

Rayan Al Abdallah

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

AI/ML Engineer & Data Science Intern at Auvia Group specializing in LLMs/RAG and NLP with Python, LangChain, and Hugging Face; shipped RAG agents and applied AI across wearables, VR, and academic advising..

Experience

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

- Working on applied AI/ML projects across LLMs, NLP, and data engineering within cross-functional teams.
- Proactive LLM: building a retrieval-augmented generation (RAG) pipeline with LangChain and vector databases.
- Oil Simulation (VR/AR): prototyping predictive analytics for operational scenarios and safety use cases.
- Smart Watch for Students: designing ML-driven monitoring and simple data-visualization dashboards.
- Neuroscience: experimenting with EEG/BCI preprocessing and baseline modeling for signal patterns.
- Stack: Python, LangChain, OpenAI API, FAISS/Chroma, scikit-learn, TensorFlow, Hugging Face, Pandas, Git.

Projects

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, Java, JavaScript, HTML, CSS, SQL, Dart
- **ML/AI:** scikit-learn, TensorFlow, PyTorch, Hugging Face Transformers, OpenAI GPT API, LangChain
- **NLP:** Tokenization, embeddings, sentiment/emotion detection, BERT, VADER, TF-IDF, spaCy, NLTK
- **Data Handling:** Pandas, NumPy, data cleaning, preprocessing, feature engineering
- **Computer Vision:** Image preprocessing, model evaluation
- **Tools:** Jupyter Notebook, Google Colab, Git, GitHub, Matplotlib, Seaborn
- **Languages:** Arabic (Native), English (Advanced)

Certificates

- Udemy — AI & LLM Engineering Mastery: GenAI, RAG Complete Guide 2025
- AI & Digital Marketing Scholarship Program (4/2025 – 13/2025)
- Data Analysis — Simplilearn (6/2024 – 10/2024)
- Introduction to Networks — Cisco
- Switching, Routing, and Wireless Essentials — Cisco
- Leading a Team Through Change — LinkedIn

Education

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