Taehyeon Kim

Seoul, Korea kimtaehyeon610@gmail.com (permanent); potter32@kaist.ac.kr
taehyeon-k-6a1239207 https://Kthyeon.com https://github.com/Kthyeon

INTERESTS

- **Thinking/action alignment**: Training sycophancy-free models where internal reasoning genuinely drives external behaviour; inference-time and training-time solutions.
- **Self-improving training loops**: Eval-harness design, feedback wiring, and rapid model iteration via RFT/RL pipelines (RPO, DPO, AdaSTaR-style self-taught data generation).
- **Agentic AI**: Multi-step planning, reliable action-taking, policy-aligned agents — from research prototypes to production systems.
- **LLM inference**: Speculative decoding, vLLM/SGLang optimization, serving-cost reduction for production agent deployment.

EDUCATION

Ph.D., Graduate School of AI
Korea Advanced Institute of Science and Technology (KAIST) - Advisor: Prof. Se-Young Yun *Mar 2020 - Feb 2025*
- Thesis: *Efficient and Effective Inference-time Decoding Strategies for Foundational Models*

M.S., Graduate School of Data Science
Korea Advanced Institute of Science and Technology (KAIST) - Advisor: Prof. Se-Young Yun *Mar 2018 - Feb 2020*
- Thesis: *Orthogonal Feature Regularization: A Novel Approach for Training Robust Models*

B.S., Mathematical Sciences & Minor, Intellectual Property Minor Program *Mar 2013 - Feb 2018*
Korea Advanced Institute of Science and Technology (KAIST)

WORK EXPERIENCES

LG AI Research @ Seoul, South Korea *Apr 2025 – Present*
- **Tech lead:** Expert-domain AI search system — end-to-end pipeline from retrieval to agentic reasoning.
- **Agentic Orchestration:** Built multi-tool harness with parallel retrieval, LLM query decomposer/planner, and structured examination modeling (novelty & inventiveness judgment), wired via tool-use endpoints.
- **LLM alignment & RL:** SFT, DPO, GRPO-based training (RPO — adaptive thinking via root-token policy, ACL 2026 [C15]); novel DPO for small-large model collaboration; SIMPER for math/coding (in preparation).
- **Research:** Contrastive thinking decoding; sycophancy-free inference-time alignment via vLLM extension (ICML under review [U4]); adaptive self-taught data generation for RFT, co-corresponding author (NeurIPS 2025 [C12]); LLM negotiation-agent benchmark MERIT (ACL 2026 [C13]); multi-drafter speculative decoding (ACL 2026 Findings [C14]).
- **Retrieval & Reranking:** Trained domain dense retrievers (text-embedding) with cross-encoder rerankers; contrastive learning (hard-negative mining, curriculum sampling).
- **Multi-modal:** Designed VLM fine-tuning experiments for diagram understanding; evaluated cross-modal fusion (image+text) strategies for retrieval grounding.

Google Research @ NYC, USA *Oct 2023 – Dec 2023*
- **PhD Intern:** Fast parallel decoding techniques on large language models (published to NeurIPS24 [C10])
- *Working with: Adrian Benton, Ananda Theertha Suresh, and Kishore Papineni*

Dynamo AI (YC-backed) @ SF, USA *Jan 2023 – Aug 2023*
- **Full research contributor (6+ months):** Privacy-preserving federated learning for enterprise AI; contributed to closing B2B contracts with automotive clients while simultaneously publishing a first-author NeurIPS 2023 paper [C5] (*with Eric Lin*).

National Institute of Meteorological Sciences (NIMS) @ South Korea *Mar 2022 – Jun 2023*
- **Research Manager:** Transformer-based precipitation forecasting; NeurIPS 2022 Competition award (40+ researchers).

Qualcomm @ Seoul, South Korea *Jun 2021 – Dec 2021*
- **PhD Intern:** CV & ML research for autonomous drivingreport (*Working with: Heesoo Myeong*)
- Designing resource-efficient and accurate backbone for ADAS (published at ICML22W [W3], U.S. patent [P1]).

PUBLICATIONS

C15 **Kim, T.**, Lee, H., Jang, Y. & Lee, M. (2026). Efficiently Learning To Reason or Not to Reason: Root-token Policy Optimization for Adaptive Thinking, The 64th Annual Meeting of the Association for Computational Linguistics (***ACL 2026***) (**Oral**), Jul. 2026.

C14 **Kim, T.***, Jung, D.* & Yun, S.. (2026). Multi-Drafter Speculative Decoding with Alignment Feedback, The 64th Annual Meeting of the Association for Computational Linguistics (***ACL 2026 Findings***), Jul. 2026.

