Seunghyuk Oh

Phone: (+82) 50-6768-1768 · Email: seunghyukoh0@gmail.com
Google Scholar   LinkedIn   Github

EXPERTISE & INTERESTS

Optimizing Large Language Models in the test time; Efficient long context modeling; Efficient decoding, speculative decoding; Test-time computation, verification & refinement; Tabular learning

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST)   Seoul, South Korea
M.S. Student in Graduate School of AI; Advisor: Prof. Jinwoo Shin   *Expected Aug 2025*
B.S. in Electrical Engineering (Minor: Computer Science)   *Aug 2023*

PROJECTS

ReVISE, learn to self-verify and self-refine   *Dec 2024 – Feb 2025*
- Suggested a framework that enables LLMs to refine reasoning at test time through intrinsic self-verification.

Mamba drafters for speculative decoding   *Aug 2024 – Feb 2025*
*Collaborated with **Amazon Science***
- Explored Mamba's efficacy as a fast, and efficient drafter for speculative decoding.

HOMER, hierarchical context merging for long context modeling   *Aug 2023 – Sep 2023*
*Collaborated with **NAVER***
- Proposed a training-free, efficient long-context modeling technique leveraging a divide-and-conquer strategy.

WORK EXPERIENCE

Machine Learning Engineer Intern at ONEUNIVERSE   *Jan 2023 – Jul 2023*
- Developed a persona-controllable chatbot system to enhance user engagement and customization.
- Implemented an LLM-based evaluator for assessing the quality of the conversation.

Co-Founder & CTO at Weebut   *Jan 2021 – Dec 2022*
- Developed and operated iN!T; recruitment solution based on senior developer's evaluations.
- Built a lung cancer diagnostic solution, MEDIAR, to segment and classify lesions in whole slide images.

PUBLICATIONS

C: Conference, J: Journal, P: Pre-print, *: Equal contribution

[P2] D. Choi, **Seunghyuk Oh**, S. Dingliwal, J. Tack, K. Kim, W. Song, S. Kim, I. Han, J. Shin, A. Galstyan, S. Katiyar, S. Bodapati, "Mamba Drafters for Speculative Decoding," Pre-print.

[P1] H. Lee*, **Seunghyuk Oh***, J. Kim, J. Shin, J. Tack, "ReVISE: Learning to Refine at Test-Time via Intrinsic Self-Verification," Pre-print.

[J1] W. Song, J. Tack, S. Mo, **Seunghyuk Oh**, J. Shin, "Sparsified State-Space Models are Efficient Highway Networks," **Transactions on Machine Learning Research (TMLR)** 2025.

[C2] J. Nam*, K. Kim*, **Seunghyuk Oh**, J. Tack, J. Kim, J, Shin, "Optimized Feature Generation for Tabular Data via LLMs with Decision Tree Reasoning," **Neural Information Processing Systems (NeurIPS)** 2024.

[C1] W. Song*, **Seunghyuk Oh***, S. Mo, J. Kim, S. Yun, J. Ha, J. Shin, "Hierarchical Context Merging: Better Long Context Understanding for Pre-trained LLMs," **International Conference on Learning Representations (ICLR)** 2024.