Joon Hyeok Kim

Philadelphia, PA hozy@seas.upenn.edu LinkedIn GitHub

Research Interests

Deep Generative AI, Mechanistic Interpretability, Bayesian Optimization, Ethical AI

Education

University of Pennsylvania *Aug. 2024 – May 2026*
Master in Computer and Information Technology *Philadelphia, PA*

- GPA: 4.0/4.0
- ML Coursework: Generative Models, Bayesian Optimization, Ethical Algorithm Design, ML

Sogang University *Feb. 2013 – Feb. 2019*
BA in Economics *Seoul, Korea*

- GPA: 4.04/4.3, Summa Cum Laude, 2nd of the class

Publication

Joon Hyeok Kim, Yonghyun Park, Mattis Dalsætra Østby, and Jiatao Gu. "Grokking of Diffusion Models: Case Study on Modular Addition." *ICLR 2026 Workshop on Deep Generative Models: Theory, Principle, and Efficacy (DeLTa)*. [OpenReview], 2026.

Research Experience

Generative Machine Learning Research (GMLR) Lab *Aug. 2025 – Present*
Supervised by Prof. Jiatao Gu *Philadelphia, PA*

- **Grokking of Diffusion Models: Case Study on Modular Addition (2026)**: Co-led research on Rectified Flow models (DiT backbone) by implementing training framework and conducting mechanistic interpretability (MI) analysis during inference to investigate grokking and algorithmic phase transitions.
- **Literature Review and Theoretical Framework Design**: Conducting extensive literature review and theoretical analysis on Discrete Diffusion (CTMC) and Flow Matching (e.g., DFM, Editflow) to propose efficient multi-token prediction frameworks within AutoRegressive structures.

Independent Research: Diversity-Guided Bayesian Optimization (DGEBO) *Jan. 2025 – Jul. 2025*
Supervised by Prof. Jake Gardner *Philadelphia, PA*

Developed a Bayesian Optimization framework extending DGEMO by implementing posterior variance normalization and diversity-informed sampling to enhance Pareto front approximation, validated through benchmarks on ZDT datasets using `gpytorch` and `pymoo`.

Professional Experience

Software Engineer *Jul. 2019 – Aug. 2024*
LG CNS *Seoul, Korea*

- Managed system integration and ERP development for the LG Electronics Global ERP (GERP) system.
- Developed PL/SQL procedures and Java applications; designed RDBMS structures and optimized complex queries for enterprise-level performance.

Teaching Experience

Graduate Teaching Assistant *Aug. 2025 – Dec. 2025*
University of Pennsylvania *Philadelphia, PA*

- **CIT 5920: Mathematical Foundations of Computer Science**: Led weekly recitations for 40+ graduate students, delivering lectures on discrete mathematics, including set theory, induction, and graph theory.
- Provided technical mentorship through office hours to strengthen students' foundational reasoning and abstract problem-solving capabilities.

Skills and Certificates

ML Theory Studies: Studied independently, covering Decision Trees, Neural Networks, Bayesian Learning, Instance-based Learning, PAC, Analytical Learning (EBL, ILP, Domain Theories), Reinforcement Learning.

ML Packages: PyTorch, TorchVision, Slurm, Wandb, GPyTorch, BoTorch, PyMoo

Programming Languages: Python, PL/SQL, C/C++, Java, JavaScript, R

Database Management: Oracle DBMS, MariaDB, MongoDB

Tools: Linux, AWS Cloud, Microsoft Azure, Docker, Django, Stata, Oracle ERP

Certificates

- AWS Certified Solutions Architect - Associate (SAA-C03), Jul. 2023
- Microsoft Azure AI: Fundamentals (AI-900), May 2023
- Oracle Database 12c Administrator Certified Professional, Feb. 2019

Awards

First Place Winner, NSF HDR ML Hackathon University of Pennsylvania, Dec. 2024
First Place Winner, LG CNS Bootcamp Webservice Contest, Jun. 2019
Second Place, KPC Big Data Analysis Contest, Feb. 2019

Scholarships

Global Korea Scholarship, Ministry of Education, Republic of Korea *Jul. 2023*

OK Bae&Jung Student Scholarship ,OK Bae&Jung Scholarship Foundation *Jan. 2018*

Mirae Asset Exchange Student Scholarship, ParkHyeonJoo Foundation *Feb. 2016*

Extra-Curricular Activities

Instructor, Fife-Penn STEM & CS Academy *Dec. 2024 - Jun. 2025*
Delivered introductory coding lessons using Scratch to students in Kindergarten through 2nd grade at St. Francis Xavier School, fostering early interest in computer science through hands-on activities.

George Washington CSED Week Volunteers *Dec. 2024*
Volunteered as a CS educator for UPenn's high school Python game design workshop, teaching basic programming concepts and simple game development

Advanced Level Econometrics and Financial Economics Study Group *Dec. 2017 – Sep. 2018*
Held weekly seminars with presentations on subjects in econometrics and financial economics.
Subjects covered and presented in econometrics:

- Vector Auto-Regressive Model Analysis Application to Time Series Data
- Cointegration and Vector Error Correction Model
- (Generalized) Auto-Regressive Conditional Heteroskedasticity
- Fixed Effect and Random Effect Model for Panel Data Analysis

Subjects covered and presented in macro-economics:

- Consumption–Capital Asset Pricing Model
- Credit Rationing Model by J. Stiglitz, et al.
- Bank Run Model by D. Diamon, et al.
- Central Banking Theories by D. Romer

Sogang Student Supporters for International Affairs (SSIA) *Aug. 2017 - Jun. 2018*
Mentored international students offering guidance on academic studies (particularly math and economics), introduced aspects of Korean culture, and assisted with university administrative processes.