

# Lee, Jung Hyun

Address: NAVER 1784, 95 Jeongjail-ro, Bundang-gu, Seongnam, South Korea (13561)

E-mail: onliwad101@gmail.com / Personal website: <https://onliwad101.github.io>

## EDUCATION

---

**Korea Advanced Institute of Science and Technology (KAIST)**, Sep. 2019 – Aug. 2021 Daejeon, South Korea

**Master of Science in the Graduate School of AI (advisor: Prof. Eunho Yang)**

- GPA: 3.92 / 4.3 (96.2 / 100)
- Thesis: Cluster-Promoting Quantization with Bit-Drop for Minimizing Network Quantization Loss

**Pohang University of Science and Technology (POSTECH)**, Mar. 2011 – Feb. 2019 Pohang, South Korea

**Bachelor of Science in Mathematics, minor in Industrial and Management Engineering**

- GPA: 3.82 / 4.3 (95.2 / 100) - *Magna Cum Laude*
- Leave of absence for mandatory military service (Jan. 2013 – Oct. 2014)

## PUBLICATIONS & ACADEMIC PAPERS (\* Equal Contribution)

---

### Preprints

[8] **Jung Hyun Lee\***, June Yong Yang\*, Byeongho Heo, Dongyoon Han, Kang Min Yoo. Token-Supervised Value Models for Enhancing Mathematical Reasoning Capabilities of Large Language Models. In Preparation.

[7] **Jung Hyun Lee\***, Jeonghoon Kim\*, June Yong Yang, Se Jung Kwon, Eunho Yang, Kang Min Yoo, and Dongsoo Lee. LRQ: Optimizing Post-Training Quantization for Large Language Models by Learning Low-Rank Weight-Scaling Matrices. Under Review.

[6] HyperCLOVA X Team. HyperCLOVA X Technical Report. Preprint.

### Peer-reviewed Articles

[5] Byeonghu Na, Yeongmin Kim, HeeSun Bae, **Jung Hyun Lee**, Se Jung Kwon, Wanmo Kang, Il-chul Moon. Label-Noise Robust Diffusion Models. International Conference on Learning Representations (ICLR), 2024.

[4] Jeonghoon Kim\*, **Jung Hyun Lee\***, Sungdong Kim, Joonsuk Park, Kang Min Yoo, Se Jung Kwon, and Dongsoo Lee. Memory-Efficient Fine-Tuning of Compressed Large Language Models via sub-4-bit Integer Quantization. Neural Information Processing Systems (NeurIPS), 2023.

[3] **Jung Hyun Lee\***, Jeonghoon Kim\*, Se Jung Kwon, and Dongsoo Lee. FlexRound: Learnable Rounding based on Element-wise Division for Post-Training Quantization. International Conference on Machine Learning (ICML), 2023.

[2] Kyung-su Kim\*, **Jung Hyun Lee\***, and Eunho Yang. Compressed Sensing via Measurement-Conditional Generative Models. *IEEE Access*, 2021.

[1] **Jung Hyun Lee\***, Jihun Yun\*, Sung Ju Hwang, and Eunho Yang. Cluster-Promoting Quantization with Bit-Drop for Minimizing Network Quantization Loss. IEEE/CVF International Conference on Computer Vision (ICCV), 2021.

## RESEARCH & WORK EXPERIENCE

---

**NAVER Cloud**, Mar. 2022 – Present

Seongnam, South Korea

### Research Scientist, Foundation Research Team

- Developed a new post-training weight-rounding mechanism, FlexRound [3] that can flexibly quantize pre-trained weights of not only computer vision models but also language models including Llama, based on the magnitude of each weight
- Introduced PEQA [4], a method that fine-tunes only the quantization step sizes of quantized LLMs to (i) reduce both the model size and the number of training parameters during fine-tuning, and (ii) accelerate inference latency after fine-tuning
- Proposed a new post-training weight quantization method for LLMs, LRQ [7] that learns low-rank weight-scaling matrices instead of dense ones to decrease learnable parameters, thus enhancing the generalization capability of quantized LLMs

- Improved and evaluated the mathematical capabilities of HyperCLOVA X [6], a family of Korean-specialized LLMs
- Presented token-supervised value model (TVMs) [8], new token-level verifiers trained to estimate the probability of reaching the correct final answer for each token in a solution

**Samsung Research**, Jul. 2021 – Mar. 2022

Seoul, South Korea

**Software Engineer, Data Research Team**

- Had programming training in algorithms and data structures as a newly-hired employee and successfully completed the training course by earning its own programming certification
- Analyzed customers' buying behavior patterns, such as purchase frequency, time and occasion; sorted out loyal customers and recommended brand-new electronic products to them

**Machine Learning and Intelligence Laboratory, KAIST**, Apr. 2019 – Aug. 2019

Daejeon, South Korea

**Research Intern (advisor: Prof. Eunho Yang)**

- Conducted preliminary research into the impact of neural network pruning on the interpretability of neural networks via Layer-wise Relevance Propagation
- Implemented recent algorithms proposed in deep learning and machine learning papers, reproduced the experimental results, and brainstormed how to improve those algorithms for performance enhancement

## ACADEMIC SERVICES

---

**Conference Reviewer:** NeurIPS (2022-2024), ICLR (2024-2025), ICML (2024), ACL Rolling Review (2024), AAAI (2025)

## HONORS & AWARDS

---

- **TOP 2 in the research track at the N INNOVATION AWARD 2023**, an internal excellence in technology awards ceremony hosted by NAVER
- **National Scholarship for Science and Engineering** from Korea Student Aid Foundation in 2011, which covered full tuition and included an additional stipend

## EXTRACURRICULAR ACTIVITIES

---

- Served as a mentor at POSTECH by helping freshmen's adaptation and teaching calculus and applied linear algebra in 2016
- Acted as a captain and playing a coach for POSTECH baseball club; won the 2<sup>nd</sup> prize in the university competition in 2012

## OTHER INFORMATION

---

- TOEFL IBT score 98 (Reading: 27, Listening: 23, Speaking: 21, Writing: 27)