Nuno Gonçalves

nuno.m.goncalves@tecnico.ulisboa.pt +351 963784966 nunomtg.github.io nunomtg

Education

Instituto Superior Técnico *Sept 2022 – July 2025*
Master of Science – Data Science and Engineering *(Est.)*
- GPA: 18.2/20.0

Technical University of Munich *Oct 2023 – Feb 2024*
Erasmus Program – Mathematics Department
- GPA: 18/20

Instituto Superior Técnico *Sept 2019 – June 2022*
Bachelor of Science – Physics

Experience

Deep Learning Research Intern *Lisbon, PT*
SARDINE Lab *March 2024 –*
- Optimized the Entmax activation function for GPUs using Triton, achieving a 15x speed improvement.
- Designed and implemented a custom algorithm inspired by FlashAttention2, enabling Entmax usage in attention layers for long-context Transformers.

ML Engineer Intern *Lisbon, PT*
Talka.ai *Oct 2023 – Jan 2024*
- Annotated data to incorporate new gestures into a vision-language model, enhancing its ability to classify gestures, behaviors, and speech for insights in sales meetings.
- Designed and implemented a custom data augmentation pipeline, increasing recall by up to 26% for under-represented classes.

Software Engineer Intern *Lisbon, PT*
CERN/LIP *Jan 2022 – July 2022*
- Optimized the GPU-based version of a critical algorithm for identifying particle showers within the Large Hadron Collider (LHC) at CERN.
- Improved the number of events processed per second by approximately 2x compared to the previous implementation.
- Developed a custom GPU-agnostic heuristic to optimize kernel launch parameter selection for performance.

Research Intern *Lisbon, PT*
INESC (Systems and Computing Engineering Institute) *Dec 2021 – March 2022*
- Developed a Neural Network for signal classification capable of achieving 90% accuracy in detecting magnetic nanoparticles within a magnetic flow cytometry lab-on-a-chip device.

Publications

Nuno Gonçalves, Marcos V. Treviso, André F. T. Martins. *AdaSplash: Adaptive Sparse Flash Attention.* To appear in *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 2025. **Spotlight**.

Awards

IST Academic Excellence Award 2024
Awarded for achieving the highest academic distinction across all programs at Instituto Superior Técnico.

***Celfocus* Academic Merit in Machine Learning** 2023
Recognized as one of the top five students in the Machine Learning course for exceptional academic performance.

Projects

Portuguese Word Embeddings

- Implemented and trained word models (HAL, CBOW, Skip-Gram) in a Portuguese corpus, creating contextualized vector representations of portuguese words.

Recommendation Systems In Pure NumPy

- Built and compared different recommendation systems from scratch, including the famous Funk's SVD and multiple non-negative matrix factorization (NMF) methods, fully implemented in pure NumPy.
- Predicted user ratings on the MovieLens dataset and visualized latent feature clustering.
- Explored connections between matrix factorization techniques and auto-encoders for latent representation.

Technologies

Languages: Python, Triton, C, C++, Rust, SQL

APIs and Technologies: CUDA, PyTorch, Git