C13 Oh, J., Aghazada, M., Shin, Y., Yun, S. & **Kim, T.**† (2026). MERIT Feedback Elicits Better Bargaining in LLM Negotiators, The 64th Annual Meeting of the Association for Computational Linguistics (***ACL 2026***), Jul. 2026.

W9 Kim, Y., Yi, E., Kim, M., Yun, S.†, **Kim, T.**† (2025). Guiding Reasoning in Small Language Models with LLM Assistance, ***CoLM 2025*** Workshop: Test-time Scaling and Reasoning Models.

C12 Koh, W., Oh, W., Jang, J., Lee, M., Kim, H., Kim, A., Kim, J., Lee, J., **Kim, T.**†, Yun, S.† (2025). AdaSTaR: Adaptive Data Sampling for Training Self-Taught Reasoners, *Advances in Neural Information Processing Systems 39 (**NeurIPS 2025**)*, Dec. 2025.

C11 Koh, W., Yoon, J., Lee, M., Song, Y., Cho, J., Kang, J., **Kim, T.**, Yun, S., Yu, Y., Lee, B. (2025). C^2: Scalable Auto-Feedback for LLM-based Chart Generation, *The 2025 Annual Conference of the Nations of the Americas Chapter of the ACL (**NAACL 2025**)* (**Oral**), May. 2025.

J2 **Kim, T.**, Kim, D. & Yun, S.. (2024). FLR: Label-Mixture Regularization for Federated Learning with Noisy Labels, Transactions on Machine Learning Research (**TMLR**).

W8 **Kim, T.***, Jung, D.* & Yun, S.. (2024). A Unified Framework for Speculative Decoding with Multiple Drafters as a Bandit, NeurIPS 2024 Workshop: Efficient Natural Language and Speech Processing (ENLSP-IV).

C10 **Kim, T.**, Suresh, AT., Papineni, K., Riley, M., Kumar, S. & Benton, A., (2024).Exploring & Improving Multi-token Prediction (Block Draft) in Language Modeling, *Advances in Neural Information Processing Systems 38 (**NeurIPS 2024**)*, Dec. 2024.

C9 Ho, N.*, Bae, S.*, **Kim, T.**, Jo, H., Kim, Y., Schuster, T., Fisch, A., Thorne, J. and Yun, S. (2024). Block Transformer: Global-to-Local Language Modeling for Fast Inference, *Advances in Neural Information Processing Systems 38 (**NeurIPS 2024**)*, Dec. 2024.

C8 Yi, E.*, **Kim, T.***, Jeung, H., Chang, D. and Yun, S. (2024). Towards Fast Multilingual LLM Inference: Speculative Decoding and Specialized Drafters, *The 2024 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2024**)*, Nov. 2024.

W7 **Kim, T.**, Suresh, AT., Papineni, K., Riley, M., Kumar, S. & Benton, A., (2024). Exploring and Improving Drafts in Blockwise Parallel Decoding, ICML 2024 Workshop: Efficient Systems for Foundation Models (Es-FoMo).

C7 **Kim, T.***, Kim, J.*, Lee, G.* & Yun, S.. (2024). Instructive Decoding: Instruction-Tuned Large Language Models are Self-Refiner from Noisy Instructions, *Twelfth International Conference on Learning Representations (**ICLR 2024**)* **(Spotlight, Top 5%)**, May. 2024.

C6 Yang, Y., **Kim, T.** & Yun, S.. (2024). Leveraging Normalization Layer in Adapters with Progressive Learning and Adaptive Distillation for Cross-Domain Few-Shot Learning, *The 38th Annual AAAI Conference on Artificial Intelligence (**AAAI 2024**)*, Feb. 2024.

C5 **Kim, T.**, Lin, E., Lee, J., Lau, C. & Muguntahn, V.. (2023). Navigating Data Heterogeneity in Federated Learning: A Semi-Supervised Federated Object Detection, *Advances in Neural Information Processing Systems 37 (**NeurIPS 2023**)*, Dec. 2023.

C4 Gruca et al. Weather4cast at NeurIPS 2022: Super-Resolution Rain Movie Prediction under Spatio-temporal Shifts, *Proceedings of the **NeurIPS 2022 Competition and Demonstration Track**, in Proceedings of Machine Learning Research 220:292-313 (2022)*.

W6 **Kim, T.**, Kang, S., Shin, H., Yoon, D., Eom, S., Shin, K. & Yun, S.. (2022). Region-Conditioned Orthogonal 3D U-Net for Weather4Cast Competition, NeurIPS 2022 Workshop: Weather4Cast Competition.

