ARIJIT SAMAL

+33 0755028593 | arijit.samal@student-cs.fr | linkedin.com/in/arijit-samal1 | github.com/1-ARIjitS | devpost.com/arijits19 | arijitportfolio-sigma.vercel.app | scholar.google.com

EDUCATION

CentraleSupélec, University of Paris Saclay, Paris, France *2023 – 2025*
- MS in Big Data Management and Analytics (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, G.P.A.: 9.57 / 10.00

EXPERIENCE

Capital Fund Management (CFM), Hedge Fund *March 2025 – present*
Data Scientist *Paris, Île-de-France*
- Built a multi-agent system to auto resolve data pipeline alerts for the Data Referential Equity team *Project Link*
- Engineered an ensemble RAG pipeline (Graph RAG, RAPTOR, hybrid search, reranking, Agentic RAG)
- Created autonomous coding agents using ReWOO, LLMCompiler, and LATS planners
- Implemented an agentic swarm orchestration with dynamic handoffs between agents
- Achieved 97% RAG accuracy and 94% agent performance on production CFM alerts
- Reduced mean time to resolution from 60 min to 2 min (fast) / 5 min (normal), cutting handling time by 92%

MobilityDB *July – September 2024*
Open Source Developer *Brussels, Belgium*
- Improved JMEOS Java bindings for the MEOS spatiotemporal library, a PostgreSQL extension *Project Link*
- Contributed 30K+ LOC across JMEOS, pyMEOS, and MobilityDB; increased JUnit test coverage by 70%
- Built CI/CD pipelines, reducing build and integration times by 30% and exposing documentation to 500+ users

Health Technologies Lab (HTL), IBME, University of New Brunswick (UNB) *May – October 2023*
Deep Learning Researcher *Fredericton, Canada*
- Translating Foot Pressure Maps to 3D Human Poses *Project Link*
- Captured foot pressure maps using 100Hz tiles; mapped to 3D human 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 loss, enabling non-invasive person identification with 95% accuracy

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, OpenCV, Streamlit, FastAPI, LLM, RAG, GenAI, Langchain, Agents, ADK, Langgraph
- **Data Engineering:** SQL, PostgreSQL, Apache Airflow, PySpark, Docker, Shell, GCS, MinIO, Big data, Git, GitLab
- **Graph databases:** SPARQL, Cypher, Neo4J, GraphDB, OrientDB

PROJECTS AND PUBLICATIONS

SpicyBytes *February – July 2024*
BIG DATA, GEN-AI, VLM, LLM, DL, ML *Project Link*
- Scraped 1M+ product listings from 1K+ stores in 60+ PIN codes in Barcelona to cut food waste, pitched at UPC
- Airflow orchestrated Delta Lake Medallion architecture with PySpark on MinIO (temporal) and GCP (persistent)
- Automated inventory management via VLM, dynamic pricing, sales forecasting, recommendations, sentiment analysis

ThermalVision *December 2022 – August 2023*
Published in Elsevier ScienceDirect Smart Health Journal *Publication Link*
- Developed and deployed a real-time temperature tracking system in dense environments using edge devices
- Achieved 94% thermal face detection accuracy and R2 score of 0.96 in real-time temperature estimation

ACHIEVEMENTS

- Erasmus Mundus Scholarship (BDMA), 1 of 24 globally
- Mitacs GRI Fellow (2023), research at UNB, Canada
- MLH HackUPC 2024 Overall 1st, InterSystems Challenge Winner for Best use of GenAI, MLH Best Use of MATLAB
- MLH HackUPC 2025 Revolut AI in Financial Applications, 3rd place
- HackerEarth Get a Room ML Hackathon, Global Rank 108