MAYSSA BEN ROMDHANE

Ingénieure IA — RAG, Agents LLM & Computer Vision

Monastir, Tunisie | mayssabenromdhane1999@gmail.com | +216 93 653 187
linkedin.com/in/mayssa-ben-romdhane | github.com/Mayssabr

Ingénieure IA, 2 ans d'expérience sur des systèmes LLM en production. Chez HealthCare Innovation, conçoit l'architecture IA et les pipelines RAG et agentiques de 10 plateformes HealthTech, avec des LLM open-weight déployés en souveraineté sur GPU et une évaluation continue du grounding et des hallucinations, après une première expérience R&D chez Continental en pipelines de segmentation sémantique et en analyse de fusion caméra-LiDAR 3D en zero-shot, marquée par une auto-annotation réduisant de plus des deux tiers le coût d'étiquetage.

EXPÉRIENCES PROFESSIONNELLES

Ingénieure IA — RAG & Systèmes agentiques | HealthCare Innovation, Monastir *Janv. 2026 - Présent*

- Conçu et industrialisé la stack IA de **10 plateformes HealthTech/HSE** en production : architecture RAG, orchestration d'agents et harness d'évaluation.
- Déployé un **RAG souverain on-premise** — LLM open-weight sur GPU dédié, recherche hybride sémantique + mot-clé, citation systématique des sources — réduisant la latence de **40 %**.
- Mis en place l'évaluation continue (**RAGAS**, **LLM-as-judge**, Langfuse) et la recherche Qdrant/ChromaDB sur **10 000+ documents**, avec **+25 % de pertinence**.
- Orchestré des systèmes multi-agents **LangGraph** avec validation humaine (HITL), reprise sur incident et traçage de bout en bout.

Ingénieure IA générative | Freelance, Monastir *Mars 2025 - Déc. 2025*

- Développé une application Stable Diffusion/DALL-E avec fine-tuning **LoRA** : une image en **moins de 5 s**.
- Déployé un backend FastAPI et une interface React via Docker, avec optimisation de l'inférence GPU.

PFE ingénieur - Computer Vision R&D | Continental, Lindau, Allemagne *Févr. 2024 - Nov. 2024*

- Conçu un pipeline de segmentation sémantique de route (ADAS) par **foundation models en zero-shot**, atteignant **99 % de précision**.
- Automatisé l'annotation image/nuage de points, réduisant le coût d'étiquetage de **70 %**.
- Évalué et comparé les architectures de segmentation sous **PyTorch** sur données automobiles : **+8 % de mIoU** pour le modèle retenu.

Stage ingénieur | Wisi, Monastir *Juin 2023 - Août 2023*

- Développé un streaming adaptatif multi-bitrate et un serveur local de test avec Docker, Node.js et Shaka Packager.

PFE de licence | Carthage Solutions / CHU Monastir *Févr. 2021 - Juin 2021*

- Développé **de bout en bout** un modèle de **posologie prédictive par patient** pour le service de pharmacologie du CHU de Monastir (Python, R, scikit-learn), dont les prédictions ont été **validées par les médecins** du service.

PROJETS

- **Détection précoce du mélanome (2023, académique)** : classification de lésions cutanées — VGG16/ResNet-50, SVM, transfer learning et augmentation Albumentations.
- **Classification ECG (2022, académique)** : prétraitement du signal, conversion en spectrogrammes et classification CNN avec TensorFlow/Keras.
- **Classification de vibrations (2022, académique)** : comparaison de SVM, régression logistique et réseaux de neurones sur des données de capteurs vibratoires.

FORMATION

Diplôme d'ingénieur en informatique appliquée - Systèmes intelligents
École nationale d'ingénieurs de Sousse (ENISo) | 2021 - 2024

Licence fondamentale en informatique
Faculté des sciences de Monastir | 2018 - 2021

COMPÉTENCES

- **Computer Vision & Deep Learning :** PyTorch, TensorFlow, OpenCV, ViT/Swin, CNN (ResNet, YOLO, U-Net), SAM/SEEM/CLIP, Image Segmentation, Object Detection, LiDAR 3D, Camera-LiDAR Fusion, ONNX, TensorRT, ONNX Runtime.
- **IA générative & NLP :** LLMs, RAG, LoRA/QLoRA, LangChain, LangGraph, MCP, AI Agents, Prompt Engineering, embeddings, recherche hybride BM25, Qdrant, ChromaDB, Hugging Face, déploiement on-premise de LLM open-weight.
- **Évaluation & observabilité :** RAGAS, LLM-as-judge, Langfuse, jeux de référence versionnés, grounding et anti-hallucination.
- **MLOps & Cloud :** Docker, Kubernetes, CI/CD, AWS, FastAPI, REST APIs, Redis, BullMQ, Git, Linux.
- **Programmation & Data :** Python, C++, TypeScript, SQL, NumPy, pandas, Machine Learning (scikit-learn).

LANGUES

Arabe langue maternelle | Français C1 | Anglais C1