Jacopo Zacchigna

+39 3922428077 | Trieste, Italy | jaczac2002@gmail.com | github.com/Jac-Zac | linkedin.com/in/jacopo-zacchigna | jac-zac.github.io

Research Interests

Generative Models, Mechanistic Interpretability, Multimodal Learning, Computer Vision, Bayesian Inference

Publications

Application of Computer Vision to Automated Metadata Extraction from Natural History Specimen Labels: A Case Study on Herbarium Specimens https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5247575

- Preprint under review at **Pattern Recognition Letters**
- Fine-tuned *Donut-base* multimodal transformer on >40k herbarium label images, achieving **85% Tree Edit Distance accuracy**, presenting the **first** end-to-end approach.
- Designed fully **OCR-free workflows** for digitization; tracked experiments in **Weights & Biases** and released reproducible code, and guidelines for future improvements

Selected Projects

BayesianFlow — Uncertainty Estimation in Generative Models

- Proposed *firs extension* of BayesDiff uncertainty propagation to **Flow Matching** on the ODE of the vector field.
- Implemented **Last Layer Laplace Approximation** on U-Net; validated on Fashion-MNIST with interpretable pixel-wise uncertainty maps

LoRA & DoRA in TinyGrad

- Implemented *LoRA* and *DoRA* from scratch in *TinyGrad*
- Reduced fine-tuning parameter updates by >99% through low-rank adaptation in linear layers

Representation Extraction & Visualization

- Developed PyTorch hook pipeline to extract hidden layer activations and project into 2D via *UMAP* / *t-SNE*
- Visualized generalization dynamics in *Plotly* dashboards across training epochs

Education

Università degli Studi di Trieste Trieste, Italy

Master's Degree in Data Science and Artificial Intelligence *Sep 2024 — Present*

- Relevant Coursework: Probabilistic ML, Deep Learning, NLP, Advanced Topics in ML, Multi-Objective Optimization, HPC, Theory of Deep Neural Networks

University of Trieste Trieste, Italy

Bachelor's Degree in Artificial Intelligence and Data Analytics *Sep 2022 — Jul 2024*

- Thesis: **Toward automated extraction of metadata from herbarium specimen labels**
- Built ML models with Bayesian & frequentist approaches; covered mathematics and computer science fundamentals

Work Experience

AI Consultant May 2024 — Jan 2025

Obloo Ventures *Trieste, Italy*

- Mentored interns in AI research & engineering best practices
- Designed *GenAI Retrieval-Augmented Generation (RAG)* systems for automated knowledge retrieval from company data lake
- Led development of initial ML prototypes and automated Python-based data workflows

Skills

- **Programming & Tools**: Python, C, C++, R, MPI, PyTorch, HuggingFace, Git, Slurm, Linux, Docker, W&B
- **AI & Data Science**: Deep Learning, NLP, Computer Vision, Generative Models, Bayesian Inference
- **Languages**: Italian (Native), English (C1)