

# Otmane El Bourki

📍 France • 📞 +33775735751 • ✉ otmane.elbourki@gmail.com • 🌐 oel • 🌐 oel • 🌐 oelbourki.site

## EXPÉRIENCE

**Stagiaire Ingénieur Intelligence Artificielle** — Stage de fin d'études de Master Jan 2024 – Juin 2024  
Tetricks – Solutions d'hospitalité alimentées par l'IA 🌐 Sidi Slimane, Maroc

- Conception et mise en œuvre d'une application de chat basée sur RAG pour traiter les demandes des invités, en utilisant des LLM open-source. Amélioration des performances de RAG avec une recherche hybride sur la base de données vectorielle et optimisation du déploiement en utilisant les services GCP, y compris Cloud Run.
- Développement d'un pipeline utilisant VLM, permettant la génération automatisée de descriptions pour les chambres d'hôtel avec **99,9% de disponibilité**. Le modèle est servi en utilisant Ollama sur une machine virtuelle, accessible via une API Flask sur Cloud Run.

**Stagiaire Ingénieur Intelligence Artificielle** — Stage d'été Juin 2023 – Oct 2023  
Tetricks – Solutions d'hospitalité alimentées par l'IA 🌐 Sidi Slimane, Maroc

- Architecture et déploiement de pipelines ML en production pour la classification des demandes sur GCP, permettant une **réduction de 47% des coûts** opérationnels grâce à l'optimisation des workflows automatisés et à une utilisation efficace des ressources.
- Mise en œuvre des techniques d'optimisation des LLM, y compris la distillation et la quantification, offrant des temps d'inférence **3X plus rapides** tout en maintenant la performance du modèle.
- Conception d'un système intelligent de traitement des e-mails utilisant LangChain et des LLM personnalisés, entraînant une **amélioration de 60%** de la précision du traitement des demandes par rapport à la méthode de base basée sur regex.

**Stagiaire Ingénieur Machine Learning** — Stage de fin d'études de Licence Fév 2022 – Juin 2022  
Département CS d'Ibn Tofaïl – Institution de Recherche 🌐 Kenitra, Maroc

- Conception et développement d'un système de reconnaissance des émotions utilisant l'apprentissage profond, atteignant **85% de précision** sur le dataset FER2013, et mise en œuvre des modèles de vision SOTA atteignant **90% de précision moyenne** grâce à l'optimisation des hyperparamètres.
- Amélioration de la robustesse du modèle de **20%** grâce à des techniques d'augmentation des données et des stratégies de régularisation (par exemple, dropout). Développement et déploiement d'une application de bureau pour fournir le modèle dans une interface utilisateur conviviale.

## ÉDUCATION

**Expert en Architecture Informatique** — Option Data (M2, RNCP39774) Oct 2024 – Août 2025  
École 42 Paris 🌐 Paris, France

- Principaux domaines : IA, Data Science, Microservices, Conception de Systèmes, Infrastructure as Code, et pratiques DevOps

**Master en Intelligence Artificielle** — (Bac + 5) Oct 2022 – Août 2024  
Université Ibn Tofaïl 🌐 Kenitra, Maroc

- Cours : Algorithmes Avancés de ML, Vision par Ordinateur, Traitement du Langage Naturel, Traitement d'Image, Programmation en Parallèle

**Licence en Mathématiques et Informatique** — (Bac + 3) Sep 2018 – Juil 2022  
Université Ibn Tofaïl 🌐 Kenitra, Maroc

- Études principales en Structures de Données, Algorithmes, Algèbre Linéaire et Génie Logiciel

## PROJETS PHARE

🌐 **Chatbot Médical RAG** | RAG, Meditron 7B LLM, Qdrant, Chainlit, Leadership, Communication |

- Conception et développement d'une application de question-réponse (QA) basée sur RAG, exploitant le Meditron 7B LLM, la base de données vectorielle Qdrant et PubMedBERT pour fournir des informations médicales précises et fiables en temps réel. Intégration d'une interaction fluide via l'interface Chainlit et mise en œuvre d'une recherche vectorielle scalable pour optimiser les performances.
- Supervision et gestion d'une équipe de trois étudiants pour mener à bien ce projet dans le cadre de mon projet transversal de master, favorisant une collaboration efficace et le développement des compétences techniques de l'équipe.
- Optimisation du déploiement et de l'exécution du système en utilisant Docker pour la containerisation, avec prise en charge GPU pour une utilisation performante du modèle, tout en mettant en place des configurations robustes pour des tests et une surveillance continus.
- Développement de compétences en leadership, communication, et résolution de problèmes en coordonnant les efforts de l'équipe, en gérant les délais et en assurant la qualité des livrables.

🌐 **Système de Classification des Maladies Rénales de A à Z** | AWS, MLflow, DVC, MLOps |

- Conception et déploiement d'un système simple de diagnostic médical atteignant **94% de précision**, intégration des pratiques MLOps pour les tests automatisés, la surveillance et les tests A/B, et mise en place de pipelines CI/CD robustes en utilisant GitHub Actions pour un réentraînement et déploiement continus des modèles sans temps d'arrêt.

- 🔗 **Plateforme d'Automatisation de Newsletters Alimentée par l'IA** | *CrewAI, Exa, LangChain, Streamlit* |
- Conception et développement d'un système automatisé de génération de newsletters en utilisant CrewAI pour gérer les agents de recherche en IA pour récupérer et résumer les dernières nouvelles, intégration d'Exa pour l'agrégation avancée de données, LangChain et ChatGPT pour générer des résumés concis, et une interface Streamlit pour une interaction et gestion sans faille.

COMPÉTENCES

---

**Intelligence Artificielle & Apprentissage Automatique :** Machine Learning, Deep Learning, NLP, Computer Vision, Reinforcement Learning, Large Language Models, Prompt Engineering, RAG, Fine-tuning, Embeddings, PyTorch, TensorFlow, Hugging Face, LangChain, CrewAI, OpenCV  
**Infrastructure & DevOps :** AWS (SageMaker), GCP (Vertex AI, Cloud Run), Docker, Kubernetes, Model Deployment, Model Monitoring (MLflow), Model Serving, A/B Testing  
**Données & Développement :** Python, C++, Git/GitHub, REST APIs, FastAPI, SQL, Vector Databases (Qdrant, Chroma)  
**Compétences Interpersonnelles :** Pensée analytique, Résolution collaborative de problèmes, Adaptabilité, Apprentissage continu  
**Langues :** Anglais (Professionnel complet — C2), Français (Professionnel — B2)

CERTIFICATIONS

---

**Generative AI Winter School** — Ecole Centrale Casablanca — 03/2024, **Practical Data Science** — AWS — 04/2022,  
**ML in Python with scikit-learn** — Inria — 07/2021, **Mathematics for Machine Learning** — ICL — 04/2021,  
**Professional TensorFlow Developer Certificate** — Deep Learning Specialization — DeepLearning.AI — 10/2020