MYOUNGCHUL KIM

Data Scientist

English: Fluent | Japanese: Fluent, JLPT N1 | Korean: Native
kmc8907@gmail.com | 080-9801-7956 | Greater Tokyo Area, Japan | MyoungchulK | myoungchulkim
troopl.com/kmc8907 | 0000-0002-8624-5564

ABOUT ME

I'm an Astrophysicist used to analyzing large datasets to find astronomical signals and recently graduated from Le Wagon Data Scientist & AI boot camp. I'm seeking to utilize programming skills backed by my scientific background and data science knowledge.

EXPERIENCE

Data Scientist

Fracta

November 2024 – Ongoing | Tokyo, Japan

Bringing AI to Infrastructure

November 2024 – Ongoing | Role: Data Scientist | Team size: 4

- **Applied AI and machine learning to large infrastructure**, such as a network of water or gas pipes.
- Predicted which pipes are most likely to fail, **Likelihood of Failure (LOF)**, and provided to the client.
- Implemented **the automation to the ML pipeline** to increase efficiency.

Teaching Assistant for Data Science & AI

Le Wagon

October 2024 – Ongoing | Tokyo, Japan

Data Science & AI Bootcamp

October 2024 – Ongoing | Role: Teaching Assistant | Team size: 13

- **Instructed Python for Data Science**, emphasizing data extraction, manipulation, visualization, and statistics, and linear algebra principles.
- **Provided expert guidance in Machine Learning and Deep Learning**, with a focus on Scikit-Learn workflows and neural network architecture design.
- **Delivered advanced training in ML Engineering**, covering Python package development for GCP and the ethical implications of AI technologies.

Python Data Scientist/Analyst

Turing

July 2024 – January 2025 | Remote

Multilingual SFT and RLHF Implementation for Next-Generation AI

July 2024 – January 2025 | Role: LLM Dataset Creation | Team size: 6

- Performed **Supervised Fine-Tuning (SFT)** and **Reinforcement Learning with Human Feedback (RLHF)** across **English, Korean, and Japanese** datasets to improve AI model alignment with human preferences, enhancing the accuracy and reliability.
- Analyzed model performance, focusing on **Loss Patterns** and **Output-Based Evaluations** to enhance robustness.
- Curated model responses through **Comparative Analysis** and **Preference Pair evaluation**, ensuring accurate and contextually appropriate outputs.

AI Training for Japanese Writers

Outlier

July 2024 – October 2024 | Remote

Fine-Tuning and Validation of AI-Generated Japanese Voice

July 2024 – October 2024 | Role: LLM evaluation | Team size: 12

- Conducted **validation of AI-generated voices**, focusing on voice quality through data augmentation, pitch, and speech rate to ensure natural and contextually accurate outputs.
- Evaluated voice model performance, ensuring clarity, consistency, and cultural accuracy.

Graduate Research Assistant, Ph.D.

International Center for Hadron Astrophysics (ICEHAP)

April 2017 – December 2023 | Chiba University, Japan

Search for Ultra-high Energy Neutrinos from Askaryan Radio Array (ARA) by Template Method

April 2018 – December 2023 | Role: Project Leader | Team size: 6

- Classified astronomical signal by **Principal Component Analysis (PCA)**, after obtaining features from ∼**200 TB** of radio-frequency dataset.
  - Optimized the PCA based on **Frequentist Statistics** and **Pseudo Experiment**.
  - Implemented the **Fast Fourier Transform (FFT)**, **Interferometry**, and the **Matched Filter**, for **feature extraction**.
- Implemented analysis pipeline used by **large CPU & GPU clusters** developed with **scientific Python & C++ packages**. It leads to **wide use by international collaborators**.
- Evaluated results by calculating **statistical significance**, including **systematic uncertainty**, and **Monte Carlo simulation**.

Development of In Situ Antenna Model for Simulation

April 2017 – March 2019 | Role: Data Management | Team size: 4

- Performed **high-precision calibration for the radio-frequency antenna** to measure property.
- **Extracted feature pattern** from raw data after noise & signal analysis.

