HYUHNG JOON KIM

1, Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea, 08826
https://heyjoonkim.github.io ⋄ heyjoonkim@europa.snu.ac.kr

RESEARCH INTERESTS

Natural Language Processing (NLP) based on Machine Learning and Deep Learning. Specifically interested in:

- Reliable models through improving **abstention abilities** (reducing hallucinations).
- **Uncertainty quantification** of language generations.
- Understanding and advancing **Large Language Models (LLMs)**.

EDUCATION

Seoul National University *Sep. 2020 -*
M.S. & Ph.D. in Artificial Intelligence (integrated) (Advisor: Sang-goo Lee)

Soongsil University *Mar. 2014 - Aug. 2020*
B.S. in Computer Science and Engineering

PUBLICATIONS

International Conferences

1. **Hyuhng Joon Kim**, Youna Kim, Sang-goo Lee, and Taeuk Kim, When to Speak, When to Abstain: Contrastive Decoding with Abstention, *The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL 2025**)*
2. **Hyuhng Joon Kim**, Youna Kim, Cheonbok Park, Junyeob Kim, Choonghyun Park, Kang Min Yoo, Sang-goo Lee, and Taeuk Kim, Aligning Language Models to Explicitly Handle Ambiguity, *The 2024 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2024**)*.
3. Youna Kim, **Hyuhng Joon Kim**, Cheonbok Park, Choonghyun Park, Hyunsoo Cho, Junyeob Kim, Kang Min Yoo, Sang-goo Lee, and Taeuk Kim, Adaptive Contrastive Decoding in Retrieval-Augmented Generation for Handling Noisy Contexts, *Findings of the Association for Computational Linguistics: EMNLP 2024 (**Findings of EMNLP 2024**)*.
4. **Hyuhng Joon Kim**, Hyunsoo Cho, Sang-Woo Lee, Junyeob Kim, Choonghyun Park, Sang-goo Lee, Kang Min Yoo, and Taeuk Kim, Universal Domain Adaptation for Robust Handling of Distributional Shifts in NLP, *Findings of the Association for Computational Linguistics: EMNLP 2023 (**Findings of EMNLP 2023**)*.
5. Hyunsoo Cho, **Hyuhng Joon Kim**, Junyeob Kim, Sang-Woo Lee, Sang-goo Lee, Kang Min Yoo, and Taeuk Kim, Prompt-Augmented Linear Probing: Scaling Beyond The Limit of Few-shot In-Context Learners, *The 37th AAAI Conference on Artificial Intelligence (**AAAI 2023**)*.
6. Hyunsoo Cho, Choonghyun Park, Junyeop Kim, **Hyuhng Joon Kim**, Kang Min Yoo, Sang-goo Lee, Probing Out-of-Distribution Robustness of Language Models with Parameter-Efficient Transfer Learning, *The 12th Joint Conference on Lexical and Computational Semantics (***SEM 2023**)*.
7. Kang Min Yoo, Junyeob Kim, **Hyuhng Joon Kim**, Hyunsoo Cho, Hwiyeol Jo, Sang-Woo Lee, Sanggoo Lee, and Taeuk Kim, Ground-Truth Labels Matter: A Deeper Look into Input-Label Demonstrations, *The 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2022**)*.

International Workshops

1. **Hyuhng Joon Kim**, Hyunsoo Cho, Junyeob Kim, Taeuk Kim, Kang Min Yoo, Sang-goo Lee, Self-generated in-context learning: Leveraging auto-regressive language models as a demonstration generator, *Workshop on Large-scale Pre-trained Language Models (**LPLM 2022**) at NAACL 2022*.

Preprint (arXiv)

1. Choonghyun Park, **Hyuhng Joon Kim**, Junyeob Kim, Youna Kim, Taeuk Kim, Hyunsoo Cho, Hwiyeol Jo, Sang-goo Lee, Kang Min Yoo, Investigating the Influence of Prompt-Specific Shortcuts in AI Generated Text Detection, *arxiv:2406.16275*
2. Youna Kim, **Hyuhng Joon Kim**, Minjoon Choi, Sungmin Cho, Hyunsoo Cho, Sang-goo Lee, Taeuk Kim, UniKnow: A Unified Framework for Reliable Language Model Behavior across Parametric and External Knowledge, *arXiv:2502.13648v2*

EXPERIENCES

Recommender System — *Seoul National University*
Teaching Assistant — *Spring 2023*

Database — *Seoul National University*
Teaching Assistant — *Fall 2022*

Software Developer — *Tomato System*
Development Intern — *2020*

Software Developer — *Naver Corp.*
Development Intern — *2019*

SKILLS

- Programming Languages: Python, C, JAVA, Javascript, SQL
- Deep Learning Libraries: PyTorch, Transformers
- Tools: Git, Vim
- Languages: Korean (Native), English (Intermediate)

EXTRA INFORMATION

- Personal Page: `https://heyjoonkim.github.io/`
- Github: `https://github.com/heyjoonkim`
- Google Scholar: `https://scholar.google.com/citations?user=NHFFB4gAAAAJ&hl=en`
- Semantic Scholar: `https://www.semanticscholar.org/author/Hyuhng-Joon-Kim/2166352356`