

Hyunsoo Cho



RESEARCH INTERESTS

Machine Learning

- Anomaly Detection, Distributional Shift, Self-Supervised Learning, Unsupervised Learning

Natural Language Processing

- Foundation Model, Prompting, Weakly-Supervised Learning, Parameter-efficient Transfer Learning, Black-box Language Model

EDUCATION

- **Ph.D.** Computer Science and Engineering, Seoul National University Mar. 2019 — Feb. 2023
 - **Dissertation:** *A Deep Representation Learning for Unsupervised Anomaly Detection.*
 - Outstanding Ph.D. Dissertation Award
 - **Advisor:** Sang-goo Lee
 - **Committee:** U Kang, Sang-goo Lee, Seung-won Hwang, Sungzoon Cho (SNU), Taeuk Kim (HYU)
- **M.S.** Computer Science and Engineering, Seoul National University Sep. 2016 — Aug. 2018
 - **Dissertation:** *Visual Question Answering using Simple Natural Language Processing.*
 - **Advisor:** Sang-goo Lee
- **B.S.** Computer Science and Engineering, Chung-ang University Mar. 2012 — Aug. 2016

WORK EXPERIENCE

Naver AI Lab.

- Visiting Researcher Apr. — Aug. 2022, Mar. 2023 — Current

RESEARCH EXPERIENCE

HYPERCLOVA X: The Next Generation Korean Hyperscale Conversational Language Model

Visiting Researcher, *Naver Corp.*

Mar. 2023 — Current

- Synthetic instruction data generation
- Enhancing language model's coding capability
- Code capability evaluation

Enhancing Natural Language Understanding via Pre-trained Language Models

Project Manager, with *Naver Corp.*

Apr. 2021 — Mar. 2023

- Prompt optimization.
- Parameter-efficient transfer learning.
- Unsupervised anomaly detection.

An Advanced Study of QA Technologies for Intelligent Assistant System

Project Manager, with *Hyundai Motors.*

Mar. 2020 — Feb. 2021

- Advancing vehicle-embedded intention classifier.
- Detecting anomalous user request from spoken language.

Text and Sentiment Analysis for Fashion Trend Analysis

Project Manager, with *Korea Creative Content Agency (KOCCA).*

Jul. 2019 — Jun. 2020

- Analyzing sentiment in fashion postings through meta information.
- Extracting prominent keywords and entities from fashion postings.

PUBLICATIONS

Pre-print & Under review

- [1] **Exploring the Influence of Factual Correctness on Instruction Fine-tuning Foundational Language Models**
Hyunsoo Cho, Bruce W. Lee, Kang Min Yoo, Sang-goo Lee
Ongoing research
- [2] **Elevating the Knowledge Depth Harnessing Human Cognitive Pathways**
Hyunsoo Cho*, Bruce W. Lee*, Kang Min Yoo.
Ongoing research
- [3] **Integrating Two Perspectives on Distributional Shift in Natural Language Understanding.**
Hyuhng Joon Kim, **Hyunsoo Cho**, Sang-Woo Lee, Junyeob Kim, Choonghyun Park, Sang-goo Lee, Kang Min Yoo, Taeuk Kim.
Under review

