

Md Kamrul Islam

+33652243655 | mdkamrul.islam@student-cs.fr | [linkedin.com/in/kamrulkonok](https://www.linkedin.com/in/kamrulkonok) | <https://github.com/kamrulkonok>

RESEARCH INTERESTS

Big Data Analytics, Machine Learning, Deep Learning, Reinforcement Learning, Natural Language Processing, Supervised/Unsupervised Learning, Multilingual NLP, Retrieval-Augmented Generation (RAG)).

EDUCATION

CentraleSupélec, Université Paris-Saclay

Master Sciences de l'Ingénieur

Gif-sur-Yvette, France

Sept 2024 – Present

Universitat Politècnica de Catalunya

Master Erasmus Mundus in Big Data Management and Analytics

Barcelona, Spain

Feb 2024 – July 2024

Université libre de Bruxelles

Master of Science in Computer Science and Engineering

Brussels, Belgium

Sept 2023 – Jan 2024

Sichuan University

Bachelor of Engineering in Software Engineering

Chengdu, China

Mar 2018 – Dec 2021

WORK EXPERIENCE

Research Intern

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

Paris, France

Apr 2025 – Present

- Thesis Title: "Data Sharing Policy Recommendation with Large Language Models (LLMs)"

Big Data Research Project ([Project Link](#))

Laboratoire Interdisciplinaire des Sciences du Numérique (LISN), CentraleSupélec

Gif-sur-Yvette, France

Oct 2024 – Feb 2025

- Research Title: "Enhancing Self-Supervised Learning for Image Clustering Using Geometric Deep Learning"
- Developed a novel deep clustering architecture that integrates Group Equivariant CNNs to enhance self-supervised learning for medical image analysis using the NIH chest X-ray datasets.
- Demonstrated that encoding geometric symmetries directly in the network architecture eliminates the need for explicit data augmentation while improving clustering performance and generalization.
- Built and optimized a scalable training pipeline using PyTorch's DistributedDataParallel and automatic mixed precision technologies, enabling efficient model training across multi-GPU systems.

Technical Skills: *Pytorch, OpenCV, Latex, HPC*

Software Engineer Intern

Chengdu Suncap Co., Ltd

Chengdu, China

Dec 2020 – May 2021

- Developed and optimized data pipelines using *Apache Spark* to process and analyze large data sets.
- Improved data preprocessing workflows to support algorithm accuracy, achieving a *10% increase* in predictive model performance.
- Collaborated with cross-functional teams using *Agile Scrum* methodology to ensure best practices in code quality and version control.

Technical Skills: *SciKit-Learn, Apache Spark, Agile Scrum, Jira, Git*

TECHNICAL SKILLS

Programming Languages: Python, SQL, Cypher, SPARQL

Machine Learning & Deep Learning: Pytorch, Scikit-Learn, OpenCV

Natural Language Processing: HuggingFace Transformers, spaCy, NLTK, BERT, GPT-based models

Big Data & Data Processing: Apache Spark, Hadoop, Apache Airflow, Microsoft Azure, Microsoft Fabric, HPC

Databases: SQL Server, PostgreSQL, Neo4j, GraphDB

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

Project Management & Collaboration: Agile, Scrum, Git, Jira

PROJECTS

DigiScan360

Feb 2024 – June 2024

Developed a competitive intelligence platform and pitched it as a startup prototype at UPC's entrepreneurship initiative.

[Project Link](#)

Technical Skills: PySpark, LLMs SQL Server, Microsoft Fabric, Azure Data Factory, Power BI, GraphDB, SPARQL

Anomaly Detection in Diesel Train Cooling Systems

Sept 2023 – Dec 2023

Developed unsupervised models to detect anomalies in train cooling systems for the Belgian National Railway Company (SNCB).

[Project Link](#)

Technical Skills: Pandas, Numpy, Matplotlib, Seaborn, scikit-learn, Tableau, Anomaly Detection

PostgreSQL Extension for Chess Game Analysis

Sept 2023 – Dec 2023

Created a custom PostgreSQL extension for storing and analyzing chess games.

[Project Link](#)

Technical Skills: PostgreSQL, C, SQL, Linux, Database Systems, Indexing

Brain Tumor Detection and Classification by Using CNN

Sept 2021 – Dec 2021

Achieved 98% accuracy in brain tumor detection using a custom CNN model.

[Project Link](#)

Technical Skills: TensorFlow, Keras, Python, OpenCV

AI-Based Disease Prediction System

Mar 2021 – June 2021

Built a web-based symptom checker predicting over 40 diseases with machine learning algorithms.

[Project Link](#)

Technical Skills: Python, Django, ReactJS, Machine Learning, REST APIs

EXTRA-CURRICULAR ACTIVITIES

Twelfth European Big Data Management & Analytics Summer School (eBISS 2024)

Padova, Italy

University of Padova

July 2024

- Selected to participate in a 5-day fully-funded summer school sponsored by the European Commission, focused on advanced topics in Big Data, AI, and Business Intelligence.
- Presented a research poster titled “Applying Knowledge Graphs in Retrieval-Augmented Generation (RAG) Systems”.

AWARDS & SCHOLARSHIPS

European Union

Sept 2023 – Sept 2024

Erasmus Mundus Partner Country Scholarship

- Awarded a fully-funded scholarship (1 of 24 selected from 837+ applicants), covering tuition, living expenses, and academic support for the Erasmus Mundus Joint Master Degree program.

Chinese Government

Mar 2018 – Mar 2022

The Belt and Road Initiative Scholarship

- Received a comprehensive 4-year scholarship covering tuition, accommodation, and academic resources during undergraduate studies.

LANGUAGES

English: C1 (Advanced proficiency), **Chinese (Mandarin):** B1 (Intermediate proficiency), **French:** A2 (Elementary proficiency)