



Marius Roger

Massy 91300, France
marius.roger@outlook.fr
linkedin.com/in/marius-roger

Machine Learning Engineer

Looking for a first position in Data Science, focused on ML Ops or R&D.

Experience

April 2024

September 2024

ML Engineer R&D intern *GitGuardian*

- Adding context to secrets (security incidents) with NLP (HuggingFace, llama.cpp, SkyPilot, AWS EC2)
- Building a relevant generic secrets dataset (Prodigy, Snowflake, DVC, OpenAI)
- Contributing to the secrets detection engine (Python, git, PDM, Linear)

October 2022

June 2023

Research assistant *LMU Munich*

- Large-scale analysis and comparison of French & German LLMs (HuggingFace)
- Review and writing of research articles on NLP techniques and biases : [article](#)
- Review of teaching material and grading of the Deep Learning for NLP class

April 2022

July 2022

ML Engineer R&D intern *HrFlow.ai*

- Bibliography and experiments on recommender systems, graph neural networks and LLMs
- PoC on the matching of candidates and job offers (PyTorch, HuggingFace)

March 2021

August 2021

NLP research intern *Dauphine PSL*

- Design of recurrent models (LSTM & GRU) with Neural ODE (PyTorch)
- Training and evaluation on the PennTreeBank dataset with the perplexity metric

Education

2019 – 2024

ENS Paris-Saclay Diploma

Top research-oriented school, Computer Science, Data Science and Machine Learning program

2023 – 2024

Master 2 IASD *Dauphine PSL, Mines Paris, ENS PSL*

Machine Learning, Data extraction and storage, NoSQL and distributed computing

- Data science project : Recommender systems and GANs with NumPy, PyTorch and git
- Databases project : Extraction of researchers relationships from DBLP with XML SAX and Neo4j
- NoSQL project : Implementation of iSAX (time series indexing) in Python

2021 – 2022

Master 2 MVA *ENS Paris-Saclay, Institut Polytechnique de Paris*

Machine Learning, Natural Language Processing, Time Series analysis, Audio processing

- Research paper review : NLP, signal processing, time series analysis, graph NNs, biostatistics

2020 – 2021

Master 1 MPRI *ENS Paris-Saclay*

Computer Vision, Deep Learning, Software Engineering, Optimization, Markov Chains

- Software engineering project : Modular audio synthesis/processing program for Linux in C++

2017 – 2019

Classes Préparatoires aux Grandes Écoles *Lycée Camille Guérin*

Maths, physics, and computer science training for competitive entrance to french Grandes Écoles

Skills

Programming

Python (PyTorch, Pandas, Transformers), Scala (Spark), SQL, OCaml, Bash

Tools

git, SSH, LaTeX, Notion, DVC, Linear, SkyPilot, Snowflake

Relational

Collaborative, Organized, Flexible, Autonomous, Analytic, Curious

Languages

English (C2), Spanish (B2), French (native)

Interests

Music production, electronics for audio processing, handball