Angelo Nardone

Curriculum Vitae

Personal Information

Date of birth	May 17, 1997
Gender	Male
Address	Via Antonio Rosmini, 5, 56123, Pisa (PI), Italy
Phone	(+39) 3453584086
Personal email	angelo.nardone17@gmail.com
GitHub	github.com/Angelido
Linkedin	linkedin.com/in/nardone-angelo

Education

2025–Present **Ph.D. in Computer Science**, *University of Pisa*,
Department of Computer Science,
Supervisor: Paolo Ferragina

2022–2025 **Master's Degree in Computer Science**, *University of Pisa*,
Department of Computer Science,
Curriculum: Artificial Intelligence
Thesis: *Lossless Compression of Source Code using Large Language Models*
Explored novel lossless compression techniques applied to source code in the context of the Software Heritage project, introducing new LLM-based compression pipelines and evaluating 30 different models.
Supervisor: Paolo Ferragina
Final grade: 110/110 cum Laude.

2016–2022 **Bachelor's Degree in Mathematics**, *University of Pisa*,
Department of Mathematics,
Thesis: *Approximate Near Neighbor Searching: Methods for Searching Similar Time Series using the Fréchet Distance*
Supervisor: Roberto Grossi
Final grade: 101/110

Publications and Conferences

2025 **Poster Presentation**, *International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics (CIBB) 2025, 20th Edition*,
Milan, September 10–12, 2025,
Presentation of the short paper "MIEO: encoding clinical data to enhance cardiovascular event prediction", focused on the use of neural models for encoding clinical data to improve cardiovascular event prediction.
Available at: MIEO: encoding clinical data.

Talks and Workshops

2026 **Talk**, *Sequences in London 2026*,
City, University of London (Bayes Business School),
London, February 5–6, 2026
Talk title: *Lossless Compression of Source Code using Large Language Models*.

2026 **Talk**, *A^3Lab Inaugural Workshop: Algorithms, AI & Society*,
Scuola Superiore Sant'Anna,
Pisa, January 21, 2026
Talk title: *Lossless Compression with LLMs*.

Experience

2025 **Research Fellowship**, *University of Pisa*,
Pisa,
Awarded the research fellowship titled *"Compressione dati tramite Large Language Models (LLM)"*, focused on the study and development of compression techniques based on large language models, with particular application to **source code compression**.
The fellowship lasts 6 months, starting on July 1, 2025.

Projects

- **Data Compression through LLMs (Research Fellowship Project)**: study of lossless data compression on source code using Large Language Models, with the goal of improving compression ratio while keeping throughput as close as possible to traditional methods.
  GitHub folder: Lossless Compression using LLMs

- **Clinical Data Encoding with Autoencoders (Poster Presentation Project)**: design of a custom autoencoder for generating clinical embeddings, with automatic handling of missing data and compression of patient health states. This project resulted in the short paper *MIEO: encoding clinical data to enhance cardiovascular event prediction*, presented at CIBB 2025.
  GitHub folder: Clinical Data Encoding

- **Parallel and Distributed Implementations**: implementation of well-known algorithms such as MergeSort, Softmax, and compression with Miniz, designed in parallel or distributed forms.
  GitHub folder: Parallel and Distributed Implementations

- **Data Mining on U.S. Traffic Accidents**: multivariate analysis on real-world datasets; feature engineering, clustering (K-means, hierarchical, DBSCAN), fatal event classification, and time-based analysis by city.
  GitHub folder: Data Mining

- **Sentiment Analysis on Amazon Reviews**: automatic classification using LLMs and topic modeling to identify recurring negative themes, with applications for user experience and brand reputation.
  GitHub folder: Sentiment Analysis

- **MiniHack Evolution**: training evolutionary agents in NetHack environments (via OpenAI Gym), developing strategies through genetic algorithms in a 15×15 gridworld.
  GitHub folder: MiniHack Evolution

Technical Skills

Programming Languages
- C/C++: particularly for parallel programming.
- Python.
- Java.
- Matlab.
- Javascript.

Other Languages Basic knowledge of OCaml, R, and HTML/CSS for front-end web application development.

Tools Visual Studio, LaTeX, GitHub.

Other
- Experience with `pandas` and `torch` libraries.
- Experience working with Large Language Models via HuggingFace.
- Experience with compression algorithms via libraries and custom implementations.
- Experience with Linux and macOS systems.

Certifications

2025-Ongoing **B2.2 English Language Course**, *Linguistic Center, University of Pisa*

2025 **B2.1 English Language Course**, *Linguistic Center, University of Pisa*

2015 **Category B Driver's License**

Tuesday 27th January, 2026

Signature