Gahyun Yoo

padme0421@gmail.com padme0421 gahyun-yoo-96b367346

Education

University of Southern California, MS in Computer Science — Aug 2026 – present

Seoul National University, BS in Linguistics & Computer Science — Mar 2020 – Aug 2025

Experience

Agentic AI Engineer Intern, DeepAuto — Apr 2026 – June 2026

- Major conglomerate EPC automation project
- Proposed & developed data lakehouse system for the EPC domain comprised of a tri-store knowledge base
- Developed lakehouse-connected agents that each boosted recall for instrumentation drawing retrieval by 10%pp
- Containerization for efficient and reliable forward deployment
- Improved object detection by referencing project-specific legend sheets

Research Assistant, SKI-ML Lab (PI: Jay-Yoon Lee), Seoul National University — Sept 2025 – Mar 2026

- [Research Project] Reward Model Finetuning for Diffusion Model Alignment Using Tree Structured Trajectories
- [Government Project] Development of a Korean-English Translation Model for Safety Management of Pharmaceutical Products (Ministry of Food and Drug Safety): As project lead, developed a machine translation pipeline supporting an ensemble of inference strategies using in-context examples and dictionary entries, and instruction tuning & preference optimization training, outperforming several SOTA commercial models
- [TA] Special Topic: Multimodal & Diffusion Models
- [Coursework] Special Topic: Generative Modeling with Diffusion and Flow-Based Models

Research Intern, SKI-ML Lab (PI: Jay-Yoon Lee), Seoul National University — Mar 2023 – Aug 2025

- [Research Project] Application of energy based training with quality estimators as score models for low resource machine translation, resulted in a publication at COLING 2025

Research Intern, NCSOFT NLP Center — July 2022 – Aug 2022

- Developed a decision tree classification algorithm for distinguishing literal and non-literal translation, using linguistic features, alignment score from text embedding similarity, and multilingual knowledge graphs (>90% accuracy on open source English-Chinese dataset)

Publications

Improving NMT Models by Retrofitting Quality Estimators into Trainable Energy Loss — Jan 2025
Gahyun Yoo, Jay-Yoon Lee
(International Conference on Computational Linguistics)

BridG MT: Enhancing LLMs' Machine Translation Capabilities with Sentence Bridging and Gradual MT — July 2025
Seung-Woo Choi, *Gahyun Yoo*, Jay-Yoon Lee
(Findings of the Association for Computational Linguistics)

Generalizing Numerical Reasoning in Table Data through Operation Sketches and Self-Supervised Learning May 2026
Hanjun Cho, *Gahyun Yoo*, Hanseong Kim, Jay-Yoon Lee
(Transactions of the Association for Computational Linguistics)

Awards

EFL 8th Best Speaker, SNUDA X KNC 2021 – Korea Intervarsity Debate Association Jan 2021

Bronze Medal, Korean Mathematical Olympiad – Korean Mathematical Society Jan 2014

Presentations

- Invited Seminar @ Kakao Mobility (Feb 2025)

Extracurricular Activities

AttentionX: Member of Artificial Intelligence Research Group (2023)

Waffle Studio: Member of Seoul National University Web Development Club (2021-2023), developed the backend server of an interactive web editor application

SNUDA: Finance Minister and Member of Seoul National University Debate Club (2020-2021)

Skills

Languages: Python, C++, Java

ML Frameworks: PyTorch, SGLang, VLLM, Tensorflow, Scikit-Learn

Infrastructure: Docker

Coursework

Math: Linear Algebra, Calculus 1&2, Statistical Computing and Lab, Mathematical Analysis, Discrete Mathematics

Computer Science: Data Structures, Algorithms, Automata Theory, Logic Design, Computer Architecture, System Programming

Artificial Intelligence: Basics of Deep Learning, Machine Learning and Deep Learning for Data Science 1&2, Special Topic - Generative Modeling with Diffusion and Flow-Based Models