

Aung Ko Ko Oo

Kobe, Hyogo, Japan | AungKoKoOo@my.uopeople.edu | GitHub

PROFESSIONAL SUMMARY

Machine Learning Engineer with a research focus in Natural Language Processing for low-resource languages, including published preprints on RAG-based hallucination mitigation and Deep Knowledge Tracing. Experienced building end-to-end deep learning pipelines using Python, PyTorch, and Scikit-learn, with hands-on work in computer vision, sequence modeling, and custom tokenization. Founder of an open-source AI research collective advancing transformer benchmarks for underrepresented language architectures.

RESEARCH AREAS

- Artificial Intelligence, Machine Learning, Natural Language Processing (NLP) for Low-Resource Languages, Computer Vision, Robotics, Human-Robot Interaction (HRI)

PUBLICATIONS

- Aung Ko Ko Oo. (2026). Integrating Natural Language Processing with Deep Knowledge Tracing in Intelligent Tutoring Systems for Low-Resource Languages: A Case Study on Burmese STEM Curricula. Preprint. DOI: 10.21203/rs.3.rs-9991485/v1 - Built an algorithmic pipeline pairing Deep Knowledge Tracing (DKT) frameworks with syntax analyzers to optimize adaptive technical learning structures for the Burmese language.
- Aung Ko Ko Oo. (2026). Mitigating Hallucinations in Large Language Models for Burmese Domain-Specific Question Answering. Preprint. DOI: 10.21203/rs.3.rs-10023313/v1 - Engineered a domain-specific Retrieval-Augmented Generation (RAG) fine-tuning pipeline to reduce factual errors in generative models handling complex, low-resource language inputs.

EDUCATION

Massachusetts Institute of Technology (MIT)
MicroMasters Program in Statistics and Data Science (Online).
University of the People
Bachelor of Science in Computer Science. GPA: 3.8/4.0.

Cambridge, MA, USA
Jan 2026 - Present
Pasadena, CA, USA
Expected May 2027

PROJECTS

Intelligent Tutoring System (ITS) NLP Engine

Apr 2026 – Present

Lead Developer

- Built and deployed deep learning text processing pipelines using Python and PyTorch to evaluate student domain-specific inputs in low-resource environments.
- Integrated custom tokenization frameworks with Scikit-learn classification models to map continuous student response logs for real-time proficiency tracking.

Low-Resource Language Vision and Sequence Benchmarks

Independent Creator

- Implemented NumPy-optimized computer vision pipelines and sequence-to-sequence architectures to handle irregular character layouts and linguistic segmentation.
- Developed standalone tools for dataset ingestion and specialized downstream model evaluation.

EXPERIENCE

Founder and Lead Researcher

Mar 2026 – Present

Burmese Artificial Intelligence Research Institute

Japan / Remote

- Founded an open-source research collective and built data parsing tools and transformer fine-tuning benchmarks for low-resource character architectures.
- Coordinated computational infrastructure to support localized text classification and sequence-to-sequence model development.

SKILLS

Languages: Python, SQL, Bash, HTML/CSS

Frameworks and Libraries: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, OpenCV

Core Methodology: Retrieval-Augmented Generation (RAG), Deep Knowledge Tracing (DKT), Neural Language Modeling, Sequence Engineering