David Nardi

Munich, Germany • github.com/david-inf • linkedin.com/in/dn25 • david.nardi30@gmail.com

SUMMARY

AI Engineer with an Engineering Management background. Passionate about Deep Learning foundations and NLP to improve operational productivity. Experience building pipelines for large-scale research experiments. Proven ability to work in cross-functional, fast-paced, and international teams.

EDUCATION

MSc in Artificial Intelligence Engineering, University of Florence, Italy — 2023 - Apr 2026

- **Prospected Thesis**: Large-scale optimization for over-parametrized deep models.
- **Activities:** Student representative also part of the AI Engineering degree review committee.
- **Relevant Coursework:** Generative AI and Multimodal AI, NLP & LLMs, Computer Vision, Reinforcement Learning, Machine Learning Security, Deep Learning, Continuous & Combinatorial Optimization, Statistics.
- **GPA:** 29/30

BSc in Management Engineering (IT Track), University of Florence, Italy — 2019 - 2023

- **Thesis**: Machine learning applications to industrial quality control.

TECHNICAL SKILLS

Programming Languages: Python, R
Frameworks and Libraries: PyTorch, HuggingFace (transformers, datasets, accelerate, gradio), Scikit-learn, Optuna, FAISS (integrated in RAG pipelines), spaCy, LaTeX
Development and Operations: Linux, Git, GitHub, Slurm, CometML, Wandb
Languages: Italian (Native), English (C1)

EXPERIENCE

AI Research Intern, Mathematics Institute, LMU Munich, Germany — Sep 2025 - Now

Research focus: understanding why the Adam optimizer effectively pretrains language models

- Identified a research direction with insights from 10+ **deep learning optimization** papers by conducting a structured review on the SGD-Adam performance gap in LLMs pretraining.
- Engineered experimental reproducibility by successfully replicating and extending 3 core studies by building a **Python**/**PyTorch**-based pipeline with **Optuna** for HPO and **CometML** for experiment tracking.
- Enabled scalable LLMs pretraining with stable distributed runs on a **Slurm cluster (8 GPUs)** by orchestrating job submission, monitoring, automated logging, and iterative evaluation workflows.
- Tested research quality after experts' feedback from a 60+ international attendees **summer school on machine learning** by presenting current findings.

Co-Head of Business Operations, JEFLO, Junior Enterprise Florence, Italy — Mar 2024 - Jun 2025

Student-run consulting network providing hands-on entrepreneurial experience through real-world projects.

- Reduced **project planning** effort by 40% by developing a novel internal management system.
- Selected as one of 500+ participants Europe-wide for a **4-days conference** based on resume and achievements to represent the organization.
- Broadened the scope of the service portfolio with 3 new projects by training the team (10 consultants) on **dashboarding** for a total of 5 hours.

TECHNICAL PROJECTS

Slurm jobs monitor

- Engineered and deployed a real-time Slurm job-monitoring service using **Python**, **Slurm** REST API/CLI, **HuggingFace transformers** (log-summarization pipeline), enabling researchers to receive periodic updates and instant notifications on state transitions (PENDING → RUNNING → FAILED/COMPLETED).
- Reduced manual **cluster checks** by 60% and cut the average detection time for failed experiments from 10 minutes to milliseconds required for the query, improving iteration speed and resource utilization across 10 active users and 50 daily jobs.

Argument Mining with NLP

- Led **fine-tuning and model selection** across 6 DistilBERT and SBERT variants for argument-component detection, achieving a 20% F1 improvement over baseline and reducing inference latency by 60% through model optimization.
- Built and deployed and end-to-end **inference pipeline** processing 50+ medical research articles, implementing custom scoring metrics that enabled retrieval of high-argumentative-content papers improving precision by 30% over baseline.