

# Mohammed FELLAJI

fellajimohammed@gmail.com | LinkedIn: fellajimed | GitHub: fellajimed | +33662647607

## EDUCATION

<b>Ecole CentraleSupélec - Loria lab</b> <i>PhD on Bayesian Deep Learning supervised by Miguel Couceiro and Frédéric Pennerath</i>	Metz, France 01/2022 – 07/2025
• Title: “Epistemic Calibration for Bayesian Deep Learning: Principles, Issues and Solutions”	
<b>Ecole CentraleSupélec</b> <i>Engineering School - Major: Data Science</i>	Metz, France 09/2017 – 04/2021
<b>University of Lorraine</b> <i>Master of Mathematics (M2): fundamental and applied mathematics</i>	Metz, France 09/2020 – 04/2021
<b>Ecole Polytechnique Fédérale de Lausanne (EPFL)</b> <i>Exchange program</i>	Lausanne, Switzerland 02/2019 – 07/2019
<b>Prep School</b> <i>Maths-physics Major (Preparatory classes for the top French engineering schools)</i>	Tangier, Morocco 09/2015 – 06/2017

## EXPERIENCE

<b>Teacher assistant</b> <i>CentraleSupélec/IUT Metz</i>	Metz, France 09/2022 – 04/2024
• Statistical Models, Algorithms for Data Science	
• Introduction to algorithms (VBA), Introduction to Python	
<b>Deep Learning Research Intern</b> <i>Orange Labs</i>	Lannion, France 05/2021 – 10/2021
• <b>Keywords:</b> Audio Processing, Audio Embeddings, Large Pre-trained Models, Fine-tuning, Deep Learning.	
• <b>Researched and applied</b> deep learning techniques, including large pre-trained models and fine-tuning, to generate <i>audio embeddings</i> from extensive datasets.	
• <b>Analyzed and compared</b> novel audio embedding approaches against traditional methods, demonstrating superior performance and efficiency.	
• <b>Demonstrated significant results</b> by testing audio embeddings on DCASE datasets, enabling the use of smaller models for classification and regression tasks.	
<b>Data Science Intern</b> <i>AiSight GmbH</i>	Berlin, Germany 03/2020 – 08/2020
• <b>Keywords:</b> IoT, Predictive Maintenance, Anomaly Detection, Vibrations Data, Gaussian Processes, Deep Learning.	
• <b>Engineered</b> classification and regression models using high-frequency IoT sensor data for <b>predictive maintenance</b> . Utilized data visualization, cleaning, signal processing, Gaussian processes, and deep learning algorithms, presenting results directly to clients.	
• <b>Developed and optimized anomaly detection</b> for continuous vibration data; automated client reporting via Slack/email for actionable insights.	
<b>Machine Learning Intern</b> <i>Zhor-Tech</i>	Nancy, France 08/2019 – 02/2020
• <b>Keywords:</b> IoT, IMU Sensors, Gait Analysis, Machine Learning, Signal Processing, Data Fusion.	
• <b>Analyzed and processed</b> IMU sensor data for <b>gait analysis</b> , developing multi-class classification models that predicted human movement from 3 distinct classes, enhanced by data augmentation.	
• <b>Developed</b> a precise penalty kick simulation algorithm in Python using IMU sensor data, presented at <i>CES 2020</i> in partnership with <b>Ubisoft</b> . Collaborated with the firmware team to develop and test a C version.	
• <b>Optimized C code</b> for embedded systems, reducing computation error from 20% to 3% and significantly improving memory usage.	

## PAPERS

<b>On the Calibration of Epistemic Uncertainty: Principles, Paradoxes and Conflictual Loss</b>	ECML PKDD 2024 <i>arXiv: 2407.12211</i>
<b>The Epistemic Uncertainty Hole: An issue of Bayesian Neural Networks</b>	CAp 2023 <i>arXiv: 2407.01985</i>

## PREVIOUS PROJECTS

<b><a href="https://github.com/fellajimed/DocSeer">https://github.com/fellajimed/DocSeer</a></b>	<i>Hugging Face, LangChain, ChromaDB, Faiss, Pydantic ...</i>
• An intelligent PDF analysis tool that allows you to summarize documents and ask questions about their contents using natural language. It leverages modern language models to provide fast, accurate insights from complex files.	
<b><a href="https://github.com/fellajimed/FLAIR">https://github.com/fellajimed/FLAIR</a></b>	<i>Python, PyTorch, tifffile, Scikit-learn</i>
• Developed and implemented modular Encoder-Decoder and U-Net models for <b>image segmentation</b> on the FLAIR challenge. Utilized data patching for efficient processing of large inputs and to enhance batch diversity. The implementation allow streamlining training, logging, and result aggregation.	
<b><a href="https://github.com/fellajimed/Conflictual-Loss">https://github.com/fellajimed/Conflictual-Loss</a></b>	<i>Machine Learning, PyTorch, Yaml, NumPy ...</i>
• Implementation of the published papers. Developed code for training models and aggregating results based in a modular fashion based on a Yaml file.	
<b><a href="https://github.com/fellajimed/PyPrune">https://github.com/fellajimed/PyPrune</a></b>	<i>Python, PyTorch, ONNX</i>
• A Python library to prune and quantize PyTorch models.	
<b>LLMs from scratch (WIP)</b>	<i>Python, PyTorch, uv</i>
• Implementing and training LLMs in Python from scratch.	
<b>Tweetoscope</b>	<i>Machine Learning, Python, C++, Apache Kafka, Kubernetes, Docker, Git</i>
• The goal of the project is to detect as soon as possible (simulated) tweets that are likely to become popular, where the popularity of a tweet is defined as the number of times this tweet will be retweeted.	
<b>Time series analysis of car accidents in France</b>	<i>R, Python</i>
• Conducted a time-series analysis of daily accident data in France (2005-2016), identifying and incorporating seasonality into SARIMA models for accurate future predictions.	
<b>Databases Project</b>	<i>Oracle SQL Developer, Python, Java</i>
• Based on an Airbnb dataset, the goal is to design a database schema, parse, clean, and load the data into a DBMS, write and optimize queries, and, finally, implement an interface that will access the database and offer an interactive experience querying parts of the database.	

## SKILLS

<b>Programming Languages:</b> Python, C++, SQL, R, Matlab, Java.
<b>Libraries:</b> PyTorch, TensorFlow, Scikit-learn, Pandas, NumPy, SciPy, Matplotlib, Seaborn ...
<b>Math:</b> Statistics, probability, calculus, time series, signal processing, regression, optimization, stochastic processes.

## ABOUT ME

<b>Languages:</b> English (advanced), French (bilingual), Arabic (mother tongue), German (beginner).
<b>Hobbies:</b> Football, Handball, Badminton, Running, Foosball, Reading, Traveling.
<b>Personality:</b> Autonomy, Innovation, Organization, Team spirit, Project management.
<b>Dev environment:</b> Python, PyTorch, NumPy, Joblib, Git, Vim/NeoVim, Tmux, Linux.

## AWARDS

Recipient of the Moroccan Government Merit Scholarship (2017).
Second place at the maths Olympiad organized by the Spanish school “Ramón Y Cajal” in Tangier (2012).