

Marc'Antonio Lopez

(+39) 3755742433 | marcantonio.lopez@ieee.org | linkedin.com/in/marc-antonio-lopez
github.com/LookUpMark | kaggle.com/marcantoniolopez | lookupmark.github.io

SUMMARY

Computer Engineer and M.Sc. candidate in Artificial Intelligence at Politecnico di Torino, with hands-on experience across **Machine Learning, Generative AI**, and **Agentic AI**. I build end-to-end AI systems — data pipelines, LLM fine-tuning, and agent orchestration with **LangGraph** — in Python (Pandas, Scikit-Learn, PyTorch, Hugging Face), and I apply AI-assisted coding (Claude Code, GitHub Copilot, OpenCode) within enterprise workflows. Currently a Graduate Thesis Intern at Data Reply developing production-oriented GenAI solutions, I aim to grow as a Machine Learning Engineer in large-scale production environments.

TECHNICAL SKILLS

Programming: Python, SQL
ML & Data: Scikit-Learn, PyTorch, NumPy, Pandas, Data Pipelines, Model Evaluation
Generative AI & LLMs: Transformers, Hugging Face, RAG, Prompt Engineering, Fine-Tuning, PEFT/LoRA, Quantization
Agentic AI: LangChain, LangGraph
Retrieval: Embeddings, Reranking, Hybrid Retrieval, Parent-Child Chunking
MLOps & Deployment: Docker, Weights & Biases (W&B)
Tools & AI-Assisted Dev: Git, GitHub, Claude Code, GitHub Copilot, OpenCode
Big Data: Hadoop, PySpark (*academic exposure*)
Languages: Italian (Native), English (C1 Advanced)

PROJECTS

- Multi-Turn Self-Corrective RAG (SemEval 2026 Task 8)** 2026
Python, LangGraph, Llama 3.1, RAG
– Built a Self-Corrective RAG system with LangGraph for reliable, well-grounded answers across multi-turn dialogues, combining a hybrid retrieval pipeline (Parent-Child chunking, BGE-M3 embeddings, Cross-Encoder reranking) with 4-bit NF4 quantization of Llama 3.1 8B for efficient offline inference on consumer hardware. Code: **GitHub**.
- SM-SIP: Multilingual Abstractive Summarization** 2026
Python, LLMs, Hugging Face, PEFT
– Developed a controllable abstractive summarization framework with semantic supervision, building a multilingual (IT/EN) pipeline that integrates Llama and Qwen with token-classification modules to detect and mitigate hallucinations; published the open-source package and LoRA adapters (PEFT) on the Hugging Face Hub. Code: **GitHub**.
- Explainability in AI: Concept Discovery for Medical VLMs** 2026
Python, Sparse Autoencoders
– Engineered a pipeline for unsupervised concept discovery in medical Vision-Language Models via Sparse Autoencoders, pairing a concept-naming module with an LLM-as-a-judge system to quantify concept faithfulness. Code: **GitHub**.
- DYLEM-GRID: Gesture Recognition** 2024
Python, PyTorch, LSTM, Transformers
– Designed and evaluated a BiLSTM with attention and an encoder-only Transformer (PyTorch) for dynamic hand-gesture classification, optimizing high-dimensional feature extraction on the DYLEM-GRID dataset. Code: **GitHub**.

EXPERIENCE

- Graduate Thesis Intern** February 2026 – Present
Data Reply *Turin, Italy*
– Research and design production-oriented Generative AI solutions for metadata management and semantic discovery, combining Foundation Models with Data Engineering to automate data governance. Architect scalable AI workflows (LLM-based reasoning with backend architectures) from prototyping toward deployment.
- Research Trainee in Artificial Intelligence** February 2024 – June 2024
University of Enna "Kore" *Enna, Italy*
– Developed machine-learning pipelines for dynamic gesture recognition on Leap Motion time-series sensor data, performing statistical feature analysis and contributing to data acquisition and annotation for the DYLEM-GRID dataset (400 samples, 100 participants).

EDUCATION

- Politecnico di Torino** Turin, Italy
M.Sc. in Artificial Intelligence and Data Analytics, 110/110 (expected) *September 2024 – September 2026*
– Thesis: Generative AI and Foundation Models for Automated Data Engineering and Metadata Orchestration
- University of Enna "Kore"** Enna, Italy
B.Sc. in Computer Engineering, 110/110 cum laude *October 2021 – July 2024*
– Thesis: Recurrent Neural Networks for Dynamic Gesture Recognition

CERTIFICATIONS & PUBLICATIONS

- DYLEM-GRID Dynamic Gesture Dataset** 2025
Kaggle Dataset Publication – 400 dynamic gestures, 100 participants | Kaggle
- Cambridge English C1 Advanced** 2024
Certified CEFR C1 proficiency in technical and professional English