

JANGHOON HAN

hanjh9439@gmail.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

- [1] Changho Lee*, **Janghoon Han***, Seonghyeon Ye, Stanley Jungkyu Choi, Honglak Lee, Kyunghoon Bae, “Instruction Matters: A Simple yet Effective Task Selection for Optimized Instruction Tuning of Specific Tasks”, **EMNLP 2024**.
- [2] Taesuk Hong, Youbin Ahn, Dongkyu Lee, Joongbo Shin, Seungpil Won, **Janghoon Han**, Stanley Jungkyu Choi, Jungyun Seo, “Exploring the Use of Natural Language Descriptions of Intents for Large Language Models in Zero-shot Intent Classification”, SIGDIAL 2024.
- [3] **Janghoon Han**, Dongkyu Lee, Joongbo Shin, Hyunkyung Bae, Jeessoo Bang, Seonghwan Kim, Stanley Jungkyu Choi, Honglak Lee, “Efficient Dynamic Hard Negative Sampling for Dialogue Selection”, NLP4ConvAI @ ACL 2024.
- [4] **Janghoon Han***, Changho Lee*, Joongbo Shin, Stanley Jungkyu Choi, Honglak Lee, Kyunghoon Bae, “Deep Exploration of Cross-Lingual Zero-Shot Generalization in Instruction Tuning”, **Findings of ACL 2024**.
- [5] LG AI Research, "Chat EXAONE: Enterprise AI Agent," Technical Report, August 2024. Available: https://www.lgresearch.ai/data/upload/tech_report/ko/Technical_report_ChatEXAONE.pdf.
- [6] Seongho Joo, Kang-il Lee, Kyungmin Min, Joongbo Shin, **Janghoon Han**, Seungpil Won, Kyomin Jung, “Leveraging Ensemble Techniques and Metadata for Subjective Knowledge-grounded Conversational Systems”, DSTC11 @ SIGDIAL 2023 **Best Paper**.
- [7] Dongkyu Lee, Gyeonghun Kim, **Janghoon Han**, Taesuk Hong, Yi-Reun Kim, Stanley Jungkyu Choi, Nevin L Zhang, “Local Temperature Beam Search: Avoid Neural Text DeGeneration via Enhanced Calibration”, **Findings of ACL 2023**.
- [8] Seungpil Won, Heeyoung Kwak, Joongbo Shin, Janghoon Han, Kyomin Jung, “BREAK: Breaking the dialogue state tracking barrier with beam search and re-ranking”, **ACL 2023**.
- [9] Taesuk Hong, Dongkyu Lee, **Janghoon Han**, Stanley Jungkyu Choi, Jungyun Seo, “CECADA: Cause-Effect Conjunctive Adverb-based Data Augmentation Method in Low-Resource Knowledge-Grounded Dialogue”, Knowledge Augmented Methods for NLP @ AAAI 2023.
- [10] Joel Jang, Seonghyeon Ye, Changho Lee, Sohee Yang, Joongbo Shin, **Janghoon Han**, Gyeonghun Kim, Minjoon Seo, “TemporalWiki: A Lifelong Benchmark for Training and Evaluation Ever-Evolving Language Models”, **EMNLP 2022**.

- [11] Joel Jang, Seonghyeon Ye, Sohee Yang, Joongbo Shin, **Janghoon Han**, Gyeonghun Kim, Stanley Jungkyu Choi, Minjoon Seo, “Towards Continual Knowledge Learning of Language Models”, **ICLR 2022**.
- [12] **Janghoon Han**, Joongbo Shin, Hosung Song, Hyunjik Jo, Gyeonghun Kim, Yireun Kim, Stanley Jungkyu Choi, “External Knowledge Selection with Weighted Negative Sampling in Knowledge-grounded Task-oriented Dialogue Systems”, DSTC10 @ AAI 2022.
- [13] **Janghoon Han**, Youngjoong Ko, Jungyun Seo, “External Knowledge Incorporated BERT for Retrieval-based Dialogue Systems”, KIISE Transactions on Computing Practices (KCTP) (2021).
- [14] **Janghoon Han**, Taesuk Hong, Byoungjae Kim, Youngjoong Ko, Jungyun Seo, “Fine-grained Post-training for Improving Retrieval-based Dialogue Systems”, **NAAACL 2021**.
- [15] Janghoon Han, Yeongjoon Park, Younghoon Jeong, Inkwon Lee, Jungwook Han, Seojun Park, Juae Kim, Jeongyeon Seo, “Korean Dependency Parsing Using Sequential Parsing Method Based on Pointer Network”, Human and Cognitive Language Technology 2019.

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

Professional Activities

Reviewer, ACL Rolling Review (ARR), (Oct 2023 – Present)

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)