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

Research Interest

Foundation Models, Retrieval-Augmented Generation, Information Extraction, Deep Learning, Computer Vision, Medical Imaging, Knowledge Graphs, and Big Data Analytics.

Education

Erasmus Mundus Joint Master's Degree in Big Data Management and Analytics (BDMA)
Overall Grade: 14.5/20

CentraleSupélec, Université Paris-Saclay — *Gif-sur-Yvette, France*
Master Sciences de l'Ingénieur — *Sept 2024 – Oct 2025*
Thesis: "SecBPMN-GPT: A Hybrid LLM & Rule-Based Framework for Automating Security Annotation in Business Process Models"

Universitat Politècnica de Catalunya (UPC) — *Barcelona, Spain*
Master Erasmus Mundus in Big Data Management and Analytics — *Feb 2024 – July 2024*

Université libre de Bruxelles (ULB) — *Brussels, Belgium*
Master of Science in Computer Science and Engineering — *Sept 2023 – Jan 2024*

Sichuan University — *Chengdu, China*
Bachelor of Engineering in Software Engineering — *Mar 2018 – Jan 2022*
Thesis: Brain Tumor Detection and Classification using CNN *(Best Thesis Award)*
CGPA: 3.64/4.00

Work Experience

Research Engineer Intern — *Créteil, France*
LISSI, Université Paris-Est Créteil — *May 2025 – Oct 2026*
(In close collaboration with CEA–LIST, France)

- Proposed a novel hybrid framework integrating LLM based semantic reasoning and rule-based validation for automated cybersecurity annotated BPMN model generation.
- Optimized annotation accuracy and efficiency through advanced prompt-engineering and retrieval-augmented generation (RAG) across multiple LLMs.
- Constructed a benchmark dataset and evaluation framework supporting quantitative and qualitative metrics.
- Achieved higher annotation accuracy (F1 = 0.516) while reducing annotation time by 95% compared to domain experts.

Tech Stack: *Python, LangChain, LLMs, RAG, Prompt Engineering, React.js, Flask, CI/CD*

Graduate Research Assistant **(Project Link)** — *Gif-sur-Yvette, France*
Laboratoire Interdisciplinaire des Sciences du Numérique (LISN), CentraleSupélec — *Oct 2024 – Feb 2025*

- Developed an unsupervised deep-learning framework combining AlexNet and Group-Equivariant CNNs (G-CNNs) to analyze large-scale medical imaging datasets.
- Improved clustering accuracy and feature robustness by encoding geometric symmetries, mitigating reliance on data augmentation.
- Implemented multi-GPU distributed training pipelines for scalable, high-performance experiments with early disease detection.

Tech Stack: *PyTorch, OpenCV, HPC, CUDA, LaTeX*

Software Engineer Intern — *Chengdu, China*
Chengdu Suncape Co., Ltd — *Dec 2020 – May 2021*

- Developed end-to-end data pipeline for large-scale data ingestion, transformation, and analytics in distributed environments.
- Enhanced data preprocessing and feature-engineering, boosting predictive model accuracy by approximately 20%.
- Collaborated within an Agile Scrum environment to implement best practices in modular design, version control, and continuous integration.

Tech Stack: *Apache Spark, Scikit-learn, SQL, Agile Scrum, JIRA, Git, CI/CD*

Technical Skills

Programming & Framework: Python, C/C++ (OOP), SQL, JavaScript, TypeScript, React.js, Flask
Data Science & ML: PyTorch, Scikit-learn, Hugging Face Transformers, LangChain, LLMs, OpenCV, CUDA
Big Data & Cloud: PySpark, Hadoop, Apache Airflow, Azure, Vertex AI, Fabric, HPC Environments
Graph & Database Systems: PostgreSQL, SQL Server, Neo4j, GraphDB, Cypher, SPARQL, ETL Pipelines
Analytics & Visualization: Pandas, NumPy, Power BI, Tableau, Streamlit, D3.js
Tools & Collaboration: Git, Docker, REST APIs, CI/CD, Agile Scrum

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
Tech Stack: *PySpark, LLMs, 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
Tech Stack: *Python, SQL, Scikit-learn, Pandas, Numpy, Matplotlib, Tableau*

PostgreSQL Extension for Chess Game Analysis *Sept 2023 – Dec 2023*
Created a custom PostgreSQL extension for storing and analyzing chess games. Project Link
Tech Stack: *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
Tech Stack: *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
Tech Stack: *Python, Scikit-learn, Django, ReactJS, REST APIs*

Publications

C. K. Sah, L. Xiaoli, M. M. Islam and **M. K. Islam**, "*Navigating the AI Frontier: A Critical Literature Review on Integrating Artificial Intelligence into Software Engineering Education*," *2024 36th International Conference on Software Engineering Education and Training (CSEE&T)*, Würzburg, Germany, 2024, pp. 1–5. doi:10.1109/CSEET62301.2024.10663054

Extra-Curricular Activities

Twelfth European Big Data Management & Analytics Summer School (eBISS 2024) *Padova, Italy*
University of Padova — Poster — Paper *July 2024*

- Participated in the 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)"*.

Awards & Scholarships

Erasmus Mundus Partner Country Scholarship *Sept 2023 – Sept 2024*
European Union (Fully-Funded by the European Commission)

The Belt and Road Initiative Scholarship *Mar 2018 – Mar 2022*
Chinese Government (Full Undergraduate Funding)

Languages

English: C1 (Advanced proficiency), **Chinese:** B2 (Intermediate proficiency), **French:** A2 (Elementary Proficiency)

References

Available upon request.