

# Sethulakshmi Babu

Master's Student – Data & AI | ML

Paris, France | +33 7 80 79 07 52 | sethulakshmikb760@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## Professional Profile

Data/AI engineering student (ESILV, Paris) seeking an end-of-studies AI/Software Engineering internship. I ship production ML/LLM services with clean design, tests, and reviews: Flask microservice (Random Forest + Gemini) processing 73k+ clinical cases with auto JSON/PDF; BERT resume matcher (live API); agentic scraper (12 rec/min, 95 percent uptime, Pydantic). Stack: Python, PyTorch/TensorFlow, scikit-learn, Hugging Face. English C1, French B1. Available ASAP.

## Skills

**AI, GenAI & Deep Learning:** Generative AI, deep learning and natural language processing, HuggingFace Transformers, PyTorch, TensorFlow, LLM Fine-tuning, Agentic AI, BERT Embeddings, Semantic Search.

**Data Science & Machine Learning:** Python (Pandas), Scikit-Learn, Machine Learning, Data Science, Data Visualization (Matplotlib, Plotly), SQL, algorithms, Computer Science

**Tools & Methodologies:** Cloud-Based Services, Git, Agile Development Methodologies, Docker, CI/CD, Vertex AI Pipelines.

**Soft Skills:** Problem-Solving, Teamwork, Collaboration, Communication, Technical Documentation.

## Work Experience

**CMI Consulting | Medipatient AI** Paris, France — 10/2025–01/2026  
**Student Consultant – AI Developer** | Part of the Industrial Innovation Project – M2  
[GitHub](#)

- Built hybrid ML pipeline (Random Forest + optional LLM) for 73k+ clinical cases across 8 disease classes.
- Applied data augmentation, class balancing, and feature verification → improved model stability, precision, and fairness.
- Automated model management and feature verification → reduced manual setup by 30% and ensured reproducibility.
- Delivered actionable outputs including confidence, severity, and urgency scores supporting clinical review.
- Explored literature and SOTA approaches to enhance model design and performance.

## Projects

**Path Optimization using AI | Smart Routing** 09/2024–12/2024  
[GitHub](#)

- Built A\*/Dijkstra hybrid → **35% faster routing**, **92.7% optimal paths** on constrained graphs (traffic, cost).
- Multi-objective cost function → transport/logistics trade-off analysis (time × distance × tolls).
- Interactive NetworkX dashboard → strategy comparison on real/synthetic maps.
- Dockerized Python pipeline → production-ready for fleet/rail scheduling.

**OpsPilot | AI-Powered DevOps Assistant** Personal Project  
[GitHub](#)

- Built AI-assisted log analysis combining rule-based logic and LLM fallback → reduced triage time for high-severity incidents.
- Developed human-in-the-loop remediation → ensures coverage of all known incident types.
- Implemented asynchronous FastAPI backend → supports concurrent processing and structured reporting.
- Generated actionable incident summaries with severity classification → improved operational decision-making.

**Intelligent Document Auditor | Contract Analysis** 08/2025–11/2025  
[GitHub](#)

- Production Streamlit app → processes **13 PDF pairs/min** (4.7s avg), analyzes 200+ page contracts.
- LangChain + Llama3.2 RAG → forensic insights extraction via PyPDF, offline Ollama processing.
- Battle-tested Reference.pdf (66KB) vs Suspect.pdf (111KB) → discrepancy detection workflows.
- Direct context injection → scales to enterprise document audit infrastructure.

## Education

**Master of Science in Engineering — Data and AI** 2024–Present  
ESILV Engineering School, Paris, France

**Bachelor of Science in Computer Science** Graduated 2023  
Mahatma Gandhi University, India  
Grade: 88.8% (Distinction)

## Certifications & Languages

Dataiku ML Practitioner — 2026  
Dataiku Generative AI Practitioner — 2026  
IBM watsonx.data Technical Essentials — 2025

**Languages:** English (C1), French (B1), Malayalam (Native)