

Hyunji Amy Lee

hyunji.amy.lee@kaist.ac.kr | <https://amy-hyunji.github.io/>

RESEARCH INTEREST

Natural Language Processing and Machine Learning: information retrieval, knowledge representation, knowledge-augmented LMs, and reasonings.

My research goal is to build a robust semi-parametric model that aligns with external knowledge modules thereby capitalizing on the strengths of both the model's parametric knowledge space and external nonparametric knowledge space, aiming to achieve the optimal performance. I've focused on superposing parametric and nonparametric knowledge spaces [C4][C8], exploring retrieval models to find relevant external knowledge [C2][P2][C9][W1], and developing language models that ground well to external knowledge [C4][C7]. Additionally, I aim to extend these solutions to multi-modal challenges by merging representation spaces from various modalities [P1][C1][C3].

EDUCATION

KAIST AI

Ph.D. Candidate in AI Grad School. Advisor: [Minjoon Seo](#) Mar. 2023 – present

KAIST AI

M.S. in AI Grad School. Advisor: [Minjoon Seo](#) Mar. 2021 – Feb. 2023

KAIST

B.S. in Electrical Engineering & Computer Science (Double Major) Mar. 2016 – Feb. 2021

RESEARCH EXPERIENCE

Tencent AI Lab

Research Intern, Tencent AI Lab

Mentor: Wenhao Yu

Seattle, US

Sep 2024 ~ Present

Adobe Research

Research Intern, Vision Language Lab (Mentor: David Seunghyun Yoon, Trung Bui, Franck Dernoncourt)

Improving Question Answering over Real-World Corpus

San Jose, US

May 2024 ~ Aug 2024

Allen Institute for AI (AI2)

Research Intern, Semantic Scholar Team (Mentor: Kyle Lo, Luca Soldani, Arman Cohan)

Improving Out-of-Domain Dense Retriever

Seattle, US

May.2023 - Dec. 2023

[P2]

MindLogic

Research Scientist Intern (Mentor: YongWoo Kim)

Developing a real-time korean QA Chatbot model

Seoul, Republic of Korea

Dec. 2020– Feb. 2020

Naver

Research Scientist Intern (Mentor: Sohee Yang, Wonseok Hwang, Minjoon Seo)

Developing a cost-effective parser model for semi-structured document images

Seongnam, Republic of Korea

Mar. 2020 – Dec. 2020

[C1]

SKT AI

AI Fellowship (Mentor: Eric Davis, Jaehoon Kim)

Building Efficient Koran Multi Sentiment Analysis Model ([page](#))

Seoul, Republic of Korea

June. 2019 – June. 2020

MindLogic

Research Scientist Intern (Mentor: YongWoo Kim)

Developing Virtual Try-On Model ([page](#))

Seoul, Republic of Korea

Dec. 2019 – Feb. 2019

Last Updated: Oct. 4, 2024

PUBLICATIONS

Conference and Journal Papers

[C9] Continually Updating Generative Retrieval on Dynamic Corpora

Soyoung Yoon, Chaeun Kim, Hyunji Lee, Joel Jang, Minjoon Seo
EMNLP 2024

[C8] SemiParametric Token-Sequence Co-Supervision

Hyunji Lee*, Doyoung Kim*, Jihoon Jun, Sejune Joo, Joel Jang, Kyung-Woon On, Minjoon Seo
Accepted in ACL 2024 (Long)

[C7] How Well Do Large Language Models Truly Ground?

Hyunji Lee*, Sejune Lee*, Chaeun Kim, Joel Jang, Kyung-Woon On, Minjoon Seo
Accepted to NACCL 2024 (Long)
ACL KnowledgeNLP Workshop 2024 Oral (Best Student Paper)

[C6] KTRL+F: Knowledge-Augmented In-Document Search

Hanseok Oh*, Haebin Shin*, Miyoung Ko, Hyunji Lee, Minjoon Seo
Accepted to NACCL 2024 (Long)
ACL KnowledgeNLP Workshop 2024, ACL KnowledgeLM Workshop 2024

[C5] Improving Probability-based Prompt Selection Through Unified Evaluation and Analysis

Sohee Yang, Jonghyeon Kim, Joel Jang, Seonghyeon Ye, Hyunji Lee, Minjoon Seo
Accepted to TACL 2024

[C4] Nonparametric Decoding for Generative Retrieval

Hyunji Lee, Jaeyoung Kim, Hyeon Chang, Hanseok Oh, Sohee Yang, Vlad Karpukhin, Yi Lu, Minjoon Seo
Accepted to Findings of ACL 2023 (Long)

[C3] Local 3D Editing via 3D Distillation of CLIP Knowledge

Junha Hyung, Sungwon Hwang, Daejin Kim, Hyunji Lee, Jaegul Choo
Accepted to CVPR 2023

[C2] Generative Multi-hop Retrieval

Hyunji Lee, Sohee Yang, Hanseok Oh, Minjoon Seo
Accepted to EMNLP 2022 (Long), AKBC 2022
Finalist Qualcomm Innovation Fellowship 2023

[C1] Cost-Effective End-to-End Information Extraction for Semi-Structured Document Images

Wonseok Hwang, Hyunji Lee, Jinyeong Yim, Geewook Kim, Minjoon seo
Accepted to EMNLP 2021 (Short)

Posters and Workshop Papers

[W1] InstructIR: A Benchmark for Instruction Following of Information Retrieval Models

Hanseok Oh, Hyunji Lee, Seonghyeon Ye, Haebin Shin, Hansol Jang, Changwook Jun, Minjoon Seo
ACL KnowledgeNLP Workshop 2024 Oral

Preprints / Under Review

[P2] Knowledge Entropy Decay during Language Model Pretraining Hinders New Knowledge Acquisition

Jiyeon Kim*, Hyunji Lee*, Hyowon Cho, Joel Jang, Hyeonbin Hwang, Seungpil Won, Youbin Ahn, Dohaeng Lee, Minjoon Seo

Preprint arxiv

[P2] RouterRetriever: Exploring the Benefits of Routing over Multiple Expert Embedding Models

Hyunji Lee, Luca Soldani, Arman Cohan, Minjoon Seo, Kyle Lo

Preprint arxiv

[P2] Back to Basics: A Simple Recipe for Improving Out-of-Domain Retrieval in Dense Encoders

Hyunji Lee, Luca Soldani, Arman Cohan, Minjoon Seo, Kyle Lo

Preprint arxiv

[P1] Zero-Shot Dense Video Captioning by Jointly Optimizing Text and Moment

Yongrae Jo, Seongyun Lee, Aiden SJ Lee, Hyunji Lee, Hanseok Oh, Minjoon Seo

Preprint arxiv

HONORS, AWARDS, AND SCHOLARSHIPS

Best Student Paper

ACL KnowledgeNLP Workshop 2024

2024

Qualcomm Innovation Fellowship

Qualcomm, Finalist

2023

KAIST Graduation Leadership Awards

KAIST

2021

Qualcomm-KAIST Multimodal Emotion Recognition Kaggle

Qualcomm, 5th Prize

2020

SKT AI Fellowship

SKT AI

2019

DKU (Duke & Duke-Kunshan Univ.) program as KAIST representative

Duke & Duke-Kunshan Univ.

2018

KAIST Global Leadership Awards

KAIST

2018

KAIST Leadership & Volunteering Awards

KAIST

2017