Eithan Nakache

eithannakache@gmail.com | LinkedIn | Portfolio

AI Research Intern | Machine Learning, LLMs, Agentic AI

WORK EXPERIENCE

Siemens Healthineers, Princeton, NJ, USA **March 2025 – Sep 2025**

AI Research Intern

- Designed and deployed a scalable **Vision-Language Model** pipeline to extract, structure, and index knowledge from complex, unstructured clinical guidelines (PDFs).
- Replaced traditional **OCR** and layout detection with a multimodal end-to-end solution, achieving **15x faster processing** (**0.39s** vs. **7.22s** per page) and improved robustness on noisy, heterogeneous documents.
- Boosted retrieval performance to **0.81 NDCG@5** (**+23% over baseline**), enabling reliable, agent-driven **clinical decision support**.
- Developed intelligent AI agents integrating **LLMs** and external tools to dynamically orchestrate specialized models for complex, multi-step medical reasoning tasks.
- Led **prompt engineering** and evaluation to optimize agent performance, resolve reasoning bottlenecks, and enhance robustness in **real-world clinical scenarios**.
  Tech Stack: Python, Hugging Face, Vision-Language Models, LLMs, NLP, Agentic AI

ACHIEVEMENTS

MISTRAL AI Hackathon | Top 3 (Mistral AI Model Dilution)

- Designed and built a modular platform to distill **task-specific models** from large **pre-trained LLMs**, using lightweight fine-tuning techniques.
- Combined adversarial and curriculum learning with synthetic data generation, improving downstream task performance by **up to 18%**, while reducing model size by **40%**.
  Tech Stack: Python, Hugging Face, PyTorch

French Army Hackathon 6milarity | Winner (Computer Vision Challenge)

- Classified car models using ResNet-18 architecture.
- Leveraged model explainability techniques to improve model performance.
- Achieved a **93% accuracy** rate in car model classification.
  Tech Stack: Python, Pandas, PyTorch, OpenCV, CNN

Hugging Face & Zama Privacy-Preserving AI Hackathon

- Developed an encrypted face recognition pipeline using **Fully Homomorphic Encryption** (FHE), ensuring secure, privacy-preserving on-device authentication.
- Processed images locally and encrypted all sensitive data to prevent reverse engineering.
  Tech Stack: Python, FaceNet, FHE, Concrete ML, Hugging Face

PROJECTS

NanoDiffVision – Vision Transformers with Differential Attention

- Re-implemented Vision Transformer and integrated **Differential Attention** to reduce attention noise.
- Achieved up to **96.03% accuracy** on MNIST and **86.73%** on FashionMNIST using compact model.
  Tech Stack: Python, PyTorch, PyTorch Lightning

EDUCATION

Master MVA (Mathematics, Vision, Learning) **Sep 2025 – Aug 2026**
ENS Paris-Saclay, France

Master of Engineering – Artificial Intelligence and Big Data **Sep 2020 – Aug 2025**
École pour l'informatique et les techniques avancées, Paris (EPITA) (GPA: 4.0/4.0)

Exchange Semester – Computer Science Engineering **Jan 2022 – May 2022**
California State University, Los Angeles (CSULA)

SKILLS

Programming & Tools: Python | C | C++ | Java | Bash | Slurm | SQL | NoSQL | Docker | Git | Azure | HPC
AI & Data Science: TensorFlow | PyTorch | Scikit-learn | Hugging Face | vLLM | Ollama | Langchain | Docling
Languages: French (Native) | English (Fluent) | Spanish (Notions)