Niccolò Ajroldi

niccolo@tue.ellis.eu
Ellis Institute Profile
Google Scholar
Linkedin
GitHub

BIOGRAPHY

My background is in mathematics and statistics, and their application to machine learning. My research focuses on optimization for deep learning and novel sequence modeling architectures. I am particularly interested in benchmarking optimization algorithms for training large language models, enhancing performance and efficiency. I very much enjoy developing machine learning codebases!

EXPERIENCE

Research Engineer - *Ellis Institute Tübingen and Max Planck Institute for Intelligent Systems* *Tübingen, DE*
Jan. 2023 - Ongoing

- Research on novel sequence model architectures and optimization algorithms with Jonas Geiping and Antonio Orvieto.
- Developed submissions to AlgoPerf, a benchmark for optimization algorithms, scoring third in the competition leaderboard.
- Joined MLCommons Algorithmic Efficiency team, developed open-source contributions to AlgoPerf API.
- Mentored research interns in the CaCTüS program, researching LLM optimization and LLM watermarking.
- Released a simple PyTorch codebase for pretraining modern language models.

AI Resident - *Meta, FAIR Labs* *Seattle, WA*
July 2022 - Aug. 2023

- Research on hypergradient methods for adaptive hyperparameter tuning with Lin Xiao.
- Investigated loss spikes in language model optimization and the Slingshot phenomenon.
- Gained experience in ML optimization, distributed training, and familiarity with vision and language model training.

ML Researcher - *U-Care Medical* *Turin, IT*
Dec. 2021 - June 2022

- Developed machine learning models to forecast Acute Kidney Injury in critically ill patients.
- Developed algorithms to discriminate persistent kidney injury and transient kidney injury in ICU patients.
- Conducted statistical analyses and designed data visualization APIs to support product promotion.

EDUCATION

Politecnico di Milano - *Master of Science in Mathematical Engineering & Statistical Learning*
March 2019 - Oct. 2021

- Thesis: Functional Time Series Forecasting.
- Advisor: Simone Vantini.
- Teaching assistant for "Algorithms and Parallel Computing", covering C++, OOP, data structures, parallel programming.

Politecnico di Milano - *Bachelor of Science in Mathematical Engineering*
March 2019 - Oct. 2021

- Thesis: Deep Learning Optimization Algorithms and Saddle Points.
- Advisor: Danilo Ardagna.

SELECTED PUBLICATIONS

Ajroldi, N., Orvieto, A., & Geiping, J. When, Where and Why to Average Weights? (2025). To appear at **ICML 2025**. ArXiv preprint.

Islamov, R., **Ajroldi, N.**, Orvieto, A., & Lucchi, A. (2024). Loss landscape characterization of neural networks without over-parametrization. To appear in Advances in Neural Information Processing Systems 2024 (**NeurIPS 2024**). ArXiv preprint.

Ajroldi, N., Diquigiovanni, J., Fontana, M., & Vantini, S. (2023). Conformal prediction bands for two-dimensional functional time series. Computational Statistics & Data Analysis (CSDA), 187, 107821. Article URL.

Alfieri, F., Ancona, A., Tripepi, G., Rubeis, A., **Ajroldi, N.**, Finazzi, S., Cauda, V., & Fagugli, R. M. (2023). Continuous and early prediction of future moderate and severe acute kidney injury in critically ill patients. PLOS ONE, 18, 1–22. Article URL.