Marwan Mashra

Paris, France +33 7 67 81 14 49 marwan.mashra@gmail.com
mashra.fr in/marwanmashra MarwanMashra

Machine Learning | Deep Learning | Computer Vision | Image Processing | NLP | Ethics of AI
Software development | MLOps | Data Mining| Data Pipeline | Large-scale distributed data processing
Proficiency in: *Python* • *JavaScript* • *C/C++* • *SQL*

Professional Summary

I was selected among **500 Yemeni students** for the ***HFHD Scholarship*** to pursue my studies in France. I'm currently pursuing a **master's degree in AI** at the university of **Paris-Saclay**. I have **5 years** of experience in computer science, successfully completed **4 internships**, won a **1st Prize** in a data science **Hackathon**, spearheaded, and contributed to over **20 projects**, and was student delegate for **2 years**.

Career Highlights

Machine learning: worked on different fields including Deep Learning, Computer Vision, NLP, Ethics of AI.
Research: conducted research on Ethics of AI during which I studied and reviewed many research papers.
Development: built, tested, and delivered different software, writing clean, and well-documented code.
Leadership: spearheaded several teams of 2 to 4 members in projects, and student delegate for 2 years.

Education

Master – Artificial intelligence, *University of Paris-Saclay, France*	2023 – 2021
Bachelor – Computer science, *University of Montpellier, France*	2021 – 2018
DELF B2 - French language, *Accent Français, France*	2018 – 2017

Awards

I won **1st place** for Technical Excellence among **16 teams** in **Hackathon** awarded by **TotalEnergies**. In under 48h, we developed an AI solution using **Computer Vision** to detect a car in an image and estimate its CO2 emission, obtaining the **highest score**. See on ***GitHub*** or ***Hackathon Website***

Key Skills

- ✓ Leadership
- ✓ Teamwork
- ✓ Proactivity
- ✓ Autonomy
- ✓ Adaptability
- ✓ Time management
- ✓ Humble & always hungry to learn more
- ✓ Good programming & problem-solving skills
- ✓ Good English & French communication skills

Technical Skills

Programming languages: *Python* • *JavaScript* • *C/C++* • *SQL* • *Java* • *PHP* • *HTML* • *CSS* • *OCAML* • *LaTeX*
Machine learning: *PyTorch* • *Scikit-Learn* • *TensorFlow* • *Keras* • *OpenCV* • *Transformers* • *Spacy* • *Gensim*
Python Tools: *NumPy* • *Pandas* • *Matplotlib* • *Pillow* • *SciPy* • *NLTK* • *Seaborn* • *Plotly*
Tools & Frameworks: *Docker* • *Spark* • *Hadoop* • *Map/Reduce* • *Flask* • *MongoDB* • *Elasticsearch* • *jQuery*
Collaboration Tools: *Git* • *Jira* • *Trello* • *Asana* • *Overleaf* • *Deepnote*
Environnements & IDE: *Linux* • *VScode* • *Jupyter notebooks* • *Eclipse* • *PyCharm*

Languages

English: Fluent (TOEIC 920) **French:** Fluent (DELF B2) **Arabic:** Native language

PROFESSIONAL EXPERIENCE

Research Intern, *CNRS – LISN research lab, Paris, France* **June - July 2022**

Use **Computer Vision** and **Augmented Reality** to implement **real-time style transfer on AR glasses** and explore its ability to help patients in therapy

- Starting soon…

Research Student, *University of Paris-Saclay, Paris, France* **Nov. - Jan. 2022**

A research internship at the university during which I conducted **research on Ethics of AI**, and **prepared data** for a TRUSTAI, a challenge on CodaLab about **discriminatory biases** in Machine Learning

- Investigated and reviewed the scientific literature on discriminatory biases and evaluation metrics.
- Collected, explored, and analyzed multiple datasets for the TRUSTAI challenge.
- Built a pipeline to preprocess and format the dataset and produced several visualizations.

Software Engineer Intern, *LIRMM – Research Lab, Montpellier, France* **July – Aug. 2021**

Building a **vocal assistant** with a **microservices architecture** using **Python** & **Docker**

- Designed, implemented, and tested a microservices-based vocal assistant using Python and Docker.
- Built an interactive web user interface to visualize and interact with the data base.
- Delivered a prototype on a Raspberry Pi and a 20-page documentation.

Web Developer Intern, *LIRMM – Research Lab, Montpellier, France* **June – July 2020**

Building a digital version of **Labo DataViz** using **JavaScript**, **Flask**, and **MongoDB**

- Designed a digital version of Labo DataViz presented in 2019 by UQAC.
- Coordinated a team of 3 to build, test, and deliver a responsive web platform in under 2 months.

App Developer Intern, *PERFORM VR, Montpellier, France* **June 2019**

Optimized and improved a virtual reality **android application** with **Java**

- Optimized a virtual reality mobile app and improved the correction algorithm of the gyroscope drift.

TECHNICAL PROJECTS

Variational Autoencoder, *Python • PyTorch • NumPy • Pandas • Matplotlib*

Implemented a Variational Autoencoder with PyTorch to generate MNIST images

Deep Network, *Python • NumPy • Matplotlib • Backpropagation • Momentum SGD Optimizer*

Implemented a Deep Neural Network (Multi-Layer Perceptron) from scratch without PyTorch

Hyponymy Extraction, *Python • Spacy • Prodigy • NLTK • Scikit-Learn • Pandas • NumPy*

Built & trained a model for Named-entity recognition (NER) and extracted hyponymies in tech patents

Twitter Streaming Data, *Python • Spark • Elasticsearch • TwitterAPI • Socket • TextBlob*

Used Spark & Elasticsearch to process, cluster, index, and query tweets from a data stream

Sentiment Analysis, *Python • Transformers • Gensim • NLTK • Scikit-Learn • Imblearn • Plotly*

Performed a sentiment analysis for Amazon reviews using Bert, embeddings, and other NLP approaches

Aerial Challenge, *Python • TensorFlow • OpenCV • Scikit-Learn • NumPy • Pandas • Matplotlib*

Achieved 91.4% accuracy on a multi-class geographic image classification challenge on CodaLab

To-Be Challenge, *Python • Scikit-Learn • XGBoost • LightGBM • NumPy • Pandas • Plotly*

Achieved 77% accuracy and ranked top 3 in a medical data classification challenge on CodaLab

Geoscape, *JavaScript • jQuery • Python • Flask • RedditAPI • MongoDB • HTML• CSS*

Developed all the frontend & some of the backend of a web platform for scraping images from Reddit