Tunis, Tunisia
(+216) 26168924
mastourmontassar@gmail.com

Montassar Mastour

Research IA Engineer

https://montassar-mastour.com
linkedin.com/montassar-mastour
github.com/montassar-mastour

Research AI Engineer with a strong vision and passion for advancing the frontiers of artificial intelligence. I am dedicated to exploring novel methods and architectures that push the boundaries of machine learning and contribute to real-world innovation. As an leader in various organizations, I combine technical depth with a commitment to impactful research and meaningful causes. I believe in continuous learning and personal growth, consistently investing in my development to stay at the cutting edge of this fast-evolving field. With proven experience in deep learning, computer vision, and applied research, I have designed, implemented, and evaluated AI models that not only solve complex problems but also drive scientific and technological progress.

EDUCATION

Doctorate in Artificial Intelligence – 1st Year , *Tunis Higher Institute of Management (SMART Lab - ISG)* 11/2025 - Present
Research Master Degree in Data science , *Higher Institute of Multimedia arts of Manouba* 09/2022 - 11/2024
Bachelor Degree in Computer Science , *Higher Institute of Computer Science and Management of Kairouan* 09/2019 - 06/2022
High School degree in Computer science , *Sbikha Secondary School 1997* 09/2018 - 06/2019

SKILLS

AI Development	Generative AI, LLM, RAG Systems, NLP, Vector Databases, Fine-Tuning, Prompt Engineering
Tools and Languages	Python, SQL, PyTorch, LangChain, PostgreSQL, FastAPI, Qdrant, pgvector, Git, GCP, Hugging Face, Docker
Communication	English (Fluent), french (Fluent), Arabic (Native)

TECHNICAL EXPERIENCE

Data Scientist / Freelance Project **06/2025 — 11/2025**
Data Science Research Project — Impact of Climate Change on Aquaculture Production *Online / Tunisia*

- Built a **large-scale spatio-temporal environmental database** integrating **80+ oceanographic, climatic, and biological variables** from Copernicus, ERA5, FAO, and institutional datasets (2003–2022).
- Collected, validated, and harmonized **real-world environmental and production data** from monitoring agencies, scientific reports, and field observations to ensure **accurate ground-truth inputs**.
- Employed **remote-sensing imagery** (Sentinel, Landsat) to **detect, map, and validate aquaculture site locations**, improving spatial precision and exposure assessment.
- Performed extensive **data cleaning and exploratory analysis**, including **spatial mapping, temporal trends, seasonal decomposition, anomaly detection**, and multi-year comparative studies.
- Processed and analyzed Environmental Variables such as **salinity, temperature, pH, nutrients, oxygen, radiation, winds, chlorophyll**, and other biological indicators.
- Conducted **advanced feature selection** using **correlation matrices, multicollinearity tests, variable clustering**, and importance ranking to identify key environmental drivers.
- Developed **AI-based forecasting and classification models** (**Random Forest, Ridge Regression**) with **SHAP-based explainability** to interpret environmental influence on ecosystem productivity.
- Applied **geospatial and GIS analysis**, including **vulnerability assessment** and ecosystem interactions with **blue-carbon**.
- Produced **high-impact visual analytics** (**time-series dashboards, spatial maps, environmental indicators, heatmaps**) for scientific interpretation and decision-making support.

Data Scientist Researcher / Internship | **GitHub Repository** **02/2024 — 11/2024**
Laboratoire RIADI *Manouba Campus*

- Designed and implemented a **remote-sensing image classification pipeline** using a hybrid **Bag-of-Visual-Words (BoVW) + CNN + Graph Convolutional Network (GCN)** architecture for robust feature representation.
- Used a **pretrained ResNet50 backbone** combined with a 2-layer attention-based GCN on superpixel-derived graphs (via SLIC + K-NN graph construction).
- Built end-to-end **data preprocessing, feature extraction, and model training pipelines**, including dataset download/preparation, superpixel segmentation, graph construction, CNN-GCN feature extraction, and BoVW codebook generation.
- Trained models on the NWPU-RESISC45 remote-sensing dataset; achieved expected test accuracy in the **85–92% range**.
- Developed **feature extraction and classification pipelines** using CNN-GCN for image representation, followed by classical classifiers (Random Forest / SVM) on BoVW features.
- Ensured **reproducibility and code quality**: environment setup with Conda, automated pipeline scripts, unit tests (pytest), configuration-driven design, and public open-source release under MIT license.
- Enabled **scalable processing and efficient inference**: GPU-based training (for CNN-GCN), CPU-based preprocessing, sparse-tensor optimizations, and batch-processing to handle large image volumes.
- Provided **evaluation and analysis tools**, including confusion matrices, per-class metrics, and resource usage estimates (model size, VRAM, throughput).

