Abir HARRASSE

abirharrasse.github.io

+212 688 972327 | Abir.HARRASSE@emines.um6p.ma

Education

Master's Degree in Industrial Management   *EMINES School of Industrial Management*
Minor in Data Science: **GPA: 3.98/4.0**   *September 2024 - Present*
Coursework includes: NLP, Probabilities, Machine Learning, Optimization, Statistics, Multivariable Calculus, Martingale strategies, Lebesgue Integration, Measure Theory, Linear Programming, Biomedical Engineering, Robotics, Analytics, Principles of Economics

Bachelor's Degree in Industrial Management   *EMINES School of Industrial Management*
Engineering degree program- **GPA: 3.96/4.0**   *September 2020 - September 2024*
Coursework includes: Real Analysis, General Topology, General Algebra, Linear Algebra, Electromagnetics, Electronics, Robotics

Papers

TinySQL: A Progressive Text-to-SQL Dataset for Mechanistic Interpretability Research
Abir Harrasse*, Philip Quirke, Clement Neo, Dhruv Nathawani, Amir Abdullah
(EMNLP Main Conference) Available at: arXiv:2503.12730

Disentangling and Steering Multilingual Representations: Layer-Wise Analysis and Cross-Lingual Control in Language Models
Abir Harrasse*, Florent Draye, Bernhard Schölkopf, Zhijing Jin
(ICML AIW Workshop) Available at: AIW-Workshop

Activation Space Interventions Can Be Transferred Between Large Language Models
Narmeen Oozeer, Dhruv Nathawani, Nirmalendu Prakash, Michael Lan, Abir Harrasse*, Amirali Abdullah
(ICML 2025) Available at: arXiv:2503.04429

Industry Experience

Withmartian   *September 2024 - March 2025*
Research Intern   *Bay Area, San Francisco*

- Using mechanistic interpretability methods to probe models during reasoning tasks, investigating whether human-interpretable causal algorithms are implemented. The goal is to develop a general framework for conducting mechanistic interpretability experiments. Our work was submitted to venues such as ICML 2025 and EMNLP 2025.

Research Experience

Causal Interpretability - The Empirical Inference Lab   *April 2025 - Present*
Research Intern, supervised by **Prof. Zhijing Jin**   *Max Planck Institute for Intelligent Systems*

- Research internship focused on two main aspects: studying and interpreting multilingual mechanisms in Large Language Models, and developing new tools for training data attribution estimation.

LLMs evaluating LLMs   *April 2024 - August 2024*
Research Intern, supervised by **Prof. Chaithanya Bandi**   *National University of Singapore*

- Designed adversarial multi-agent systems for LLM output evaluation, with roles like advocates and judges. This dynamic framework improved decision-making, achieving an 8% performance boost over prior methods.

RL with generative models *January 2023 - January 2024*
Research assistant, supervised by **Prof. Omar Saadi** *College of Computing-UM6P*

- Focused on model-based reinforcement learning using generative models, testing existing algorithms and optimizing their runtime and sampling efficiency.

Product Assortment Optimization *June 2023 - July 2023*
Research Intern, supervised by **Prof. Agnès Gorge** *Africa Business School - UM6P*

- Researched the theoretical aspects of assortment optimization and its impact on retailer profitability. This involved studying substitution effects and demand uncertainty, leading to the development of a new machine learning heuristic.

Talks

Training Data Attribution *August 2025*
Amazon *Tübingen, Germany*

- Talk on training data attribution in LLMs, covering verbatim and gradient-based methods, their strengths, limitations, and open challenges..

Awards and Achievements

First Class Honours, EMINES, UM6P	Graduated with first-class honours for exceptional academic performance.
Best Master's Thesis, EMINES, UM6P	Recognized for outstanding research on understanding and interpreting language processing in LLMs.
National Moroccan Merit Scholarship FAR (2020)	Awarded to the top 50 national scores in high school final examinations, representing the top 0.1% of all candidates.
UM6P Excellence Scholarship (2020)	Granted for exceptional performance in entrance examinations, placing within the top 1% of applicants.

Key Achievements and Projects

"Chariot Autonome" Robot *October 2023 - May 2024*
AI lead *EMINES-UM6P*

- Collaborated with my team to develop an autonomous robot capable of following the user in real time. Implemented AI features using human tracking algorithms and depth estimation to enhance navigation and interaction capabilities.

BCG Platinion Hackathon *October 2023*
National Winner *BCG Casablanca*

- Our mobility solution, optimizing matchmaking among transporters, suppliers, and demanders, won the top national prize and second place internationally at the hackathon with 150+ participants. It leverages Google's VRP-Solver for route optimization under multiple constraints.

Morocco IoT and AI Challenge *September 2023*
Finalist *Marrakech, Morocco*

- As a finalist in the Morocco IoT and AI Challenge, I presented **Med-Flamingo**, a medical chatbot combining visual language models and LLMs. It interacts in Moroccan Dialect, answers medical questions, and generates reports, with ongoing work on AI safety enhancements. The project is accessible through: medical chatbot.

NaMO - Preparing for IMO's 61st edition *January 2019 - March 2020*
Participant *Rabat, Morocco*
Among the top 25 participants selected for the country's most prestigious mathematical competition, where more than 1000 best high school students gathered to prepare for international contests like the International Mathematical Olympiads (IMO).

Skills

Languages	*Native:* Arabic, French. *Fluent:* English. *Beginner:* Mandarin
Programming Languages/Tools	Python, Pytorch, TransformerLens, SAELens, Nnsight, Transformers, Diffusers, Datasets, Slurm, FICO Xpress, SQL

Volunteering

MathMaroc *September 2021-September 2024*
Vice General Secretary
Collaborating with fellow members to organize events, workshops, and outreach programs to engage students and enthusiasts in the exploration and appreciation of mathematics.

The GenAI Winter School *October 2023- February 2024*
Organizer
Taking an active role in the coordination and organization of The **GenAI Winter School**, an event that welcomed distinguished researchers including **Yann Le Cun**, **Eric Xing**, and other prominent figures in the field.

Rotaract EMINES Club *January 2022 - June 2022*
Co-founder and General Secretary
Co-founded a social impact club focused on aiding local communities through initiatives such as food distribution, visits to orphanages, and educational orientation sessions in remote middle schools.

E-maths Club *January 2021 - June 2021*
Founder and President
Founded the first mathematics club at EMINES-UM6P to promote math culture through ludic activities, conferences, peer-to-peer workshops and study groups. Organized orientation sessions connecting students with alumni to explore academic and research pathways.