W5 Eom, S., **Kim, T.** & Yun, S.. (2022). Layover Intermediate Layer for Multi-Label Classification in Efficient Transfer Learning, NeurIPS 2022 Workshop: Has it Trained Yet? (HITY).

W4 Shin, J., **Kim, T.** & Yun, S.. (2022). Revisiting the Activation Function for Federated Image Classification, NeurIPS 2022 Workshop Federated Learning.

W3 **Kim, T.** & Yun, S.. (2022). Supernet Training for Federated Image Classification under System Heterogeneity, ICML 2022 Workshop on Dynamic Neural Networks (**Oral**).

J1 **Kim, T.** & Yun, S.. (2022). Revisiting Orthogonality Regularization: A Study for Convolutional Neural Networks in Image Classification, ***IEEE Access***, Jun. 2022.

W2 **Kim, T.**, Myeong, H. & Yun, S.. (2022). Revisiting Architecture-aware Knowledge Distillation: Smaller Models and Faster Search, ICML 2022 Hardware Aware Efficient Training (HAET) Workshop, July 2022.

C3 **Kim, T.***, Ko, J.*, Cho, S., Choi, J. & Yun, S.. (2021). FINE Samples for Learning with Noisy Labels., *Advances in Neural Information Processing Systems 35 (**NeurIPS 2021**)*, Dec. 2021.

C2 **Kim, T.***, Oh, J.*, Kim, N., Cho, S., & Yun, S. Y. (2021). Comparing Kullback-Leibler Divergence and Mean Squared Error Loss in Knowledge Distillation. *In the 30th International Joint Conference on Artificial Intelligence (**IJCAI 2021**)*, Aug. 2021 (acceptance rate: 13.9%)

W1 **Kim, T.***, Ahn, J.*, Kim, N.*, & Yun, S. (2020). Adaptive Local Bayesian Optimization Over Multiple Discrete Variables. Workshop at NeurIPS 2020 Competition Track on Black-Box Optimization Challenge, Dec. 2020.

C1 **Kim, T.**, Kim, J. & Yun, S. (2019). Efficient Model for Image Classification With Regularization Tricks. *Proceedings of the **NeurIPS 2019 Competition and Demonstration Track**, in Proceedings of Machine Learning Research 123:13-26*.

T **Kim, T.**. Orthogonal feature regularization : a novel approach for training robust models, Korea Advanced Institute of Science and Technology (KAIST)

UNDER REVIEW PAPERS

U5 **Kim, T.** et al. (2025). Masked-SIMPER: Improving Collaborative Reasoning, In preparation.

U4 **Kim, T.** et al. (2025). Contrastive Thinking Decoding, *International Conference on Machine Learning (**ICML 2026**)*, Under review.

U1 Bartholet, M.*, **Kim, T.***, Beuret, A. Yun, S. & Buhmann, J. (2024). Hypernetwork-Driven Model Fusion for Federated Domain Generalization, Under review.

US PATENTS

P1 **Kim, T.**, Myeon, H. (2023). Trust-Region Aware Architecture Distillation for Sample-Efficient Neural Architecture Search, QualComm Inc..

AWARDS @ COMPETITIONS

4th Award in NeurIPS 2022 Competition: Weather4Cast *Dec 2022*
- Homepage: https://www.iarai.ac.at/weather4cast/
- Subjects: Precipitation Forecast, Segmentation, Video

8th Award in NeurIPS 2020 Black-Box Optimization Challenge *Dec 2020*
- BBO-challenge homepage: https://bbochallenge.com/
- Subjects: Auto-ML, Bayesian Learning, Hyperparameter Optimization.

2nd & 3rd Awards in NeurIPS 2019 MicroNet Challenge, CIFAR-100 Track. *Dec 2019*
- MicroNet-challenge homepage: https://micronet-challenge.github.io/
- Subjects: Image Classification, Model Compression.

ACHIEVEMENTS

Google Conference Scholarship *Dec 2024*
Exploring & Improving Multi-token Prediction (Block Draft) in Language Modeling

Google Cloud Platform (GCP) Credit Award *Mar 2024*
Gemma Academic Program

Winner, 2022 Qualcomm AI Fellowship *Oct 2022*
FINE Samples for Learning with Noisy Labels.

Best Poster Awards, KAIST AI 21/22 Workshop *Jan 2022*
1. FINE Samples for Learning with Noisy Labels.
2. Comparing Kullback-Leibler Divergence and Mean Squared Error Loss in Knowledge Distillation

LEADERSHIP

Tech Lead & Research Manager @ LG AI Research *Apr 2025 – Present*
- Leading a team of 10 researchers (5 PhDs) across retrieval, LLM alignment, and agentic systems.
- Defining research roadmap, mentoring junior researchers, and managing cross-team project delivery.

