

## Data Scientist/Engineer

### Education

From  
October  
2021 to  
September  
2023

● **Master's in Computer Science, Specialization in Machine Learning for Data Science (MLSD)**

**Université Paris Cité Paris**

- Curriculum: Big Data, NLP, Machine Learning, Deep Learning, Times Series, Recommender Systems.
- Minimum of one project per module, plus extended projects.

From  
September  
2017 to  
September  
2022

● **Engineer's degree in Computer Science**

**École nationale Supérieure d'Informatique ESI ex INI Algiers**

- Curriculum: High-Performance Computing, Advanced Database Systems, Data Structures, Operations Research & Graph Theory.
- Active member of student associations like [School of AI Algiers](#) and [GDG Algiers](#), involved in workshop facilitation.

### Work experience

From  
November  
2022 to  
September  
2023

● **Data Scientist : GenIA Applications in the Medical Field (NLP, Statistics, Ethics)**

**SogetiLabs** Issy-les-Moulineaux, France

- Improved reliability of extracted information by a range of 5 to 20%.
- Designed and implemented robust, reusable, and industrializable NLP pipelines for processing medical reports, showcasing technical skills and creativity.
- Critically analyzed methods, obtained results, and future developments, demonstrating rigor and critical thinking.
- **Tools:** Python, Pandas, HuggingFace, Numpy, Matplotlib, Streamlit, Docker.

From  
November  
2022 to  
January  
2023

● **Data Consultant: Biomarker Identification in EEG Time Series**

**Spectre Biotech** (Client Project) - Paris, France

- Key contribution to a long-term collaboration with the client.
- Developed a new architecture for EEG data processing and identification of certain pathologies.
- **Tools:** Python, Pandas, ScikitLearn, Plotly, Streamlit.

From May  
2023 to July  
2023

● **Data Consultant: GenIA Application for Test Case Automation**

**Coface** (Client Project) - Paris, France

- Developed an LLM pipeline to automatically translate functional specifications into test cases and compared open-source and proprietary model results.
- Strong acceptance from the client of the proposed solution, and have been involved in internal team training for its adoption.
- **Tools:** Python, HuggingFace, Streamlit.

### Personal and Academic Projects

**Engineer's Thesis Project: Automatic Extraction of Vehicle License Plate Numbers.**

- Implemented and Deployed the a deep learning architecture on GCP, created a Rest API ensuring accessibility.
- Python, Pytorch, Tensorflow, VM, Matplotlib, Flask, CNN, RNN, Attention.

**Kaggle Competition: Deep learning architecture for predicting sleep periods.**

- Python, Tensorflow, Time Series, LSTM, CNN, BiLSTM, Matplotlib, Pandas, Parquet.

**AutolibDZ: Self-service vehicle rental platform.**

- Contributed to the implementation of the front-end in production.
- React.js, CI/CD, Scrum, UML, UI/UX.

**MedicoDZ: Medical appointment management platform.**

- Implemented the database, REST API, and contributed to the mobile application.
- PostgreSQL, UML, Express.js, Kotlin, Sequelize, CI/CD.

### Skills

**GCP, GCP ML Services, Docker, Kubernetes**

- Internal training at SogetiLabs ([Cloud Digital Leader](#) level) and significant experience with Docker and GCP.
- [GCP Intermediate ML: TensorFlow on Google Cloud](#)

**SQL, MySQL, PostgreSQL, Cassandra, Spark, Hadoop, HDFS, Java**

- Significant experience in designing, implementing, and deploying relational databases.
- Academic experience in distributed solutions: [example](#).

**Python, R, Pandas, Numpy, Matplotlib, Tensorflow, Pytorch, Statistiques, Scikit-Learn, HuggingFace**

Substantial enterprise experience and through personal projects.

### Interests

**Chess**

I adeptly practice the London System as an opening.

**Cinema and TV Shows**

Dark is my favorite series.

### Languages

**French**

C2 Level

**English**

C2 Level

### Social networks

**in** @lounes-mechouek

**🐱** @lounesmechouek