

# Ahmad Mersaghian

Riverside, CA | ahmad\_mersaghian@yahoo.com | LinkedIn | GitHub | Personal website

## SUMMARY

- Data Scientist / ML Engineer with experience in multimodal ML systems, graph neural networks, and production APIs. Delivered **32%** gains in MS/MS spectrum analysis; built RAG pipelines and multi-agent systems. Proficient in PyTorch, Transformers, PySpark, FastAPI, and Google Cloud.

## EDUCATION

**University of California, Riverside** — M.S., Computational Data Science December 2025  
(GPA: 4.00/4.00)

**Loyola Marymount University** — B.S., Biology; Minor, Computer Science May 2023  
(*magna cum laude*; GPA: 3.85/4.00)

## PROFESSIONAL EXPERIENCE

**Graduate Research Assistant**, University of California, Riverside — Riverside, CA Oct 2024 – Present

- Evaluated 5 GNN/Transformer models for MS/MS spectrum prediction; achieved **10%** higher modification-site accuracy vs. baselines trained on **400k+ NIST-20 spectra**.
- Built Product-of-Experts ensemble with custom metrics (neighborhood precision, delta-score consistency), improving average accuracy by **12%**.

**Data Science Fellow**, University of California, Riverside — Riverside, CA Jun 2025 – Aug 2025

- Built multi-agent system (Google ADK + FastAPI) for **2,000+ students**; integrated **10k+ records** to generate conflict-free schedules, reducing registration time by **40%**.
- Developed React/Next.js interface with real-time calendar and chatbot, improving task completion efficiency by **25%**.

**Undergraduate Research Assistant**, Loyola Marymount University (GRNsight) — Los Angeles, CA Jan 2021 – May 2023

- Processed gene expression time-series with STEM; added datasets to GRNsight database and improved UI responsiveness by **15%**.

## PROJECTS

**AI Student Tutor (Retrieval-Augmented Generation)** GitHub

- Implemented a **Retrieval-Augmented Generation (RAG)** pipeline using pretrained LLaMa and DeepSeek models to answer queries based on uploaded PDFs.
- Designed and implemented a **perplexity-based system** to select the best model based on query characteristics while allowing manual model selection in the UI, reducing average latency by **25%**.
- Found that for short queries LLaMa was chosen two-thirds of the time, while also exhibiting shorter response time and lower energy use on average.
- Stack: PyTorch, Transformers, FAISS, Flask/Streamlit, REST.

**End-to-End Data Pipeline (PySpark, PostgreSQL, Flask)** GitHub

- Built a scalable **ETL pipeline** with PySpark to clean and transform three large datasets for efficiency.
- Designed PostgreSQL database schema with indexing to cut query latency by **35%**; shipped a Flask API with web interface for querying and displaying analytical results.

## TECHNICAL SKILLS

**Programming:** Python, SQL, PySpark, Bash, JavaScript, HTML/CSS, Git

**Machine Learning:** PyTorch, scikit-learn, Hugging Face Transformers, Graph Neural Networks, Retrieval-Augmented Generation (RAG), Product-of-Experts

**Data & Platforms:** Pandas, NumPy, RDKit, FastAPI, Flask, PostgreSQL, Docker, React/Next.js, Google Cloud

**Soft Skills:** Problem Solving, Collaboration, Communication, Organization, Time Management

## AWARDS

Data Science Summer Fellowship, UC Riverside (2025) • Sigma Xi Scientific Research Honor Society, LMU (2023) • Rains Undergraduate Research Fellow, LMU (2022–2023) • William McLaughlin Scholarship, LMU (2021–2023) • SOAR Fellowship, LMU (2022) • Dean's List, LMU (2019–2023)