

Eunsu Kim

✉ kes0317@kaist.ac.kr | 🏠 eunsu-k1m.github.io/

Summary

I'm a Master student at KAIST School of Computing, advised by Alice Oh. My current research focuses on the evaluation of LLMs, including but not limited to, (i) What to evaluate: Exploring the direction in which LLMs should progress and examining current LLM behavior from that perspective and (ii) How to evaluate: Developing evaluation frameworks that measure the true capabilities of LLMs.

Research Interests : Natural Language Processing, LLM, LLM evaluation

Education

Korea Advanced Institute of Science & Technology

M.S. Computer Science

- Advisor : Alice Oh

Dajeon, Korea

Aug 2023–present

Korea Advanced Institute of Science & Technology

B.S. Electronic and Electrical engineering

- GPA 4.02/4.3, Summa Cum Laude
- Major GPA 4.15/4.3

Dajeon, Korea

Feb 2019–Aug 2023

Publications

* denotes equal contributions

Uncovering Factor Level Preferences to Improve Human-Model Alignment

{Eunsu Kim, Juhyun Oh}* , Jiseon Kim, Wenda Xu, Inha Cha, William Yang Wang, Alice Oh

Preprint Under Review, 2024

BLEnD: A Benchmark for LLMs on Everyday Knowledge in Diverse Cultures and Languages

Junho Myung*, Nayeon Lee*, Yi Zhou*, Jiho Jin, Rifki Afina Putri, Dimosthenis Antypas, Hsuvas Borkakoty, Eunsu Kim, Carla Perez Almendros, Abinew Ali Ayele, Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Hwaran Lee, Sham-suddeen Hassan Muhammad, Kiwoong Park, Anar Sabuhi Rzayev, Nina White, Seid Muhie Yimam, Mohammad Taher Pilehvar, Nedjma Ousidhoum, Jose Camacho-Collados, Alice Oh

Neurips D&B 2024

CLiCK: A Benchmark Dataset of Cultural and Linguistic Intelligence in Korean

Eunsu Kim, Juyoung Suk, Philhoon Oh, Haneul Yoo, James Thorne, Alice Oh

LREC-COLING, 2024

Multi-FACT: Assessing Factuality of Multilingual LLMs using FActScore

Sheikh Shafayat, Eunsu Kim*, Juhyun Oh*, Alice Oh

COLM 2024

The Generative AI Paradox on Evaluation: What It Can Solve, It May Not Evaluate

{Eunsu Kim, Juhyun Oh, Inha Cha}* , Alice Oh

EACL SRW 2024

Work In Progress

LLM-AS-AN-INTERVIEWER: Beyond Static Testing Through Dynamic LLM Evaluation

Eunsu Kim, Juyoung Suk, Seungone Kim, Niklas Muennighoff, Dongkwan Kim, Alice Oh

Work in Progress

Domestic Publications

Korean Named Entity Recognition using D-Tag

Eunsu Kim, Sujong Do, Cheoneum Park

Proceedings of the 34th Annual Conference on Human and Cognitive Language Technology, 2022

Teaching Experience

Teaching Assistant

AI Ethics

KAIST

Spring 2024

Counseling Assistant

Leader of Counseling Assistants in spring 2024

KAIST CS

Fall 2023 - Spring 2024

Tutor

General Physics 1

KAIST

Mar 2022 - Dec 2022

Tutor

Math & Science

KAIST Center for Gifted Education

Apr 2021 - Dec 2024

Tutor

AI, Education Contents Production

KAIST Center for Gifted Education

Sep 2023 - Dec 2023

Work Experience

Hyundai motors company, NLP team

Internship

Seoul, Korea

Mar,2022 – Aug,2022

- Joint model of Named Entity Recognition and Intent Classification using BERT via soft prompt tuning
- Automatically building dictionary of unlabeled NAMUWIKI data

Talks

Hanbat National University

Evaluating Large Language Models: A Perspective of Building Evaluation Benchmark Datasets

Dajeon, Korea

Jun 2024

Research Experience

KAIST NICA lab

Individual Study

Dajeon, Korea

Dec 2020 - Dec 2021

- Unmixing proteins in fluorescence cell images
- Advisor :Prof.Youngyu Yoon, KAIST electrical engineering

KAIST BISPL lab

Individual Study

Dajeon, Korea

Sep,2022 – Dec,2022

- Detecting disease of MRI lung images using Visual-language model via prompt-based method
- Advisor :Prof.Jongchul Ye, KAIST Kim Jaechul Graduate School of AI

KAIST ALINlab

Individual Study

Dajeon, Korea

Jan,2023 – Mar,2023

- Solving Tabular data classification problems using Large Language Models
- Advisor :Prof.Jinwoo Shin, KAIST electrical engineering

KAIST NMSL lab

Research Intern

Dajeon, Korea

Jan,2023 – Mar,2023

- Developing Test-Time Adaptation method that works on Transformers
- Advisor :Prof.Sungju Lee, KAIST electrical engineering

The Many Voices of ChatGPT

Research Intern

Dajeon, Korea

Mar,2023–Aug, 2023

- Exploring the Potential of ChatGPT for Korean : A Case Study of College Entrance and Proficiency Exams
- Advisor :Prof.Alice Oh and Prof.James Thorne,

Skills

Programming Python, PyTorch, Matlab, C

Scholarships & Awards

Next Generation Water Action 2022

Technical University of Denmark

1st prize of Ulsan challenge

2022

KAIST Urban Grand Challenges 2022

KAIST Civil & Environmental Engineering

Excellence Prize

2022

National Science and Technology scholarship

Korea Student Aid Foundation

2020–2022