

EDUCATION

Pohang University of Science and Technology
Master of Science in Artificial Intelligence

Sep 2023 - Aug 2025
Pohang, South Korea

Hanyang University ERICA
Bachelor of Science in Electrical Engineering

Mar 2018 - Aug 2023
Ansan, South Korea

EXPERIENCE

Natural Language Processing Group
Graduate Researcher (Advisor: Prof. Gary Lee & Hyounghun Kim)

Sep 2023 - Aug 2025
Pohang, South Korea

- * **EMNLP 2025 Findings:** Explored self-correction mechanisms in code generation and proposed the method, a reinforcement learning-based approach that enables multi-turn code correction through fine-grained and accumulated reward optimization.
- * **ACL 2025:** Explored training-based approaches for visual programming language generation in industrial automation, and developed a two-stage training strategy that combines retrieval-augmented fine-tuning and preference learning.
- * **NAACL 2025:** Explored knowledge-injected prompt compression for retrieval-augmented QA, focusing on automatically generating domain knowledge and integrating it into the document compression process.

Efficient Learning Lab
Undergraduate Researcher (Advisor: Prof. Jaeho Lee)

Jan 2023 - Feb 2023
Pohang, South Korea

- * Reviewed model pruning techniques for neural networks with emphasis on robustness under challenging conditions; implemented a simple toy example to observe pruning-accuracy trade-offs.

RESEARCH INTERESTS

In general, my research interests are in natural language processing, focusing on how data and training frameworks shape language model behavior and alignment with human intent. I am also interested in symbolic reasoning for software engineering, aiming to develop interpretable and reliable models that integrate formal methods with data-driven learning.

SELECTED PROJECTS

Code Generation Jul 2024 - Present
In our *Self-Correction* paper, we proposed a novel online reinforcement learning-based framework that generates Python code and, without external guidance, verifies and corrects errors in the iteratively generated code. This intrinsic capacity for self-correction is typically only manifested in large models, yet ours was trained to endow small models with this ability, enhancing initial response quality while achieving substantial improvements through self-correction across diverse Python-based benchmarks.

Reinforcement Learning Jul 2024 - Present
Direct Preference Optimization: In our *Industrial Program Generation* paper, we developed the automation of software used in real-world industries targeting factory automation, with support from Hyundai Mobis. Because every programmer working in factories implements programs in vastly different ways based on their individual preferences, we implemented an assistant model to help humans implement programs consistently. To achieve this, we implemented preference learning. We emphasized that this project marked the first time a language model learned ladder diagram programs.

Multi-Turn Reward: To enable iterative interaction with the user, we explored and developed a reward function tailored for multi-turn scenarios. My contribution involved extending the REINFORCE optimization-based learning algorithm to multi-turn and implementing an accumulated reward function that aggregates rewards across the entire trajectory. The code is available on [here](#).

Retrieval-Augmented Generation

Sep 2023 - Sep 2024

In our *Prompt Compression* paper, we implemented a novel entity-based prompt compression method that combines predicting masked spans from the encoder with the advantage of causal generation from the decoder to mitigate hallucinations in medical-domain QA. This approach demonstrated computationally efficient and robust performance even in unseen settings.

PUBLICATIONS (* → Equal contribution)

Self-Correcting Code Generation Using Small Language Models

Jeonghun Cho, Deokhyung Kang, Hyounghun Kim, and Gary Geunbae Lee

Findings of EMNLP 2025. [[paper](#)][[code](#)]

Retrieval-Augmented Fine-Tuning With Preference Optimization For Visual Program Generation

Deokhyung Kang*, Jeonghun Cho*, Yejin Jeon, Sunbin Jang, Minsub Lee, Jawoon Cho, and Gary Geunbae Lee

Proceedings of ACL 2025. [[paper](#)][[press](#)]

K-COMP: Retrieval-Augmented Medical Domain Question Answering With Knowledge-Injected Compressor

Jeonghun Cho and Gary Geunbae Lee

Proceedings of NAACL 2025. [[paper](#)][[code](#)]

Medical Domain Retrieval-Augmented Question Answering through Entity-based Context Extraction

Jeonghun Cho and Gary Geunbae Lee

Proceedings of Korea Computer Congress 2024. [[paper](#)]

*Best paper (Top 1.7%)

PATENTS

Data processing device for user prompt-based code automatic generation

Sunbin Jang, Gary Geunbae Lee, Deokhyung Kang, and Jeonghun Cho

Patent No: 10-2025-0072930, Publication Date: 2025-06-04

Method and Apparatus for Entity-Based Retrieval-Augmented Question Answering

Jeonghun Cho and Gary Geunbae Lee

Patent No: 10-2024-0133650, Publication Date: 2024-10-02

HONORS

Best Paper Award

Jun 2024

Korea Computer Congress 2024

Academic Scholarship

Fall 2022, Spring 2023

Hanyang University

TEACHING EXPERIENCE

Principles of Software Construction

Spring 2025

Teaching Assistant

POSCO Youth AI & Big Data Academy

Fall 2024

Teaching Assistant