Domenico Lacavalla

✉ domenicolacavalla8@gmail.com 🔗 domenico-lacavalla ⌥ D0men1c0 Ⓜ domenicolacavalla8
D0men1c0 🏠 D0men1c0

Experience

Leonardo Labs   Jan. 2026 - Present
AI Researcher & Applied Scientist   *Turin, Italy*
- •

IBM   Oct. 2023 - Jan. 2026
Data Scientist & Applied Scientist   *Bari, Italy*
- Implemented CPU multimodal retrieval (Grounding DINO, SBERT-Whisper) boosting speed 31% and recall to 83%.
- Orchestrated ancient map geocoding (68.6% acc) fusing GenAI correction and DBSCAN under custom validation.
- Bootstrapped a few-shot SBERT classifier achieving 73.4% accuracy via 10-shot LLM labeling; deployed via ONNX.
- Saved $1,000+/mo on 5M+ samples via a hybrid LLM/SBERT anomaly detection pipeline for sentiment analysis.
- Boosted summarization throughput 5x and cut waste 21.7% via parallelized LLM inference on A100 HPC.

Google Summer of Code (GSoC) 2024 HumanAI   May 2024 - Sept. 2024
Open Source Contributor   *United States, Remote*
- Extracted 5 baseline categories by analyzing 500k Dark Web discussions using advanced NLP models.
- Synthesized multimodal models integrating BERT (170 topics) and CLIP/Vision Transformer.
- Captured topic evolution and validated clustering using LSTMs for dynamic sentiment analysis.
- Delivered trend forecasts by deploying 8 predictive models on Hugging Face based on analytical insights.
- Read more in this blog post and explore the project on GitHub Repo .

Education

UNIBA (Università Degli Studi di Bari 'Aldo Moro')   Sept. 2023 - Present
MSC in Data Science   *Bari, Italy*
- **Relevant Coursework:** Machine Learning, Deep Learning, Data Mining, Statistics, Linear Algebra, Optimization

UNIBA (Università Degli Studi di Bari 'Aldo Moro')   Oct. 2020 - July 2023
BSC in Computer Science, graduation with 110/110 cum Laude   *Bari, Italy*
- Experimental thesis on reproducibility in recommendation systems, applied to the ClayRS codebase.

Projects

Knowledge Injection XAI | *Python, PyTorch, Spark, DINOv2, LoRA*   Jan. 2026 | Repo
- Developed a distributed framework predicting OOD robustness using only XAI metrics from clean data.
- Established a Spark pipeline integrating DINOv2, LoRA, and XGBoost to forecast model failures.
- Achieved 0.739 ROC-AUC, showing low-rank adapters resist corruptions 19% better than high-rank ones.

Text2Motion | *Python, PyTorch, Transformers, CUDA, OpenCV*   Oct. 2025 | Repo
- Developed a three-stage pipeline automating 3D animation from static meshes to text-aligned motion.
- Designed a PointNet++ model for automatic rigging and skinning with joint and weight prediction.
- Formulated a topology-aware transformer diffusion model using SBERT and T5 encodings with Geodesic+InfoNCE.
- Achieved 92% armature classification accuracy on Truebones Zoo; published in *"Text2Motion"* (Article).

Gemma Model Benchmark Suite | *PyTorch, Transformers, Scikit-learn, TF*   May. 2025 | Repo
- Authored Medium article (10min+ read) detailing framework architecture, customization & performance.
- Developed modular LLM suite covering 4+ families, 5+ tasks (MMLU), 15+ metrics (ROUGE), and custom scripts.
- Optimized for Colab T4 (4 min/500 samples, Gemma 2B 4-bit) to enable experimentation on consumer hardware.

Extracurricular Activities

Superhero Valley | *Mentee* – Selected for exclusive mentorship with Big Tech leaders   Dec. 2025
LauzHack @ EPFL | *Participant* – Prototyped vision AI assistant [Link]   Dec. 2024
Samsung Innovation Campus | *Top 25 Student* – Selected for AI & IoT program   Apr. 2023

Technical Skills

Development: Python, Java, PyTorch, TensorFlow, Hugging Face, Pandas, Scikit-learn, SQL, OpenCV, NumPy
Tools & General: AWS, Docker, OpenShift, PostgreSQL, Git, Italian (Native), English (Proficiency)