# JANGHOON HAN

hanjh04@naver.com
hanjanghoon.github.io

## RESEARCH INTEREST

**Natural language processing**

        Instruction tuning
        Crosslingual transfer
        End-to-End dialogue system
        Post-training, Selection, and Continual Learning

## EDUCATION

| | | |
|---|---|---|
| Mar. 2019 ~ Feb. 2021 | **Sogang University** | Seoul, Korea |

*Master of Science in Computer Science and Engineering*
*Natural Language Processing Lab, advised by Jungyun Seo*
Thesis: Single-turn Post-training for Multi-turn Retrieval-based Dialogue Systems
*(CGPA 4.19/4.3)*

| | | |
|---|---|---|
| Mar. 2013 ~ Feb. 2019 | **Sogang University** | Seoul, Korea |

Department of Computer Science and Engineering
*Bachelor of Science in Computer Science and Engineering*
*ABEEK (Accreditation Board for Engineering Education of Korea)*
*Graduated* Cum laude *(CGPA 3.41/4.3, Major GPA 3.45/4.3 )*

## PUBLICATIONS

- C. Lee\*, **J. Han\*,** S. Ye, SJ Choi, H. Lee, K. Bae, "Instruction Matters, a Simple yet Effective Task Selection Approach in Instruction Tuning for Specific Tasks"**, EMNLP 2024**

- T. Hong, Y. ahn, D. Lee, J. Shin, S. Won, J. Han, SJ. Choi, J. Seo, "Exploring the Use of Natural Language Descriptions of Intents for Large Language Models in Zero-shot Intent Classification", SIGDIAL 2024

- **J. Han**, D. Lee, J. Shin, H. Bae, J.Bang, S. Kim, SJ. Choi, H. Lee, "Efficient Dynamic Hard Negative Sampling for Dialogue Selection", NLP4ConvAI @ ACL2024

- **J. Han\*,** C. Lee\*, J. Shin, SJ Choi, H. Lee, K. Bae, "Deep Exploration of Cross-Lingual Zero-Shot Generalization in Instruction Tuning", **Findings of ACL 2024**

- S. Joo, K. Lee, K. Min, J. Shin, J. Han, S. Won, K. Jung, "Leveraging Ensemble Techniques and Metadata for Subjective Knowledge-grounded Conversational Systems"**,** DSTC11 @ SIGDIAL 2023 **Best paper**

- D. Lee, G. Kim**,** J. Han, T. Hong, YR. Kim, SJ. Choi, NL. Zhang, "Local Temperature Beam Search: Avoid Neural Text DeGeneration via Enhanced Calibration"**, Findings of ACL 2023**

- S. Won, H. Kwak, J. Shin, J. Han, K. Jung, "BREAK: Breaking the dialogue state tracking barrier with beam search and re-ranking", **ACL 2023**

- T. Hong, D. Lee, J. Han, J. Choi, J. Seo, "CECADA: Cause-Effect Conjunctive Adverb-based Data Augmentation Method in Low-Resource Knowledge-Grounded Dialogue", Information Sciences, Knowledge Augmented Methods for NLP @ AAAI 2023

- J. Jang, S. Ye, C Lee, S. Yang, J. Shin, J. Han, G. Kim, M. Seo, "TemporalWiki: A Lifelong Benchmark for Training and Evaluation Ever-Evolving Language Models", **EMNLP 2022,** Spa-NLP @ ACL 2022 (poster)**.**

- J. Jang, S. Ye, S. Yang, J. Shin, J. Han, G. Kim, Stanley. J. Choi, M. Seo, "Towards Continual Knowledge Learning of Language Models"**, ICLR 2022 (poster),** Spa-NLP @ ACL 2022 (oral).

- **J. Han**, J. Shin, H. Song, H. Jo, G. Kim Y. Kim, Stanley. J. Choi, "External Knowledge Selection with Weighted

Negative Sampling in Knowledge-grounded Task-oriented Dialogue Systems" DSTC10 @ AAAI 2022.

