

Seanie Lee

DATE OF BIRTH: 1992.04.17, NATIONALITY: KOREAN

126 Yangjae-dong, Seocho District, Seoul

☎ (+82) 10-4475-2273 | ✉ lsnfamily02@kaist.ac.kr | 🏠 seanie12.github.io | 📺 seanie12 | 🎓 Seanie Lee

Education

KAIST (Korea Advanced Institute of Science and Technology)

PH.D IN ARTIFICIAL INTELLIGENCE

- Supervised by [Sung Ju Hwang](#) and [Juho Lee](#)
- Research interest: Safety Alignment and Efficiency of Large Language Models.

Daejeon, S.Korea

Mar. 2022 - present

KAIST (Korea Advanced Institute of Science and Technology)

M.S. IN ARTIFICIAL INTELLIGENCE

- Supervised by [Sung Ju Hwang](#) and [Juho Lee](#)
- Master Thesis: [Data augmentation for natural language processing](#)

Daejeon, S.Korea

Mar. 2020 - Feb. 2022

Yonsei University

B.A. IN LIBRARY AND INFORMATION SCIENCE

Seoul, S.Korea

Mar. 2011 - Feb. 2018

Experience

Mila

INTERNSHIP

- Research internship at Mila, advised by [Yoshua Bengio](#).

Montreal, Canada

January 2024 - June 2024

Apple

INTERNSHIP

- Research internship at Siri team, hosted by [Anders Johannsen](#).

Cambridge, UK

May 2023 - September 2023

Singapore National University

INTERNSHIP

- Remote internship at [Deep Learning](#) lab, supervised by [Kenji Kawaguchi](#).

Singapore

July 2022 - September 2022

Korea Advanced Institute of Science and Technology

TEACHING ASSISTANT

- Deep Reinforcement Learning, AI611
- Mathematics for AI, AI503
- Deep Learning, AI502

Daejeon, S.Korea

Mar. 2020 - Dec. 2021

42 Maru

INTERNSHIP

- Research on Question Answering, Semi-supervised Learning, Domain Generalization

Seoul, S.Korea

Feb. 2019 - Jan. 2020

Awards

2023 **Apple AI/ML PhD Fellowship**, Recipient of [Apple Scholars in AI/ML](#)

Cupertino, US

2022 **Google Travel Grant**, NeurIPS 2022

US

2019 **Silver Medal**, Named Entity Recognition in [NAVER NLP Challenge](#)

Seoul, Korea

Presentation

Seminar at Korea University.

PRESENTATION OF LARGE SCALE SET-ENCODING

- Synthetic Data Generation for LLM Safeguards
- ICLR 2025, ACL Findings 2025

Seoul, Korea

May. 2025

Seminar at Hanyang University.

PRESENTATION OF LARGE SCALE SET-ENCODING

- Synthetic Data Generation for LLM Safeguards
- ICLR 2025, ACL Findings 2025

Seoul, Korea

April. 2025

Tech. Talk, Nuremberg Institute of Technology Georg Simon Ohm.

PRESENTATION OF LARGE SCALE SET-ENCODING

- Scalable Set Encoding with Universal Mini-Batch Consistency and Unbiased Full Set Gradient Approximation
- ICML 2023

Nürnberg, Germany

Oct. 2023

Tech. talk, Samsung SDS.

PRESENTATION OF LARGE SCALE SET-ENCODING

- Scalable Set Encoding with Universal Mini-Batch Consistency and Unbiased Full Set Gradient Approximation
- ICML 2023

Seoul, South Korea

22.May. 2023

Tech. talk, NAVER corp.

PRESENTATION OF INFO-HCVAE

- Generating Diverse and Consistent QA pairs from Contexts with Information-Maximizing Hierarchical Conditional VAEs
- ACL 2020 Long paper

Online, South Korea

04.Dec. 2020

Publication

(* indicates equal contribution)

PREPRINT

FedSVD: Adaptive Orthogonalization for Private Federated Learning with LoRA

Arxiv

SEANIE LEE*, SANGWOO PARK*, DONG BOK LEE*, DOMINIK WAGNER, HAEBIN SEONG, TOBIAS BOCKLET, JUHO LEE, SUNG JU

2025

HWANG

- [\[paper\]](#)

Calibrated Decision-Making through LLM-Assisted Retrieval

Arxiv

CHAEYUN JANG, HYUNGI LEE, SEANIE LEE AND JUHO LEE

2024

- [\[paper\]](#)

CONFERENCES

SafeRoute: Adaptive Model Selection for Efficient and Accurate Safety Guardrails in Large Language Models

ACL Findings

SEANIE LEE*, DONG BOK LEE*, DOMINIK WAGNER, MINKI KANG, HAEBIN SEONG, TOBIAS BOCKLET, JUHO LEE, SUNG JU

2025

HWANG

- [\[paper\]](#)[\[code\]](#)

Personalized Fine-Tuning with Controllable Synthetic Speech from LLM-Generated Transcripts for Dysarthric Speech Recognition

Interspeech

DOMINIK WAGNER, ILJA BAUMANN, NATALIE ENGERT, SEANIE LEE, ELMAR NÖTH, KORBINIAN RIEDHAMMER AND TOBIAS

2025

BOCKLET

- [\[paper\]](#)

HarmAug: Effective Data Augmentation for Knowledge Distillation of Safety Guard Models

ICLR

SEANIE LEE*, HAEBIN SEONG*, DONG BOK LEE, MINKI KANG, XIAOYIN CHEN, DOMINIK WAGNER, YOSHUA BENGIO, JUHO LEE,

2025

SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

Learning Diverse Attacks on Large Language Models for Robust Red-teaming and Safety Tuning

