




Abderraouf Lahmar

-  GitHub Profile
-  jsm6k4@inf.elte.hu
-  LinkedIn Profile

PROFILE

Passionate and enthusiastic AI Engineer with a focus on NLP. I am currently pursuing a Master’s in Data Science at ELTE University, where I conduct research on large language model (LLM) alignment for low-resource languages and cultures, as well as on effective reasoning in hardware-constrained environments. I firmly believe there is always room for improving and adapting algorithms effectively in any use case. I am dedicated to advancing methods that make AI systems more reliable, interpretable, and with a positive impact. I am eager to contribute to research opportunities or pursue a PhD in NLP and applied AI.

EDUCATION

- Eötvös Loránd University, Hungary.** 2024 – 2026
Master’s degree in data science. Cumulative average: 4/5
- 08 Mai 1945 University, Algeria** 2021 – 2024
ranked 1st, Bachelor’s in Information Systems with an honors degree. Average: 16.14/20 (Excellent)
- Mahmoud Ben Mahmoud, Algeria.** 2018 – 2021
Experimental sciences specialty. Average: 16.89/20 (Very good)

EXPERIENCE

- Data science Laboratory** Jan 2025 – Present
Researcher Eötvös Loránd University, Hungary
– I am currently interning as a researcher in Generative AI, focusing on distillation techniques for large language models (LLMs), including self-knowledge distillation, LLMs synthetic-data generation, and reinforcement learning-based approaches.

RESEARCH PROJECTS

- A Benchmark for LLMs Alignment with low-resource Culture** July – August 2025
NeurIPS Workshop(Under review)
– Developed a multilingual benchmark to evaluate large language models’ alignment with some general low-resource cultural norms.
- Iterative Reinforcement Learning Self-Alignment** Jan 2025 – Present
Paper (In progress)
– Investigating methods for continuous alignment of large language models (LLMs) with synthetic human preferences.
– Optimizing LLM behavior using a reward model trained on contrastive self-generated samples with methods like DPO and PPO, enabling iterative self-improvement and preference alignment.
- Graduation project: Smart Door Lock System for Low-Connectivity Areas** Feb – June 2024
Bachelor Thesis
– Developed a novel offline smart door lock system designed for regions with limited or unreliable internet connectivity.
- Self-supervised Speech Representation Learning Technical Survey** November 2024
[Survey]
– A small report about a speech representations in speech models.
- Self-Attention using feature engineering in Reasoning tasks** August 2025 - Present
In progress
– Working on adapting the self-attention mechanism for low-cost deployment in LLM mathematical reasoning tasks, by probing and analyzing the model’s attention layers.

AWARDS

- Honors Certificate** Certificate of Honor awarded by the University Director. 2021–2024
- Top Rank** Achieved 1st place in Bachelor’s program. 2021–2024
- Faculty Top Rank** Achieved 1st place in the STEM faculty at the bachelor level. 2021–2024
- Stipendium Hungaricum Scholarship** Nominated by the Algerian government. 2024–2026

PERSONAL PROJECTS

- **"Attention is all you need" from scratch implementation (2025)** [GitHub]
From-scratch PyTorch implementation of the "Attention Is All You Need" paper, fully optimized for training and inference.
- **Data Poisoning and Evasion Techniques (2025)** [GitHub]
Experiments on data poisoning attacks targeting binary classification models.
- **Mail spam prediction (2023)** [GitHub]
Creating a spam mail detection system using logistic regression model with the help of TF-IDF vectorizer.
- **NLP papers re-implementation** [GitHub]
Re-implementation of some papers related to LLM's self-alignment paradigm.

TECHNICAL SKILLS AND INTERESTS

Technical Skills: Software engineering, Generative AI, Machine learning, RL, Data Analysis, Web scraping, mobile development, API Design, Git, Docker Github, Web development, Database management with SQL, MySQL Database, Pandas, Numpy, Java.

Soft Skills: Team leading skills, scientific writing skills, effective communication, time management.

Field of Interest: Machine learning, NLP, data mining, data science, Reinforcement Learning, LLMs Alignment.

COURSES AND CERTIFICATES

- **Natural Language Processing with Classification and Vector Spaces** [Coursera]
- **Natural Language Processing with Probabilistic Models** [Coursera]
- **Advanced Learning Algorithms** [Coursera]
- **Supervised Machine Learning: Regression and Classification** [Coursera]
- **Introduction to Statistics** [Coursera]
- **Hugging face courses (Transformers,...etc)**

RELEVANT READINGS

- **Speech and Language Processing** By Dan Jurafsky and James H. Martin.
- **Natural Language Processing with PyTorch Build Intelligent Language Applications Using Deep Learning.**
- **Reinforcement Learning: An Introduction** By Andrew Barto and Richard S. Sutton.
- **An Introduction to Statistical Learning.** By Gareth James.
- **Introduction to Machine Learning** By Ethem Alpaydm.
- **Linear Algebra and Its Applications** by Gilbert Strang.
- **Fundamentals of Statistical Signal Processing** Steven M. Kay.