1st Representative of doctoral students @ KAIST AI *Mar 2021 – Feb 2022*
- Construct organizations
- Being an intermediator between the professors and the students

Lab master @ OSI LAB *Mar 2020 – Feb 2021*
- Optimization and Statistical Inference Laboratory (OSI LAB), KAIST.
- Construct organizations (e.g., seminar, lab policy, funding)
- Being an intermediator between the advisor and the students

Vice Captain @ KAIST Representative Cheerleading Group *Oct 2014 – Nov 2015*
- A.K.A. Encouraging Leaders of KAIST (ELKA)
- KAIST festival planning and promotion
- 2015 Gwangju Summer Universiade - University Student U Cheering Festival **1st prize**
- Team building, Funding, Management.

INVITED TALKS

Naver AI Lab Seminar *Jan 2025*
- LLM Inference-Time Strategies and its Beyond.

Microsoft Research Seminar *Jan 2025*
- Instructive Decoding and its Beyond.

Squeezebits Efficient AI Seminar *Jun 2024*
- Towards Fast LLM Inference: Speculative Decoding and Efficient Architectures

KCC 2024 *Jun 2024*
- Navigating Data Heterogeneity in Federated Learning: A Semi-Supervised Federated Object Detection

HyperConnect Seminar *Jan 2023*
- Personalized Federated Learning on System Heterogeneity under Label Noise

KSC 2022 *Dec 2022*
- FINE Samples for Learning with Noisy Labels

Naver Clova AI Seminar *Jun 2019*
Orthogonal Feature Regularization: A Novel Approach for training robust model

SKILLS & OTHERS

Reviewer **Conference**: NeurIPS {2022, 2023, 2024}, ICML {2022, 2023, 2024}, AutoML-Conf 2022, ECCV 2022, KDD 2023, ICLR {2024, 2025}, ARR 2024.
Journal: Pattern Recognition.

Coding Python, PyTorch, LaTeX, ...

AI in Art Surplus human, Vania Oh, Junbeom Shin, **Taehyeon Kim**, 2022.
- AI with Weird Wonderland, 22.03.10 - 22.03.16, CELINE PARK GALLERY, Organized by Next interface lab.

TEACHING EXPERIENCES

AI Lecturer @ ELICE, Mathematics in ML *Jan 2022 – Feb 2022*
- Teaching the mathematics for gradient descent in ML (4 lectures, theory and code quiz, exams)
- Attendee: LG Employees.

AI Mentor @ Curinc, Introduction to Deep Learning *Jun 2021 – Jul 2021*
- Teaching the overview of deep learning in Curinc, Seoul, Korea (70 hours during 7 weeks)
- Attendee: Undergraduate at UC Berkeley, Boston University, and Florida International University.

TA @ LG, Computer Vision & Deep Learning *Oct 2020 – Nov 2020*
- Attendee: LG Employees in the LG AI CAMP Module(3), LG, Academy, Korea
- Image classification and semantic segmentation using public COVID dataset in Kaggle.

Advisory Committee @ National Science Museum, Deep Learning *Mar 2020 – Oct 2020*
- AI Exploration Program, National Science Museum, Korea.
- Attendee: Advanced students (Science High School)
- Subjects: reinforcement learning, object detection, image classification, evolutionary algorithm.

TA @ KAIST Data Science, Introduction to Deep Learning *Mar 2019 – Jul 2019*
- Dept. Knowledge Service Engineering, KAIST, Korea.
- Attendee: Graudate Students

TA @ LG, Computer Vision & Deep Learning *Jan 2019 – Jan 2019*
- Theories and practices for deep learning in the LG AIB Intermediate CAMP, LG Academy, Korea.

Lecturer @ Samsung, Python Basics *Jan 2019 – Jan 2019*
- Python code implementation in the Samsung SW Academy Start CAMP, Samsung SW Academy, Korea.
- Subjects: Chatbot, basic python (e.g., for loop, condition)

Lecturer @ KAIST, Regular Leadership Course *Mar 2017 – Dec 2017*
- KAIST Lecture: Personality/Leadership 3 - Liberal Arts Required (2017 Spring & 2017 Fall)
- Subjects: Sports leadership, Team leadership.
- Students: undergraduate freshman in KAIST

Scholarship & Fellowship

Graduate School of AI, KAIST Funding *Mar 2020 – Present*
$ 20,000 per year

National scholarship for graduate studies, Korea Student Aid Foundation *Mar 2018 – Feb 2020*
$ 16,000 per year

Full Academic scholarship, KAIST *Mar 2013 – Aug 2017*
$ 3,000 per year