

# Chris Essomba

## Data Scientist / Machine Learning Engineer

✉ cessomba.ir2025@esaip.org ☎ +33 7 79 45 59 83 ⚡ [github.com/chrisessomba](https://github.com/chrisessomba) 🌐 <https://chrisessomba.vercel.app/>

### Professional Experiences

---

#### French National Institute for Industrial Environment and Risks

*Data Scientist - Apprenticeship*

October 2024 – present

Verneuil-en-Halatte, France

- Developed a Deep Learning model to predict the hazard level of chemical substances based on their molecular structure.
- Designed and deployed an ETL pipeline leveraging Large Language Models (LLMs) to automatically extract toxicological data from scientific documents.
- Built a substance classification tool based on chemical ontology and structural similarity to support comparative chemical analyses.

#### University Polytechnic of Madrid (CITSEM, research center)

*Research Assistant - Internship*

July 2024 – October 2024 | Madrid, Spain

- Developed a computer vision model for fish detection in experimental videos.
- Implemented a tracking algorithm (Kalman Filter) to monitor trajectories and automatically collect behavioral metrics.

### Academic project

---

#### Credit Scoring Model Development (Random Forest)

March 2025 – May 2025

- Developed a predictive model to automate loan approval decisions, based on exploratory data analysis (EDA) of customer and financial data.
- Performance: Achieved a 0.91 F1-score on the 'Denial' class, demonstrating an excellent balance between precision and recall in risk detection.
- Interpretability (XAI): Identified business risk thresholds using Partial Dependence Plots (PDP), notably revealing that the probability of denial increases sharply when the debt-to-income ratio exceeds 35%.

### Education

---

#### University of Paris Cité

*Master's Degree - Applied Mathematics in Data Science*

September 2025 – Present | Paris, France

#### ENS Paris Saclay

*External Auditor - Mathematics, Learning, Vision (MVA)*

October 2025 – Present | Paris, France

Courses taken: Advanced Deep Learning, Reproducible Research Project

#### ESAIP, engineering school

*Engineer's Degree - Computer Science (Focus on Big Data and AI)*

September 2022 – September 2025

Angers, France

### Technical skills

---

#### Programming Languages

- Python, R, SQL, Spark, NoSQL, Java

#### Machine Learning Toolkit

- PyTorch, Scikit-Learn, NLP, LLM, RAG

#### Data Visualization

- Matplotlib, Seaborn, Power BI, Tableau

#### Cloud Computing

Azure cloud (DataFactory, Databricks, Kubernetes)

### Certifications

---

Azure AI Fundamentals (AI-900) | Azure Data Engineer Associate (DP-203) | IBM Data Science

### Languages

---

French – Native language | English – Advanced level C1 | Spanish – Advanced level C1