

Md Kamrul Islam

Paris, France — +33 6 52 24 36 55 — mdkamrul.islam@student-cs.fr

kamrul-portfolio-sigma.vercel.app — linkedin.com/in/kamrulkonok — github.com/kamrulkonok

PROFESSIONAL SUMMARY

Results-driven AI and Data Scientist with experience in building LLM, Deep Learning pipelines, and Big Data analytics. Skilled in Python, PyTorch, cloud platforms, with a track record of deliver research and industry projects. Adept at collaborating in multidisciplinary teams to transform complex datasets into actionable insights.

WORK EXPERIENCE

AI Research Intern

Paris, France

Laboratoire Images, Signaux et Systèmes Intelligents, Université Paris-Est Créteil

Apr 2025 – Present

- Building an end-to-end pipeline for extracting security and data-sharing requirements from multimodal inputs to automatically generate BPMN workflows enriched with semantically valid cybersecurity annotations.
- Designed and A/B tested advanced prompt engineering, improving valid security labels extraction rate from ~60% to ~95%.
- Developed an evaluation suite including F1 classification against ground-truth security labels, and LLM-as-a-judge scoring for final output quality.

Technologies: Python, LangChain, LLMs, RAG, React-JS, SecBPMN

Graduate Research Assistant (Project Link)

Gif-sur-Yvette, France

Laboratoire Interdisciplinaire des Sciences du Numérique (LISN), CentraleSupélec

Oct 2024 – Feb 2025

- Designed a deep clustering architecture integrating Group Equivariant CNNs to improve medical image analysis using NIH chest X-ray datasets.
- Showed that encoding geometric symmetries reduced need for explicit augmentation while improving clustering performance.
- Built a scalable multi-GPU pipeline with PyTorch DDP and mixed-precision training.

Technologies: PyTorch, OpenCV, LaTeX, HPC

Software Engineer Intern

Chengdu, China

Chengdu Suncape Co., Ltd

Dec 2020 – May 2021

- Developed and optimized Apache Spark data pipelines for large-scale processing and analytics.
- Improved preprocessing workflows, achieving a ~20% increase in predictive model performance.
- Collaborated in Agile Scrum to ensure best practices in code quality and version control.

Technologies: Scikit-learn, Apache Spark, Agile Scrum, Jira, Git

EDUCATION

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

Paris, France

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

2023 – Present

- Master Thesis: "Data Sharing Policy Recommendation using LLMs" (Ongoing)

Bachelor of Engineering in Software Engineering

Chengdu, China

Sichuan University

2018 – 2022

- Average Score: 87/100
- Bachelor Thesis: "Brain Tumor Detection and Classification using CNN"

TECHNICAL SKILLS

Programming Languages & Frameworks: Python, C/C++(OOP), SQL, JavaScript, React-JS

Artificial Intelligence: Pytorch, TensorFlow, Scikit-Learn, LLMs, GenAI, Langchain, OpenCV, Agentic AI

Big Data & Cloud Platforms: PySpark, Hadoop, Apache Airflow, Microsoft Azure, Microsoft Fabric, HPC

Databases: SQL Server, PostgreSQL, ETL, Neo4j, GraphDB

Data Analytics & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Power BI, Tableau, Streamlit, D3.js

Project Management & Collaboration: Git, Docker, CI/CD, Agile, Scrum, Jira

PROJECTS

DigiScan360 (Project Link)

- Built a competitive intelligence platform and pitched as a startup prototype
- Technologies:** PySpark, LLMs, Microsoft Fabric, Azure Data Factory, Power BI, GraphDB

Feb 2024 – Jun 2024

Anomaly Detection in Diesel Train Cooling Systems (Project Link)

- Developed unsupervised models for anomaly detection in SNCB train cooling systems.
- Technologies:** Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, Tableau

Sep 2023 – Dec 2023

Brain Tumor Detection and Classification Using CNN (Project Link)

- Achieved 98% accuracy in tumor detection using a custom CNN.
- Technologies:** TensorFlow, Keras, Python, OpenCV

Sep 2021 – Dec 2021

AI-Based Disease Prediction System (Project Link)

- Built a full-stack web symptom checker predicting 40+ diseases with ML and DL methods.
- Technologies:** Python, TensorFlow Django, ReactJS, REST APIs

Mar 2021 – Jun 2021

LANGUAGES

English: C2, **Chinese (Mandarin):** B2, **French:** A1