

# ARIJIT SAMAL

+33 0755028593 | [arijit.samal@student-cs.fr](mailto:arijit.samal@student-cs.fr) | [linkedin.com/in/arijit-samall](https://linkedin.com/in/arijit-samall) | [github.com/1-ARIjitS](https://github.com/1-ARIjitS) | [devpost.com/arijits19](https://devpost.com/arijits19) | [arijitportfolio-sigma.vercel.app](https://arijitportfolio-sigma.vercel.app) | [scholar.google.com](https://scholar.google.com)

## EDUCATION

### Erasmus Mundus Masters in Big Data Management and Analytics (BDMA)

2023 – Present

- CentraleSupélec (CS), Université Paris-Saclay, Paris, France – M2 in BDMA

### Indian Institute of Science Education and Research (IISER), Bhopal, India

2019 – 2023

- Bachelor's in Electrical Engineering and Computer Science (EECS) and Minor in Data Science and Engineering (DSE)
- G.P.A.: 9.57 / 10.00

## EXPERIENCE

### Capital Fund Management (CFM)

March 2025– present

Data Science Intern

Paris, Île-de-France

- Building a multi-agent system to auto-resolve data pipeline alerts for the data referential equity team
- Technologies: Python, OracleDB, LangChain, LangGraph, Google ADK, FastAPI, Streamlit, A2A

### MobilityDB

July – September 2024

Open Source Developer

Brussels, Belgium

- Improved JMEOS, a Java binding for the MEOS spatiotemporal library
- Contributed 30K+ lines of code to JMEOS and MobilityDB repositories
- Boosted testing coverage by 70% using JUnit for MEOS data types
- Automated documentation deployment using GitHub Pages, streamlining API visibility for 500+ users
- Built CI/CD pipelines with GitHub Actions, cutting build and integration times by 30%

[Project Link](#)

### Health Technologies Lab (HTL), IBME, University of New Brunswick (UNB)

May – October 2023

Research and Development Intern

Fredericton, Canada

- Worked on Translating Foot Pressure Maps to 3D Human Poses
- Captured foot pressure maps using 100Hz tiles; mapped to 3D poses with 33 keypoints
- Used video from 8 cameras as supervision; developed Encoder-Decoder, CRNN, and CNN+LSTM architectures
- Evaluated models using MPJPE and MSE, enabling non-invasive person identification with 95% accuracy

[Project Link](#)

## SKILLS

- **Languages and Programming:** C/C++, Python, Java, OOPs, Data Structures and Algorithms (DSA)
- **Data Science:** NumPy, Pandas, Seaborn, Matplotlib, Scikit-learn, PyTorch, TensorFlow, Machine Learning, Deep Learning, Data analytics, Data visualization, OpenCV, Langchain, Streamlit, LLM, RAG, GEN-AI, Langchain, Agents, ADK, Langgraph
- **Data Engineering:** SQL, PostgreSQL, Apache Airflow, PySpark, Docker, Shell, GCS, MinIO, Big data
- **Graph databases:** SPARQL, Cypher, Neo4J, GraphDB, OrientDB

## PROJECTS AND PUBLICATIONS

### MediReels

October 2024

LLM, GEN-AI

[Project Link](#)

- Mistral X Alan Hackathon; Developed a platform for generating engaging short videos on trending medical topics.
- Reduction in content creation time from several days of research, creation, and editing to under 10 minutes.

### SpicyBytes

February – July 2024

BIG DATA, GEN-AI, VLM, LLM, DL, ML

[Project Link](#)

- Scraped 1M+ listings from 1,000+ stores in 60+ PIN codes in Barcelona to cut food waste for 200K+ students; pitched at UPC.
- Scalable Airflow orchestrated Delta Lake Medallion architecture with PySpark on MinIO (temporal) and GCP (persistent)
- Automated inventory management via multilingual OCR; dynamic pricing; Prophet sales forecasting; Llama-driven recommendations; sentiment analysis.

### Klinic

May 2024

LLM, KG, GEN-AI, RAG, DL

[Project Link](#)

- 1st place at MLH HackUPC 2024; Platform assisting clinicians and researchers in navigating previous clinical trials
- Created a knowledge graph with 500K+ entries from NIH MedGen and clinical trials datasets
- Enhanced accuracy of queries by using KG embeddings, pykeen training, and Graph RAG, leading to a reduction in hallucinations while summarizing trials and extracting statistical insights using gpt

### ThermalVision

December 2022 – August 2023

Published in Elsevier ScienceDirect Smart Health Journal

[Publication Link](#)

- Developed and deployed real-time temperature tracking system in crowded environments using edge devices
- Achieved 94% thermal face detection accuracy and R2 score of 0.96 in real-time temperature estimation

## ACHIEVEMENTS

- Selected as 1 of 24 global recipients of the Erasmus Mundus Scholarship for the BDMA Master's program.
- Awarded Overall 1st Place at MLH HackUPC 2024, Europe's largest hackathon.
- Awarded 3rd place in the Revolut Challenge for AI in Financial Applications at MLH HackUPC 2025.
- Winner of the InterSystems Challenge for Best Use of GenAI and MLH Best Use of MATLAB at HackUPC 2024.
- Received the Mitacs GRI Fellowship (2023); completed research internship at the UNB, Canada.
- Selected as 1 of 50 participants for the Zama X Hugging Face Privacy-Preserving AI Hackathon.
- Achieved Global Rank 108 in the HackerEarth Get a Room ML Hackathon.
- Secured 1st Place in the ArmaCode Alpha 1 Coding Contest at IISER, Bhopal.