging Face, LangChain, OpenAI API, RAG systems**

NLP: **embeddings, BERT, emotion/sentiment detection, text preprocessing, TF-IDF, spaCy, NLTK**

Data: **Pandas, NumPy, preprocessing, feature engineering, visualization**

Tools: **Git, GitHub, Jupyter, Colab, Matplotlib, ChromaDB, FAISS**

Languages: **Arabic (Native), English (Advanced)**

Certificates

- AI & LLM Engineering Mastery — Udemy (2025)
- AI & Digital Marketing Scholarship Program (2025)
- Data Analysis — Simplilearn
- Cisco: ITN + SRWE
- LinkedIn: Leading a Team Through Change

Education

- Lebanese International University (LIU) — Bachelor in Computer Science (2021–2025)
- Amir Shakib Arsalan High School — Lebanese Baccalaureate in Life Science (2020)