MOHAMED MARTINI

NH | 617-955-4276 | leverai@leverai.tech | LinkedIn | Website

SUMMARY

SWE with proven industry skills in developing and deploying real-world ML solutions, cloud-native backend applications and microservices design. Passionate about architecting scalable, reliable, and maintainable data intensive applications that leverage AI and timeless engineering principles.

SKILLS

- **Programming Languages:** Expert in Python, SQL. Familiar with C/C++, Flutter.
- **AI/ML:** LangChain, LangGraph, RAG, TensorFlow, PyTorch, NumPy, Pandas, Hugging Face.
- **Data:** ETL, Apache Beam, BigQuery, SQL/NoSQL, GCS, Kafka, PySpark, Redis. Familiar with Airflow.
- **Backend:** REST (FastAPI), Microservices, Websockets, gRPC,, Auth/Authz, Load Balancing, Rate Limiting, Scalability.
- **Containerization & CI/CD:** Docker, Docker Compose, GitHub Actions, Python workspace management.

PROFESSIONAL EXPERIENCE

Lever AI **Jul 2025 - Present**

Founder & Principal Engineer

- Founded a consulting and SaaS development LLC specialized in AI-powered, data-driven software solutions.
- Showcase Project: Designed and shipped a trivia-like estimation game to Android and iOS featuring:
  - **AI-native ETL pipeline**: Used LLM APIs for question generation, enrichment, embedding for deduplication, and grounded answering. Used **semaphores** for mitigating API rate limits. Stored questions in **Postgres** with **pgvector**.
  - **High consistency** for a multi-player game mode using atomic **Firestore** transactions.
  - **Low-latency** fair question assignment using optimized schema design, indexing, and SQL queries.
  - **Full observability** via end-to-end distributed tracing with **OpenTelemetry**, correlating Flutter frontend events to FastAPI backend spans via Cloud Trace and Cloud Logging.
  - **GitFlow CI/CD pipelines** with automated unit, integration, and smoke testing, and dev/prod deployments.
  - **Leveraged AI coding agents** (Claude, Gemini) to accelerate development while maintaining full architectural oversight.

Pison Technology **Oct 2022 - Jul 2025** (2 years, 10 months)

Machine Learning Engineer III

- Promoted from MLE to MLE III in April 2025.
- Ran **1,200+ TensorFlow experiments for sleep stage classification** in TensorFlow, **improving the baseline macro F1-score by 15%**. Mitigated class imbalance with **Focal and Contrastive Loss**. Designed a **custom attention layer for feature mining**, improving performance while reducing the size of the feature set by an order of magnitude.
- Developed real-time biosignal gesture detection models. Wrote Cythonized extension modules, and vectorized bottleneck functions, significantly improving the speed of online prediction and offline modeling.
- Led a “Hackaweek” project with two colleagues to implement a state-of-the-art **domain-invariant classification network** with adversarial training in PyTorch, following the approach outlined in its original manuscript.
- Led the adoption of **BigQuery** for analytics pipelines. Decided **partitioning and clustering** for efficient queries.
- Deployed and monitored three services using Google Model Registry and **Cloud Functions**.
- **Improved the speed of a batch ETL pipeline** using Apache Beam by more than **190 folds**.
- Combined standalone Python repos into a single **uv workspace**, allowing teams on different operating systems to seamlessly cooperate, reducing the size of deployed containers, and streamlining code installation, testing, and deployment.
- Maintained cross-team communication to ensure consistency and optimal integration of ML solutions.

University of Massachusetts Lowell **Jan 2021 – June 2022** (1 year, 1 month)

Research Assistant

- Implemented **Reinforcement Learning** agents and environments for cooperative multi-agent Search and Rescue tasks.
- Implemented **object detection** training and inference pipelines. Used Generative AI for data augmentation.

EDUCATION

University of Massachusetts Lowell

- *M.S. Computer Engineering* | *GPA 4.0* **2022**
- *B.S. Electrical Engineering* | *GPA 3.8* **2021**