Professional Unity VR Development Program Associate/ Apprenticeship | GitHub Repository **10/2024 — 01/2025**
XR Bootcamp *Online / San Francisco, USA*

- Completed a highly selective 4-month professional VR development program training job-ready XR software engineers.
- Developed and prototyped VR, AR, MR, and XR applications using Unity and C#.
- Programmed advanced functionality in Unity, including interaction systems and gameplay mechanics.
- Applied rendering optimization techniques to improve VR application performance.
- Collaborated in XR development teams, following industry workflows and best practices.

Mobile Developer / Internship | GitHub Repository **02/2022 — 06/2022**
CompiTechnology *Online*

- Developed **Hassassia**, a smart food recommendation mobile application using **Flutter (Dart)**, helping users with allergies and dietary preferences make safe and healthy food choices.
- Integrated a **barcode scanner** and **text-based search** to retrieve product nutritional and allergen information from the **Open Food Facts API**, combined with local NoSQL storage for offline functionality.
- Designed the app with a **dual database system** (local NoSQL + remote MySQL) to ensure offline access, data synchronization, and efficient storage management.
- Developed backend services in **PHP + MySQL** for user management, preference storage, and data retrieval.
- Followed the **Waterfall development methodology**, including requirements analysis, UML design (class & sequence diagrams), implementation, testing, and maintenance.
- Used **Git (Github)** for version control, enabling collaborative development and reliable code management.
- Conducted module-level testing and validation to ensure correct functionality and robust user experience.

PUBLICATION

Montassar Mastour and M. Farah, **"Bag-of-Visual-Words Feature Extraction using Graph Convolutional Networks,"** in *2024 IEEE 7th International Conference on Advanced Technologies, Signal and Image Processing (ATSIP)*, vol. 1, pp. 536–541, 2024. [Link]

PROJECTS

RAG-Based Knowledge Retrieval System | GitHub Repository **08/2025 — 10/2025**
Personal / Research Project

- Developed a full-featured **cloud-deployed AI question-answering system** using a modern **Retriever-Augmented Generation (RAG)** architecture, combining semantic vector search with LLM-based generation.
- Implemented a complete **document ingestion and preprocessing pipeline**: support for PDF/website content import, text chunking, embedding generation, and vector-database indexing (PGVector/Qdrant).
- Built backend and API layer with **FastAPI**, enabling file uploads, query handling, and orchestration of retrieval + workflow.
- Integrated **semantic retrieval + LLM response generation** (via LangChain + user-configurable LLM backend) to turn user queries + retrieved context into precise, grounded answers.
- Ensured **persistent metadata and document storage** using a relational or NoSQL database for document management and retrieval tracking.
- Containerized the application using **Docker / Docker-Compose**,facilitating environment setup, portability, and easy deployment.
- Prepared **CI/CD configuration** (with automated build & deployment workflows) to support scalable, maintainable cloud hosting on Microsoft Azure.
- Designed the system for **scalability and modularity**: separation of ingestion, retrieval, generation, storage, and API layers; environment-driven configuration; and readiness for large document corpora.

CERTIFICATES

Oracle Cloud Infrastructure 2025 Certified Data Science | CERTIFICATE 09/2025
Oracle Cloud Infrastructure 2025 Certified Generative AI Professional | CERTIFICATE 09/2025
Oracle Cloud Infrastructure 2024 Certified AI Foundations Associate | CERTIFICATE 08/2024
Fundamentals of Deep Learning (NVIDIA) | CERTIFICATE 06/2024
Generative AI with Diffusion Models (NVIDIA) | CERTIFICATE 04/2025
Building RAG Agents with LLMs (NVIDIA) | CERTIFICATE 04/2025
Building Transformer-Based NLP Applications (NVIDIA) | CERTIFICATE 06/2024
Applications of AI for Predictive Maintenance (NVIDIA) | CERTIFICATE 06/2024
XR Foundation Bootcamp (XR Bootcamp) | CERTIFICATE 04/2024
Google Cloud Skill Badges (Google Cloud) | BADGES 2024/2025