- J. Shin, G. Kim, H. Jo, H. Song, J. Han, Y. Kim, Stanley. J. Choi, "A Method of Automatic Data Construction for Unseen Values in Multi-domain Dialogue State Tracking" DSTC10 @ AAAI 2022.

- **J. Han**, Y. Ko and J. Seo, "External Knowledge Incorporated BERT for Retrieval-based Dialogue Systems", KIISE Transactions on Computing Practices (KCTP) (2021).

- **J. Han**, T. Hong, B. Kim, Y. Ko and J. Seo, **"**Fine-grained Post-training for Improving Retrieval-based Dialogue Systems", **NAACL 2021.**

- **J. Han**, Y. Park, Y Jeong, I. Lee, J.-U. Han, S. Park, J. Kim and J. Seo, "Korean Dependency Parsing Using Sequential Parsing Method Based on Pointer Network" Proceedings of the 31st Annual Conference on Human and Cognitive Language Technology, pp. 533-536, 2019.

- C. Kim, S. Kim, J. Park, H. Son, J. Han and J. Kim, "Embedded-based Word Recognition Method for Unattended Order Payment Systems for the Visually Impaired", Proceedings of the Korean Information Science Society Conference, pp. 1762-1764, 2018.

## WORK EXPERIENCE

### Industry Experience

Researcher, Language Lab, LG AI Research (Mar 2021~ Present)

- Conversational Large Language Model Development – ChatEXAONE (2023~Present)
  - ✓ Conducting Preliminary Research on Robustness to Diverse Instructions (Present)
  - ✓ Technical Lead within the Conversational LLM Team for Instruction-tuned Model Enhancement (2024)
  - ✓ Independently Conducted the Prototyping and Initial Validation of Instruction-tuned Models (2023)
- Multi-task prompt tuning for English-Korean large-scale bi-lingual language model (2022)
  - ✓ 100+ Korean nlp tasks, 5000+ Korean prompt for multi-task prompt tuning
  - ✓ Optimizing Seen, Unseen task during training
- Advancement of knowledge selection and ranking modules in dialogue systems (2022)
  - ✓ Tilda (virtual ai human)'s persona and knowledge conversation.
  - ✓ FAQ based knowledge grounded conversation for AI banker.
- Development of Korea super large-scale language model (2021)
  - ✓ Korean morph bpe tokenizer, Model parallism in lingvo lib, Light-weight tuning and tuning solution.

Researcher, Natural Language Processing Lab, Sogang University (Mar 2019~ Feb. 2021)

- Development of data augmentation through Multi-modal information (2019)
  - ✓ Development of keyword extraction system (Text Rank)
- Automatic creation of text data labels for machine learning and development of verification tools (2019)
  - ✓ Development of Korean Dependency Parser

Intern, Smart Factory AI Big data team, LG CNS (June 2018~ Aug. 2018)

- Deduction of key factors for production performance through clustering

### Teaching Experience

Instructor, Industrial Professionals Training Course, National IT Industry Promotion Agency of Korea (Dec. 2020)

- Machine learning for constructing Smart factories

Teaching Assistant, Department of Computer Science and Engineering, Sogang University (Mar 2019~ Feb. 2020)

- C-Programming
- Computer Science Engineering Lab 2 (Best Teaching Assistant Award)

## AWARDS AND HONORS

- 2019 Korean Information Processing System Competition: Korean Dependency Parsing (2019. 10)
  - ✓ **Grand prize** (**Minister of Culture, Sports and Tourism Commendation**)

- **Grand prize**, LG CNS Internship Workshop Start Up Monster (2018)
- Academic Excellence Scholarships, Sogang University (2017)

## PATENTS

- J. Han, Y. Park, Y Jeong, I. Lee and J. Seo, "System and method for dependent parsing"

## SKILLS

- Python, C, JAVA
- Pytorch, Distributed Training (Deepspeed, Ray, SLURM), High performance Inference (vLLM, TensorRT)

## EXTRACURRICULAR ACTIVITIES

- Vocalist, *ABYSS*, Hip hop club (Mar. 2013 ~ Feb. 2015)
- Ski club, *WALCS* (Mar. 2014 ~ Feb. 2015)