Ibrahim Merad (Data & AI Scientist)

+33 6 65 03 87 27 | imerad7@gmail.com | 29 years old | French and Algerian citizenship

I enjoy developing data-powered machine learning/AI solutions to real world problems and have significant experience with Computer Vision and NLP models. I have a strong background in mathematics, statistics and optimization together with key skills in Computer Science, programming and algorithms.

EXPERIENCE

Kaukana Ventures Jul. 2024 – Present
Data & AI Scientist Paris, France

- Development of Deep Learning and AI based solutions in various R&D Projects (see KV projects below)

Université Paris Cité Sep. 2022 – Dec. 2023
Teaching assistant (ATER) Paris, France

- Teaching and research roles at the Mathematics department. Tutorials and labs for algorithmic and numerical methods courses.

Laboratoire de Probabilités Statistiques et Modélisation (LPSM) Apr. 2020 – Aug. 2020
Research Intern Paris, France

- Study of the training convergence of attentive recurrent neural networks and extension to contrastive methods in the unsupervised setting. Continued for a PhD thesis.

Medical Image Processing Lab (MIPLab), Biotech campus Apr. 2019 – Aug. 2019
Research Intern Geneva, Switzerland

- Investigated applications of optimal transport to signal processing on fMRI data and extensions for Slepian functions.

PROJECTS

Firetracking (KV project)

- Early wildfire detection solution based on computer vision.
- Curated datasets and trained multiple vision models for object detection and video classification.
- Developed backend postprocessing of inference results for sequential filtering and aggregation of raised alerts.

Agentic screening (KV project)

- Developed an agentic solution automating M&A screening. LLM-powered, autonomous internet navigation for information collection, curation, and structured restitution.

Rare Language translation (KV project)

- Trained multiple LLMs on Monégasque translation, a very low resource language. Communicated results through scientific publication (see dedicated section below).

Python Libraries

- `linlearn` : Efficient linear learning library implementing statistically robust estimation methods and multiple optimization algorithms.
- `wildwood` : Implements the Wildwood algorithm extending and improving effiency in random forests with fewer trees.

SKILLS

Programming Languages : Python, C/C++, Java, Matlab, SQL, Ocaml, LaTeX.

Tools/Libraries : git, docker, AWS S3, Scitkit Learn, Pytorch (Lightning), Huggingface Transformers, Tensorflow, Numpy/Numba, Pandas, OpenCV ...

Languages : French (bilingual), English (fluent), German (conversational), Arabic (mother tongue). Spanish/Russian/Japanese (intermediate level)

EDUCATION

Université Paris Cité Nov. 2020 – Dec. 2023

PhD in Applied Mathematics (Machine Learning) Paris, France

- "Robust Algorithms and other Contributions to Machine Learning", directed by Pr. Stéphane Gaïffas and Pr. Emmanuel Bacry. Defended on December 12th 2023.

ENS Paris-Saclay Sep. 2019 – Aug. 2020

MSc in Applied Mathematics Paris, France

- **Master 2 (MVA)**: Mathematics, Vision and Machine Learning

ENS Ulm 2017 – 2020

ENS Diploma Paris, France

- Major in Mathematics and minor in Computer Science.

Sorbone Université (previously UPMC) Sep. 2019 – Aug. 2020

Bachelor in Mathematics and Computer science Paris, France

- Double Major in Mathematics and Computer science

Scientific Publications

About contrastive unsupervised representation learning for classification and its convergence. *Preprint arXiv:2012.01064*. Joint work with Yiyang Yu, Stéphane Gaïffas and Emmanuel Bacry.

WildWood: a new Random Forest algorithm. *IEEE Transactions on Information theory*. Implementation in the `wildwood` Python library. Joint work with Stéphane Gaïffas and Yiyang Yu.

Robust supervised learning with coordinate gradient descent. *Springer's Statistics & Computing journal*. Implementation in the `linlearn` Python library. Joint work with Stéphane Gaïffas.

Robust methods for high-dimensional linear learning. *Journal of machine learning research (JMLR)*. Implementation in the `linlearn` Python library. Joint work with Stéphane Gaïffas.

Robust Stochastic Optimization via Gradient Quantile Clipping. *Transactions of Machine Learning Research (TMLR)*. Joint work with Stéphane Gaïffas.

Convergence and concentration properties of constant step-size SGD through Markov chains. *Electronic Journal of Statistics (EJS)*. Joint work with Stéphane Gaïffas.

Language verY Rare for All. *Proceedings of the First Workshop on Language Models for Low-Resource Languages (LoResLM)*. Joint work with Amos Wolf, Ziad Mazzawi and Yannick Léo.