

Mathieu Nalpon

mathieunlp@gmail.com ♦ mathieunlp.github.io ♦ +33 677786248

EDUCATION

- **CentraleSupélec - Paris Saclay University** Paris, France
Master's degree in Mathematics and Data Science Sept. 2019 – Oct. 2022
- **Cachan School** Cachan, France
Bachelor of Engineering Sept. 2016 – July. 2019

EXPERIENCE

- **Weborama** Paris, France
Machine Learning Engineer Oct. 2023 - Current
 - **WeboMind - URL recommendation:** Contextual targeting tool generating lists of relevant URLs from a prompt for campaign creation.
 - * Embedded crawled web pages using ME5 model.
 - * Indexed vectors using IVF and product quantization optimize search (similarity) and storage in Milvus.
 - * Fine-tuned a reranker to improve the scoring of urls.
 - * Benchmarked and switched our search to hybrid search with HNSW for our use case.
 - * CI/CD pipeline and deployed the product with APIs.
 - **Data engineering for AutoML:** Build data pipeline to predict the conversion ratio of users given the next add campaign.
 - * Create a package to collect data and optimize the SQL queries.
 - * Build a linear regression and lightGMB models fine-tuned with Optuna and plug the data pipeline with it. Orchestrate with Airflow.
 - * Deployed in production with monitoring (ML metrics: ROC, Acc; API metrics: latency, memory)
- **ArianeGroup Defense & Space** Paris, France
Data Scientist Oct. 2020 - Oct. 2022
 - **Autonomous neutralization of a launcher in flight:** Research on state-of-the-art solution for autonomous neutralization of a launcher in flight using Mahalanobis distance.
 - **Optimize the computation of time before neutralization:** Apply statistical methods and geometry to improve the resulting time before neutralizing a launcher by taking into account the spread of debris from the explosion.
 - **Prediction of the time left for an off-trajectory rocket before neutralization:** Created from scratch an algorithm that computes, given a trajectory and a specific deviation in 3D space, the time left for a rocket before it's neutralization.

PROJECTS

- **Fine-tuning SAM (Meta) for ring segmentation:** Used a Lora adapter with a hand-made dataset to solve the problem of segmenting single and pair of rings which was an issue with SAM 1.
- **Je-sais-voter:** Built a radar-graph to help people understand the political ambitions of the different french parties given their program.
- **SAMA-drums:** Isolated drum sound using Segment Anything Audio and automatically generate a music sheet to help training the earing and notation of music learners.
- **NanoMedChat:** Created a local medical chat fine-tuned with medical data. The training strategy is the same as the one used in the nanochat repository.
- **ChessVLM:** Developed a VLM to read the notes taken from the position paper in a tournament and put them in chess.com for breaking down the game. This would help the players wanting to review their game to quickly see the positions.

TECHNICAL SKILLS

- **Languages:** Python, Bash, JavaScript, C++
- **Libraries:** PyTorch, Optuna, Ray, FAISS
- **Frameworks & Tools:** Git, Docker, Kubernetes, Snowflake, ElasticSearch, Sentry, Prometheus, Grafana, MLFlow, Google Cloud, Milvus, Airflow