ICLR

SEANIE LEE, MINSU KIM, LYNN CHERIF, DAVID DOBRE, JUHO LEE, SUNG JU HWANG, KENJI KAWAGUCHI, GAUTHIER GIDEL,

2025

YOSHUA BENGIO, NIKOLAY MALKIN, MOKSH JAIN

- [\[paper\]](#)[\[code\]](#)

Optimized Speculative Sampling for GPU Hardware Accelerators

EMNLP

DOMINIK WAGNER, SEANIE LEE, ILJA BAUMANN, PHILIPP SEEGER, KORBINIAN RIEDHAMMER, TOBIAS BOCKLET

2024

- [\[paper\]](#)[\[code\]](#)

Drug Discovery with Dynamic Goal-aware Fragment

SEUL LEE, **SEANIE LEE**, KENJI KAWAGUCHI, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICML

2024

Effective and Efficient Conversation Retrieval for Dialogue State Tracking with Implicit Text Summaries

SEANIE LEE, JIANPENG CHENG, JORIS DRIESEN, ALEXANDRU COCA, ANDERS JOHANSEN

- [\[paper\]](#)

NAACL

2024

Self-Supervised Dataset Distillation for Transfer Learning

DONG BOK LEE*, **SEANIE LEE***, JOONHO KO, KENJI KAWAGUCHI, JUHO LEE, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICLR

2024

DiffusionNAG: Task-guided Neural Architecture Generation with Diffusion Models

SOHYUN AHN*, HAYEON LEE*, JAEHYEONG JO, **SEANIE LEE**, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICLR

2024

Scalable Set Encoding with Universal Mini-Batch Consistency and Unbiased Full Set Gradient Approximation

JEFFREY WILLETTE*, **SEANIE LEE***, BRUNO ANDREIS, KENJI KAWAGUCHI, JUHO LEE, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICML

2023

Margin-based Neural Network Watermarking

BYUNGJOO KIM, SUYOUNG LEE, **SEANIE LEE**, SOOEL SON, SUNG JU HWANG

- [\[paper\]](#)

ICML

2023

Self-Supervised Set Representation Learning for Unsupervised Meta-Learning

DONG BOK LEE*, **SEANIE LEE***, KENJI KAWAGUCHI, YUNJI KIM, JIHWAN BANG, JUNG-WOO HA, SUNG JU HWANG

- [\[paper\]](#)

ICLR

2023

Self-Distillation for Further Pre-training of Transformers

SEANIE LEE, MINKI KANG, JUHO LEE, SUNG JU HWANG, KENJI KAWAGUCHI

- [\[paper\]](#)[\[code\]](#)

ICLR

2023

Set-based Meta-Interpolation for Few-Task Meta-Learning

SEANIE LEE*, BRUNO ANDREIS*, KENJI KAWAGUCHI, SUNG JU HWANG

- [\[paper\]](#) [\[code\]](#)

NeurIPS

2022

On Divergence Measures for Bayesian Pseudocoresets

BALHAE KIM, JUNGWON CHOI, **SEANIE LEE**, YOONHO LEE, JUNG-WOO HA, JUHO LEE

- [\[paper\]](#)

NeurIPS

2022

Set Based Stochastic Subsampling

BRUNO ANDREIS, **SEANIE LEE**, A. TUAN NGUYEN, JUHO LEE, EUNHO YANG, SUNG JU HWANG

- [\[paper\]](#)

ICML

2022

Sequential Reptile: Inter-Task Gradient Alignment for Multilingual Learning

SEANIE LEE*, HAE BEOM LEE*, JUHO LEE, SUNG JU HWANG

- [\[paper\]](#)

ICLR

2022

Learning to Perturb Word Embeddings for Out-of-distribution QA

SEANIE LEE*, MINKI KANG*, JUHO LEE, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ACL

2021

Contrastive Learning with Adversarial Perturbations for Conditional Text Generation

SEANIE LEE*, DONG BOK LEE*, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICLR

2021

Meta-GMVAE: Mixture of Gaussian VAE for Unsupervised Meta-Learning

DONG BOK LEE, DONGCHAN MIN, **SEANIE LEE**, SUNG JU HWANG

- [\[paper\]](#)[\[code\]](#)

ICLR

2021

Generating Diverse and Consistent QA pairs from Contexts with Information-Maximizing Hierarchical Conditional VAEs

DONG BOK LEE*, **SEANIE LEE***, WOOTAE JEONG, DONGHWAN KIM, SUNG JU HWANG

- [\[paper\]](#) [\[code\]](#)[\[video\]](#)

ACL

2020

References

Sung Ju Hwang ASSOCIATE PROFESSOR IN KAIST. e-mail: sjhwang82@kaist.ac.kr.	Advisor 2020-2025
Juho Lee ASSOCIATE PROFESSOR IN KAIST. e-mail: juholee@kaist.ac.kr.	Advisor 2020-2025
Yoshua Bengio FULL PROFESSOR AT UNIVERSITÉ DE MONTRÉAL AND SCIENTIFIC DIRECTOR OF MILA – QUEBEC AI INSTITUTE. e-mail: yoshua.bengio@mila.quebec.	Collaborator 2024-present
Kenji Kawaguchi PRESIDENTIAL YOUNG PROFESSOR IN THE DEPARTMENT OF COMPUTER SCIENCE AT NUS. e-mail: kenji@comp.nus.edu.sg	Collaborator 2022-present
Nikolay Malkin CHANCELLOR’S FELLOW AT UNIVERSITY OF EDINBURGH, SCHOOL OF INFORMATICS e-mail: nmalkin@ed.ac.uk	Collaborator 2024-present