Graduate Research Assistant, MSc.

Neutrino AstroParticle Physics Lab (NAPPL)

March 2015 – February 2017 | SungKyunKwan University, South Korea

IceCube Camera System to Study Properties of the Antarctic Ice

March 2015 – February 2017 | Role: Project Leader | Team size: 5

- Classified intrinsic camera noise appearing in extremely low-temperature conditions from images using **feature extraction**.
- Developed **Python package** to control the camera by **Raspberry Pi** and **automatic data collection**.
- Evaluated the performance of camera system by applying statistical techniques to **large image datasets**.

PROJECT

Sound to Symphony (AI Music Generation)

Le Wagon Data Science & AI Bootcamp

January 2024 – March 2024 | Le Wagon, Tokyo

Role: Data Management | Team size: 4

- Generates new music by **Recurrent Neural Network (RNN)** that can be easily customizable by musical software
- **Architectures RNN to learn musical patterns** from **large classical music datasets** that are expressed in numerical form.
- Deployed the project into the **Streamlit** by utilizing **FastAPI**
- Built the connection between generated music and musical software Abelton

EDUCATION

Data Science & AI Bootcamp

Le Wagon

January 2024 – March 2024 | Japan

- Thorough study in Python for Data Science, with expertise in data extraction, manipulation, and visualization, backed by a strong foundation in statistics and linear algebra.
- Delving into Machine Learning and Deep Learning, with practical application in building comprehensive workflows utilizing Scikit-Learn and designing neural network architectures.
- Proficiency in ML Engineering, involving the development of Python packages for large-scale data tasks in GCP, and a deep awareness of the ethical considerations surrounding AI deployment.

Completion of Ph.D. program (ABD), AstroParticle Physics

Chiba University

April 2017 – December 2023 | Japan

- Research topic: Search for Ultra-high Energy Neutrinos Using Eight Years of Data from Two ARA Stations by the Neutrino Template Method

Master of Science, AstroParticle Physics

SungKyunKwan University

March 2015 – February 2017 | South Korea

- Thesis title: Performance study of camera system for the IceCube-Gen2 detector

Bachelor in Science, Physics

SungKyunKwan University

March 2011 – February 2015 | South Korea

TECHNICAL SKILLS

Coding Tools: Python | C++ | ROOT | Vim | HTCondor | CVMFS | Latex | G-Collab & Jupyter

Data Analytics: NumPy | SciPy | Scikit-learn | Pandas | SQL | Matplotlib | Seaborn

Modelling: TensorFlow | Deep Learning | Unsupervised Leaning | NLP | CNN | GNN | Time series | Ensemble Methods | Statsmodels | LLM

Deployment: AWS | GCP | Docker | Dagster | FastAPI | Streamlit

Hardware Experience: Electronics | Optics

SELECTED PUBLICATIONS

Journal Articles

- P. Dasgupta, M. S. Muzio, *et al.*, "Progress Towards a Diffuse Neutrino Search in the Full Livetime of the Askaryan Radio Array," *PoS*, vol. ICRC, p. 1226, 2023.
- M. Kim *et al.*, "Enhanced Ultra-High Energy Neutrino Search at the Askaryan Radio Array using Template-based Techniques," *PoS*, vol. ICRC, p. 1148, 2023.
- D. Bose, M. Jeong, K. Woosik, J. Kim, M. Kim, C. Rott, *et al.*, "PINGU camera," *PoS*, vol. ICRC, p. 1145, 2015.

AWARDS

Japanese Government Monbukagakusho Scholarship (MEXT)
Graduate Research Assistant in Ph.D., 2017 – 2020

Teaching Assistant (T.A.) of Korea & Japan Joint Government Scholarship
Teaching Assistant for a freshman Korean students, 2017 – 2018

BK21+ Research student scholarship
Graduate Research Assistant in Msc., 2015 – 2017

Operating Assistant scholarship
Physics Experiment Assistant, 2016 – 2017

CK Research student scholarship
Research Assistant in Bsc., 2014

INTERESTS

Classical Music | Orchestra | Contrabass | Universe | Fourier Transform