

# Seanie Lee

REGULARIZATION · DATA AUGMENTATION

291 Daehak-ro, Yuseong-gu, Daejeon, Korea 34141

✉️ [lsnfamily02@kaist.ac.kr](mailto:lsnfamily02@kaist.ac.kr) | 🏠 [seanie12.github.io](https://seanie12.github.io) | 💻 [seanie12](#) | 🎓 [Seanie Lee](#)

## Education

### KAIST (Korea Advanced Institute of Science and Technology)

Daejeon, S.Korea

PH.D IN ARTIFICIAL INTELLIGENCE

Mar. 2022 -

- Supervised by [Sung Ju Hwang](#) and [Juho Lee](#)
- Research interest: Large Language Models, Safety Alignment, and Efficient Fine-tuning.

### KAIST (Korea Advanced Institute of Science and Technology)

Daejeon, S.Korea

M.S. IN ARTIFICIAL INTELLIGENCE

Mar. 2020 - Feb. 2022

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

### Yonsei University

Seoul, S.Korea

B.A. IN LIBRARY AND INFORMATION SCIENCE

Mar. 2011 - Feb. 2018

## Experience

### Mila

Montreal, Canada

INTERNSHIP

January 2024 - June 2024

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

### Apple

Cambridge, UK

INTERNSHIP

May 2023 - September 2023

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

### Singapore National University

Singapore

INTERNSHIP

July 2022 - September 2022

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

### Korea Advanced Institute of Science and Technology

Daejeon, S.Korea

TEACHING ASSISTANT

Mar. 2020 - Dec. 2021

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

### 42 Maru

Seoul, S.Korea

INTERNSHIP

Feb. 2019 - Jan. 2020

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

## Awards

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

Cupertino, US

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

Seoul, Korea

2019 **Bronze Medal**, Semantic Role Labeling in [NAVER NLP Challenge](#)

Seoul, Korea

## Presentation

### Tech. talk, Samsung SDS.

Seoul, South Korea

PRESENTATION OF LARGE SCALE SET-ENCODING

22.May. 2023

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

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

## Publication

---

### CONFERENCES

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

ICML

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

2024

- [\[paper\]](#)
- \*: equal contribution

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

NAACL

**SEANIE LEE**, JIANPENG CHEN, JORIS DRIESEN, ALEXANDRU COCA, ANDERS JOHANNSEN

2024

- [\[paper\]](#)
- \*: equal contribution

#### Self-Supervised Dataset Distillation for Transfer Learning

ICLR

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

2024

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution

#### DiffusionNAG: Task-guided Neural Architecture Generation with Diffusion Models

ICLR

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

2024

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution

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

ICML

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

2023

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution

#### Margin-based Neural Network Watermarking

ICML

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

2023

- [\[paper\]](#)

#### Self-Supervised Set Representation Learning for Unsupervised Meta-Learning

ICLR

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

2023

- [\[paper\]](#)
- \*: equal contribution

#### Self-Distillation for Further Pre-training of Transformers

ICLR

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

2023

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

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

NeurIPS

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

2022

- [\[paper\]](#) [\[code\]](#)
- \*: equal contribution

#### On Divergence Measures for Bayesian Pseudocoresets

NeurIPS

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

2022

- [\[paper\]](#)

#### Set Based Stochastic Subsampling

ICML

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

2022

- [\[paper\]](#)

#### Sequential Reptile: Inter-Task Gradient Alignment for Multilingual Learning

ICLR

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

2022

- [\[paper\]](#)
- \*: equal contribution

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

ACL

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

2021

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution

## Contrastive Learning with Adversarial Perturbations for Conditional Text Generation

ICLR

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

2021

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution

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

ICLR

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

2021

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

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

ACL

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

2020

- [\[paper\]](#) [\[code\]](#)[\[video\]](#)
- \*: equal contribution

## g2pM: A Neural Grapheme-to-Phoneme Conversion Package for Mandarin Chinese Based on a New Open Benchmark Dataset

INTERSPEECH

KYUBYONG PARK\*, SEANIE LEE\*

2020

- [\[paper\]](#)[\[code\]](#)
- \*: equal contribution