International

- [1] **Probing Out-of-Distribution Robustness of Language Models with Parameter-Efficient Transfer Learning.**
Hyunsoo Cho, Choonghyun Park, Junyeop Kim, Hyuhng Joon Kim, Kang Min Yoo, Sang-goo Lee.
*The 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023).*
- [2] **CELDA: Leveraging Black-box Language Model as Enhanced Classifier without Labels.**
Hyunsoo Cho, Youna Kim, Sang-goo Lee.
The 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023).
- [3] **Prompt-Augmented Linear Probing: Scaling Beyond The Limit of Few-shot In-Context Learners.**
Hyunsoo Cho, Hyuhng Joon Kim, Junyeob Kim, Sang-Woo Lee, Sang-goo Lee, Kang Min Yoo, Taeuk Kim.
Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2023).
- [4] **Ground-Truth Labels Matter: A Deeper Look into Input-Label Demonstrations.**
Junyeob Kim, Kang Min Yoo, Hyuhng Joon Kim, **Hyunsoo Cho**, Hwiyeol Jo, Sang-Woo Lee, Sang-goo Lee, Taeuk Kim.
The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP Oral 2022).
- [5] **Enhancing Out-of-Distribution Detection in Natural Language Understanding via Implicit Layer Ensemble.**
Hyunsoo Cho, Choonghyun Park, Jaewook Kang, Kang Min Yoo, Taeuk Kim, Sang-goo Lee.
Findings of the Association for Computational Linguistics: EMNLP 2022 (Findings of EMNLP 2022).
- [6] **Self-Generated In-Context Learning: Leveraging Auto-regressive Language Models as a Demonstration Generator.**
Hyuhng Joon Kim, **Hyunsoo Cho**, Junyeob Kim, Taeuk Kim, Kang Min Yoo, Sang-goo Lee
Workshop on Large-scale Pre-trained Language Models (*LPLM 2022*) at NAACL 2022.
- [7] **Masked Contrastive Learning for Anomaly Detection.**
Hyunsoo Cho, Jinseok Seol, Sang-goo Lee.
The 30th International Joint Conference on Artificial Intelligence (IJCAI 2021).
- [8] **Leveraging Class Hierarchy in Fashion Classification.**
Hyunsoo Cho, Chaemin Ahn, Kang Min Yoo, Jinseok Seol, Sang-goo Lee.
The 2nd Workshop on Computer Vision for Fashion, Art and Design 2019 (CVFAD 2019) at ICCV 2019.
- [9] **Stochastic Relational Network.**
Kang Min Yoo, **Hyunsoo Cho**, Hanbit Lee, Jeeseung Han, Sang-goo Lee.
The 1st Workshop on Statistical Deep Learning in Computer Vision 2019 (SDLCV 2019) at ICCV 2019.
- [10] **Automatic Generation of Multiple-Choice Fill-in-the-blank Question Using Document Embedding.**
Junghyuk Park, **Hyunsoo Cho**, and Sang-goo Lee.
The 19th International Conference on Artificial Intelligence in Education (AIED 2018).

Domestic

- [1] **FastText를 적용한 한국어 단어 임베딩.**
Hyunsoo Cho, Sang-goo Lee.
Korea Software Congress (KSC 2017).

PATENTS

- Anomaly Detection System and Method for Setting Threshold Thereof (international - America, China, Korea)

VISITING TALKS

- Invited speaker - Tech talk, *Naver Enterprise* Mar. 2022
- Invited speaker - Tech talk, *Kakao Enterprise* Jan. 2022
- Invited speaker - Tech talk, *Hyperconnect Enterprise* Jan. 2022
- *AIIS Fall Retreat* Nov. 2021
- Invited speaker - *The 2nd Workshop on Artificial Intelligence for Anomalies and Novelty, AI4AN 2021 at IJCAI 2021* Aug. 2021
- Invited speaker - 2018 토대연구 학술회의 <이보형 民俗樂 Odyssey> at *Korean Music Society* May. 2018

ACADEMIC SERVICES

Reviewer:

- The Annual Meeting of the Association for Computational Linguistics (**ACL**): 2022, 2023
- The North American Chapter of the Association for Computational Linguistics (**NAACL**): 2022
- The International Conference On Computational Linguistics (**COLING**): 2022
- The Conference on Empirical Methods in Natural Language Processing (**EMNLP**): 2022, 2023
- The Association for the Advancement of Artificial Intelligence (**AAAI**): 2023, 2024
- The Joint Conference on Lexical and Computational Semantics (***SEM**): 2023

AWARDS & HONORS

- **Outstanding Ph.D. Dissertation Award** - *Seoul National University* Feb. 2023
- **Best Paper Award** - *AIIS Fall Retreat* Nov. 2021
- **Academic Excellence Scholarship** - *Chung-ang University* Feb. 2015

TEACHING EXPERIENCE

- **Advanced Database** Fall. 2019
 - *Department of Computer Science & Engineering, SNU - TA*
- **Introduction to Databases** Fall. 2017
 - *Department of Computer Science & Engineering, SNU - TA*
- **IT Entrepreneurship** Fall. 2017
 - *Department of Computer Science & Engineering, SNU - TA*
- **Advanced Computer Science Seminar** Spring. 2017
 - *Department of Computer Science & Engineering, SNU - TA*

(Last update. September 2, 2023)