David (Dowon) Baek

(+1) 617-583-2851 | dbaek@mit.edu | david-baek.github.io

Summary

Second–year MIT EECS PhD student studying Large Language Model (LLM) Interpretability and AI Safety, with a strong expertise in mathematical analysis, machine learning, and algorithm design. Passionate about leveraging quantitative skills to solve challenging real-world problems.

Education

Massachusetts Institute of Technology (MIT) — Cambridge, MA, USA
Ph.D. in Electrical Engineering & Computer Science (EECS), GPA: 5.0/5.0 — *Sep 2023 – Present*
- Advisor: Max Tegmark
- Research Area: Representation Learning, Mechanistic Interpretability, AI Safety

Seoul National University (SNU) — Seoul, Korea
B.S. in Physics and Computer Science, Summa Cum Laude, GPA: 4.23/4.3 — *Mar 2017 – Aug 2023*
- Presidential Award (Ranked **1st** among graduating cohort in College of Natural Sciences)
- Includes two years on leave for compulsory military service (2020–21, Job: Cyber Security Specialist)

Technical Skills

Programming: Python, C/C++, Java, Matlab, Mathematica, LaTeX, HTML, Javascript
Libraries: PyTorch, Tensorflow[†], Numpy, Scipy, QuTiP, etc.

Publications

1. D. Baek, Y. Li, M. Tegmark, "Generalization from Starvation: Representations in LLM Knowledge Graph Learning," manuscript in preparation for submission to ICLR 2025.
2. D. Baek, Z. Liu, M. Tegmark, "GenEFT: Understanding Statics and Dynamics of Model Generalization via Effective Theory," *ICLR 2024 Workshop on Bridging the Gap Between Practice and Theory in Deep Learning*, [arXiv].
3. S. H. Park, D. Baek, I. Park, S. Hahn, "Design of Scalable Superconducting Quantum Circuits using Flip-chip Assembly," *IEEE Transactions on Applied Superconductivity*, 33(5), pp.1-6, 2023 [Link].

Experience

Tegmark AI Safety Group — Dec 2023 - Present
Graduate Research Assistant (Advisor: Prof. Max Tegmark) — *Cambridge, MA, USA*
- Research on representation learning, AI safety, and mechanistic interpretability

Applied Superconductivity Laboratory — Feb 2022 – Feb 2023
Undergraduate Research Assistant (Advisor: Prof. Seungyong Hahn) — *Seoul, Korea*
- Research on design, analysis, and simulation of superconducting quantum system

Honors & Awards

International
- Silver Medal, University Physics Competition, 2018
- Silver Medal, International Junior Science Olympiad (IJSO), 2014

Domestic
- Finalist, Samsung Collegiate Programming Cup (SCPC), 2018
- Finalist, Kakao Code Festival, 2017
- Presidential Science Scholarship, Korea Student Aid Foundation, 2017–2023

Community Service

- Chair of Publicity & Communications Committee @ Ashdown House (MIT Graduate Housing) — Nov 2023 - Present
- Vice President of Publicity @ MIT EECS Graduate Student Association — Jan 2024 - Present